

# Le SDS vers une stratégie de données en cloud hybride

## Table des matières

Table des matières.....	1
Introduction.....	1
Qu'est-ce que SDS ?.....	1
Dans quels scénarios utiliser le SDS ? .....	2
Indicateurs clés de performance (KPI) pour une implémentation réussie du SDS .....	3
La structure SDS .....	3
Le SDS, porte d'entrée du cloud hybride .....	4
IBM Spectrum Storage Suite.....	4
Exemple : Utilisation du stockage cloud en tant que « palier » .....	5
Exemple : Protection des données sur des supports multiples .....	5
Exemple : Créez votre propre stockage cloud et devenez votre propre fournisseur de services .....	5
Conclusion .....	5

## Introduction

Dans ce numéro d'Insight, Neuralytix analyse les caractéristiques du stockage défini par logiciel (Software-Defined-Storage : SDS) et son impact sur l'économie des données. Cet article examine en profondeur une solution SDS, IBM Spectrum Software Suite, et les réponses qu'elle apporte aux besoins métier de l'entreprise moderne.

Neuralytix estime qu'en 2020, 30 % des nouveaux projets informatiques seront basés sur une infrastructure convergée avec des logiciels SDS. Nous sommes convaincus des bienfaits des solutions SDS et pensons que toutes les entreprises doivent purement et simplement inclure le SDS dans leur stratégie IT.

## Qu'est-ce que le SDS ?

Une idée fausse souvent répandue sur le SDS est qu'il s'agit d'un produit individuel, c'est-à-dire un logiciel qui permet d'accéder aux fonctionnalités SDS. Dans une certaine mesure, cette définition est correcte. Le SDS doit éviter à l'entreprise de devoir utiliser une multitude de solutions individuelles provenant souvent de plusieurs fournisseurs.

L'étude de Neuralytix laisse à penser qu'il existe un stockage cloud officiel (ou fantôme) dans 100 % des entreprises.

En outre, selon nos estimations, environ 95 % de toutes les grandes entreprises recherchent une stratégie de stockage hybride comprenant un stockage sur cloud. Environ 80 % des entreprises de taille moyenne souhaitent quant à elles trouver une solution permettant d'intégrer du stockage cloud à leur stockage local.

Dans les deux cas, nos estimations indiquent qu'environ 55 % des entreprises cherchent activement à implémenter une stratégie de SDS dans les 12 prochains mois, et que 25 % envisagent de le faire dans les 36 mois à venir.

Alors même que les entreprises se tournent vers le cloud et deviennent moins dépendantes du stockage physique dédié, le SDS apporte une flexibilité et une cohérence qui lui permettent de s'adapter à différents modèles de déploiement, tout en offrant une expérience d'exploitation homogène. Cependant, la réalité est que le SDS est plutôt une structure, comprenant de nombreux éléments et dépendances.

Ses fonctionnalités sont multiples :

- 🌸 Virtualisation du stockage local et sur le cloud.
- 🌸 Orchestration de toutes les ressources disponibles, y compris les ressources cloud.
- 🌸 Système de fichiers évolutif extensible sur le cloud.
- 🌸 Services de données, par exemple présentation du stockage dans différents protocoles. Exemple : iSCSI, S3, HDFS, NFS, SMB, etc.
- 🌸 Hiérarchisation des données dans le stockage local et cloud.
- 🌸 Intégration des données locales et des données cloud.
- 🌸 Qualité de service.

Connecter toutes ces fonctions pour créer une solution cohérente, qui sera perçue comme un logiciel unique, nécessite un fournisseur possédant une bonne expérience de tous les aspects de la structure SDS.

Une façon simple de résumer le SDS et ses possibilités est de le voir comme une solution de virtualisation des données. A la base, le SDS « libère » les données de l'infrastructure matérielle. L'entreprise peut ainsi se concentrer sur la valeur de l'information pouvant être extraite des données.

La Figure 1 illustre les interactions de base entre le SDS et l'infrastructure matérielle qui l'environne :

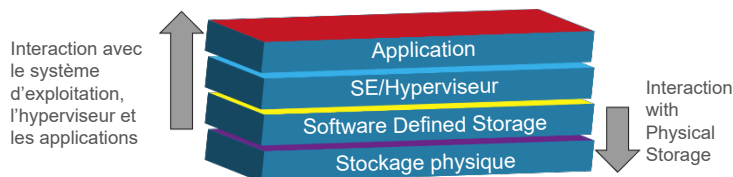


Figure 1 : Schéma fonctionnel des interactions du SDS (Neuralytix 2016)

Un élément fondamental que l'on retrouve dans toutes les solutions SDS est qu'il peut par le biais d'une seule interface gérer toutes les interactions entre la ou les applications et le système d'exploitation ou l'hyperviseur, ainsi que la capacité du stockage physique.<sup>1</sup>

La gestion de données s'en trouve ainsi incroyablement simplifiée. L'administrateur n'a plus besoin d'associer individuellement les applications aux différents systèmes de stockage. L'avantage est que cela évite la création de silos de données et une utilisation inefficace de la capacité globale du stockage. En outre, dans le cas des logiciels SDS sophistiqués, l'analytique rend possible un placement des données permettant le respect des contrats de niveaux de service (SLA). Deux exemples sont le nombre d'IOPS garanties et la suppression de la redondance des données (via par exemple la protection EC : erasure coding).

En d'autres termes, l'informatique cesse d'être une technologie de l'infrastructure pour devenir une vraie technologie de l'information. Pour l'entreprise, les résultats sont clairs : croissance et plus grand avantage concurrentiel.

## Dans quels scénarios utiliser le SDS ?

Dans son article intitulé [\*A Buyer's Guide to Software-Defined Storage \(SDS\): Establishing a Sustainable Data Strategy\*](#) (mai 2016), Neuralytix avertissait ses lecteurs que

---

*« Le SDS doit faire partie intégrante de la stratégie de toute entreprise informatique. »*

---

Nous ne pouvons qu'insister de nouveau sur ce conseil.

Outre son rôle de « libération » des données, le SDS propose d'autres fonctions tout aussi essentielles. Parce qu'elles sont définies par les logiciels, les solutions SDS accélèrent le provisionnement et le transfert des machines virtuelles et des matériels bare metal.

La prise en charge de nombreux protocoles d'accès aux données permet de transférer les données avec des méthodes d'accès classiques (SMB, NFS) mais aussi de nouvelle génération (HDFS, S3). Ces dernières sont vitales pour les applications nécessitant une évolutivité massive, notamment les applications Hadoop et les autres applications de Big Data.

Les transferts de données ne sont pas limités à des transferts 100 % locaux, mais avec les protocoles comme S3, cela permet aux données de circuler facilement entre des instanciations locales et cloud pour obtenir plus de capacité de stockage.




La libération des données, résultat direct d'un flux sans entrave des transferts de données, comporte toutefois un risque, celui d'écourter la durée avant une panne. L'adage « panne fréquente, panne rapide » s'applique ici aussi. Il indique que les expérimentations, dont beaucoup ne donneront pas le résultat souhaité, permettent néanmoins aux entreprises de tester autant de scénarios que possible pour obtenir l'état voulu.

L'expérimentation leur permet également de détecter de nouvelles opportunités insoupçonnées, avec à la clé une innovation réelle : amélioration et nouveauté des produits et solutions, meilleur service à la clientèle, augmentation des niveaux de service et de la productivité.




Les données libérées, si elles sont utilisées correctement, permettent à l'entreprise d'obtenir plus vite des connaissances et donc d'agir plus vite. Elles ont un impact direct sur les résultats métier : croissance ou avantage concurrentiel. L'objectif final est d'améliorer la santé globale de l'entreprise.

## Indicateurs clés de performance (KPI) pour une implémentation réussie du SDS

De l'avis de Neuralytix, comme dans toutes les technologies modernes de l'information, trois indicateurs clés de performance permettent de mesurer la réussite :

-  Prévisibilité
-  Reproductibilité
-  Evolutivité

L'objectif principal du secteur informatique doit être de créer de la valeur à partir de l'information. Dans cette optique, l'infrastructure doit présenter des caractéristiques qui se prêtent à la création de valeur, à l'obtention d'un avantage concurrentiel et à la croissance.

-  Dans une infrastructure optimale, les résultats et la performance doivent toujours être prévisibles. Cette prévisibilité ne doit pas être affectée par l'ajout ou le retrait d'une application, d'un nœud ou de capacité de traitement ou de stockage.
-  Les processus intervenant dans l'évolution du système doivent être reproductibles. Ici encore, quelles que soient les modifications apportées à l'infrastructure, elles doivent suivre un processus reproductible (de préférence automatisé) donnant des résultats prévisibles.
-  Dans l'idéal, l'infrastructure doit être capable d'évoluer à l'infini. A défaut, chacune de ses dimensions (application, traitement, stockage, etc.) doit, au minimum, être prévisible sur 5 à 10 ans.

Selon Neuralytix, la solution IBM Spectrum Storage Suite répond en tous points à ces trois KPI. Elle permet aux clients de mettre en place une solution de SDS performante et contribue à impulser la croissance métier, l'innovation et l'avantage concurrentiel.

## La structure SDS

Nous avons vu que le SDS est une structure et non un logiciel individuel. Cette structure comprend une ensemble de solutions individuelles interdépendantes qui, une fois réunies, créent le SDS.

Au niveau le plus élémentaire, le SDS doit proposer la virtualisation du stockage et un pool de stockage évolutif. Pour que le SDS apporte de la valeur à un datacenter, il doit pouvoir en virtualiser toutes les ressources de stockage. Toutefois, si une solution SDS se limitait à virtualiser le stockage, elle ne ferait qu'ajouter une couche superflue à une pile de stockage déjà complexe. Pour apporter une réelle valeur à la pile de stockage, une solution SDS doit donc comporter un système de fichiers capable de s'adapter à toutes les ressources de stockage disponibles. Dans l'idéal, les transferts de données doivent être évités. Le défi, pour une solution SDS, est de disposer d'un système de stockage de type scale-out qui soit à la fois mature, fiable et doté de fonctionnalités à la hauteur des besoins de l'entreprise.

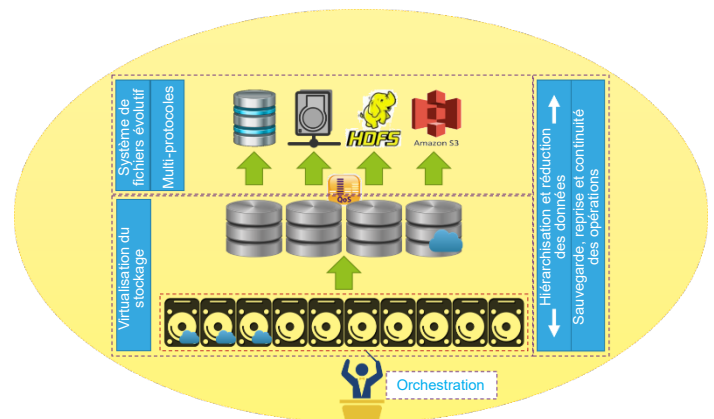


Figure 2 : La structure SDS (Neuralytix 2016)

Outre la virtualisation du stockage et un système de fichiers scale-out, une solution SDS doit pouvoir présenter les données dans un format adapté aux besoins des applications. Face à la diversification des applications, SDS est amené à devoir prendre en charge un nombre croissant de protocoles.

Si les données virtualisées sont de type bloc, la solution SDS doit pouvoir les présenter au format bloc, et ainsi de suite pour les données de fichier, d'objet, etc. En ce qui concerne les services de données les plus élémentaires que la solution SDS doit pouvoir fournir, ils doivent comprendre la présentation des données via les protocoles de bloc, de fichier, S3 et HDFS.

Ce sont là les fonctionnalités les plus simples qu'une solution SDS doit pouvoir offrir. Ensemble, elles forment la première génération des logiciels SDS.

L'arrivée à maturité des logiciels SDS a vu apparaître tout un nouveau groupe de fonctions plus performantes.

Dans l'idéal, les autres services de données qu'une solution SDS doit fournir sont la hiérarchisation des données, le dédoublonnage de données, la compression de données, la sauvegarde et la récupération, l'archivage et la continuité des opérations.

Tous les services de données doivent être associés à un niveau de qualité de service compatible avec les applications qui accèdent aux données.

La couche d'orchestration et de gestion du logiciel SDS est un indice fondamental qui révèle s'il s'agit réellement d'une solution de nouvelle génération.

L'orchestration et la gestion des données et du stockage ne doivent pas se limiter à des solutions sur site. Elles doivent aussi pouvoir s'adapter à la nouvelle tendance des clouds hybrides. En d'autres termes, la couche d'orchestration et de gestion du logiciel SDS doit pouvoir intégrer, contrôler et gérer les données locales et sur le cloud (public ou privé) par le biais d'une interface unique. Elle doit aussi traiter le stockage cloud en tant que palier au sein de la stratégie de stockage globale.

Dans l'idéal, les fonctions de gestion doivent être accessibles à partir du cloud afin de permettre de gérer à la fois le stockage local et sur le cloud. Les données doivent pouvoir être transférées librement et sans restrictions entre les cibles de stockage sur site et sur le cloud.

## Le SDS, porte d'entrée du cloud hybride

En définitive, toutes les couches d'abstraction, de gestion, d'orchestration, etc., ont pour objectif final l'abstraction des données par rapport à l'infrastructure.

C'est ce que l'on désigne souvent par l'expression « mobilité des données ». La mobilité des données permet aux données d'être transférées dynamiquement de l'utilisateur final vers les systèmes principaux, puis vers les périphériques mobiles et les systèmes de rapport et d'analyse. Les données peuvent également être transférées de et vers le cloud chaque fois que nécessaire.

De nos jours, le traitement est de plus en plus souvent exécuté sur le cloud, et un volume croissant de données est absorbé à partir de points de collecte des données basés sur le cloud. Les solutions SDS doivent donc intégrer la mobilité des données non seulement au sein du datacenter, mais aussi entre les différents datacenters via des clouds publics ou privés, ainsi qu'entre des cibles de stockage sur site et sur cloud.







Outre le traitement des données, la mobilité des données permet aussi aux entreprises de tirer parti de l'économie des cibles de stockage des données sur site et sur cloud.

Par exemple, il est préférable de stocker les données temporelles sur cloud, où les coûts sont moindres à court terme, contrairement à l'archivage à long terme qui doit privilégier le stockage sur bande, hors ligne ou hors site.

## IBM Spectrum Storage Suite

IBM Spectrum Storage Suite est une offre SDS unique en son genre. La plupart des solutions SDS du marché donnent la priorité à des solutions individuelles dédiées aux données de bloc, de fichier, d'objet ou aux machines virtuelles. IBM Spectrum Storage Suite offre en revanche flexibilité et agilité au client qui peut acheter et déployer uniquement les composants dont il a besoin. Il permet également de bénéficier de la prise en charge multi-applications, multi-protocoles, multi-charges de travail, multi-emplacements (sur site et sur le cloud) que les entreprises sont en droit d'exiger de leur logiciel SDS.

IBM Spectrum Storage Software Suite comprend six composants :

-  **Spectrum Accelerate** – Garantit une performance régulière des données et fournit une grille extensible de serveurs, idéale pour déployer et étendre rapidement les bases de données et gérer des machines virtuelles ainsi que des charges de travail transactionnelles.
-  **Spectrum Archive** – Permet de traiter les données stockées sur bande linéaires de la même façon que celles qui sont stockées sur disques magnétiques ou flash.
-  **Spectrum Control** – Offre une gestion unifiée du SDS ainsi qu'un stockage physique classique, avec des technologies de hiérarchisation et de réduction des données qui permettent d'abaisser les coûts de stockage d'environ 73 %.<sup>ii</sup>
-  **Spectrum Protect** – Fournit des fonctions rapides de sauvegarde, d'image instantanée et de reprise, étendues à toute l'entreprise et au cloud hybride.
-  **Spectrum Scale** – Fournit un système de fichiers mature à hautes performances et haute évolutivité, prenant en charge les objets et HDFS.
-  **Spectrum Virtualize** – Permet de virtualiser toutes les ressources de stockage. Protège et active les investissements existants en les intégrant à une structure SDS.

L'étude de Neuralytix laisse à penser qu'il existe un stockage cloud officiel (ou fantôme) dans 100 % des entreprises. Dans les faits, toutes les entreprises recherchent un équilibre entre le stockage local et sur le cloud des données et finissent par opter pour une stratégie de cloud hybride.

Sans IBM Spectrum Software, la plupart des entreprises vont se retrouver dans la situation où elles vont choisir une solution SDS sur site et une solution cloud entièrement séparée. Cette stratégie est loin d'être idéale.

Le principe de la mobilité des données reste crucial. Les données doivent être transférées librement et moyennant un minimum de friction entre les cibles locales et sur le cloud. Pour ce faire, la solution retenue doit pouvoir traiter de façon abstraite les cibles de stockage locales et sur cloud, à l'aide d'une plate-forme unique de gestion et d'orchestration.

IBM Spectrum Storage Software Suite est parfait pour atteindre ces objectifs.

## Exemple : Utilisation du stockage cloud en tant que « niveau »

Pour inclure un « niveau » cloud dans une stratégie de stockage, vous devez disposer d'un système de fichiers évolutif et compatible qui englobe le stockage sur site et le stockage cloud. IBM propose des solutions inédites dans ce domaine. IBM Spectrum Scale, l'un des composants de la structure IBM SDS, est un système de fichiers hautement évolutif et à hautes performances destiné aux déploiements SDS sur site ou sur le cloud. Le client peut ainsi définir le stockage cloud sous la forme d'un niveau hiérarchique dans le cadre de sa stratégie globale de hiérarchisation des données.

Si les clients préfèrent un cloud public, Spectrum Scale leur permet d'utiliser les protocoles S3 ou Swift pour transférer leurs données sur des clouds publics tels qu'Amazon Web Services (AWS) S3. Si nécessaire, le stockage de fichiers sur site peut être instantanément converti en stockage cloud d'objets (S3) pour permettre d'exploiter sur le cloud des données qui résidaient uniquement dans le datacenter.

On peut dans ce cas parler de réelle mobilité des données.

## Exemple : Protection des données sur des supports multiples

Traditionnellement, la procédure de protection des données consiste à les sauvegarder sur une cible locale permettant de les récupérer rapidement, puis de les archiver sur bande.

Toutefois, avec Spectrum Protect, IBM peut exploiter le stockage sur le cloud en tant que niveau hiérarchique et par conséquent sauvegarder les données sur plusieurs cibles (disques locaux, cloud et bande). La solution apporte flexibilité et agilité au client qui peut choisir le support de sauvegarde le plus économique disponible, en fonction de ses besoins.

## Exemple : Créez votre propre cloud de stockage et devenez votre propre fournisseur de services

L'objectif de nombreuses sociétés informatiques est de devenir le fournisseur de services de leurs utilisateurs (clients). Fournir aux clients de la capacité de stockage destinée aux données de valeur est une opération cruciale.

Le provisionnement du stockage doit respecter les indicateurs clés de performance décrits précédemment, avec notamment des niveaux maximums de prévisibilité et de reproductibilité, l'évolutivité allant de soi.

IBM Spectrum Accelerate permet de mettre en ligne des serveurs bare-metal et du stockage et de fournir de la capacité de stockage en quelques minutes.

Spectrum Accelerate s'exécute également sur le cloud sur Softlayer, le cloud IaaS d'IBM. Le stockage peut donc être provisionné en quelques minutes pour les opérations sur site et sur le cloud pour les applications cloud, le tout via une interface, un processus, une orchestration et une gestion unique.

Les fournisseurs de services et l'équipe IT peuvent en outre utiliser Spectrum Accelerate pour créer des cibles de sauvegarde pour les clients externes, car Spectrum Accelerate a été conçu en vue d'une utilisation à services partagés.

## Conclusion

Comme le montrent ces exemples, IBM Spectrum Software peut être déployé dans de nombreux cas d'utilisation. Très souvent, il est possible de combiner différents cas d'utilisation avec les composants appropriés d'IBM Spectrum Software pour créer un environnement SDS ultra-agile hautes performances.

Du point de vue de Neuralytix, ce qui distingue IBM Spectrum Software Suite des autres solutions SDS est sa parfaite richesse fonctionnelle et son agilité.

Cette suite IBM comprend toutes les composantes indispensables que l'entreprise agile et évolutive est en droit d'attendre d'une solution de stockage.

IBM Spectrum Software Suite répond aux exigences multi-protocoles, multi-applications, multi-emplacements et multi-charges de travail de l'entreprise moderne. Il a réuni dans cette suite des solutions matures garantissant un déploiement agile et comprenant toutes les fonctionnalités nécessaires.

La capacité à gérer à la fois le stockage local et le stockage sur le cloud est une caractéristique qui est de plus en plus souvent exigée par les entreprises, qui ont bien compris l'économie et les avantages du stockage cloud.

Une solution SDS ne doit pas répondre à un besoin isolé. Elle doit au contraire fournir une suite complète de services de stockage et de données. IBM a su mettre au point une solution qui répond brillamment à tous ces critères.

Neuralytix recommande vivement à ses clients la solution IBM Spectrum Suite lors de l'élaboration de leur stratégie de stockage.

## Notes de fin

<sup>i</sup> Extrait de *A Buyer's Guide to Software-Defined Storage (SDS): Establishing a Sustainable Data Strategy* (Neuralytix, mai 2016)

<sup>ii</sup> Sur la base des calculs d'IBM

### A propos de Neuralytix

Neuralytix est un grand cabinet international de conseil et d'étude de marché spécialisé dans l'informatique. Ses trois principes fondateurs sont les suivants :

Les analyses quantitatives de technologies disparates ne correspondent pas à une vision actuelle ou future du marché. Nous devons examiner les domaines technologiques et non des technologies disparates. En d'autres termes, nous devons adopter une approche globale du problème métier que nous tentons de résoudre.  
Notre rôle principal consiste à aider nos clients à atteindre leurs objectifs métier.

Nous comptons parmi nos clients des constructeurs informatiques parmi les plus réputés, ainsi que de nouveaux acteurs du secteur privé. Notre activité de conseil aux utilisateurs finaux connaît une croissance agressive, de même que notre nouvelle offre Neuralytix Investor, qui cible les prescripteurs financiers. Nos clients du secteur public totalisent à eux seuls une capitalisation de marché combinée d'environ US\$4 billions, soit une influence d'environ US\$1,5 billion sur les dépenses IT annuelles.

Nos clients apprécient le fait que leur engagement auprès de Neuralytix leur procure l'avantage d'être les premiers à agir. Nous leur permettons d'atteindre leurs objectifs métier grâce au précepte de base qui sous-tend toutes nos actions : proposer des Stratégies pour demain.

### Nous contacter

Neuralytix®  
755 Sansome Street  
Suite 360  
San Francisco, CA 94111

Téléphone : (415) 919-1188  
Email : [info@neuralytix.com](mailto:info@neuralytix.com)  
Suivez-nous sur  
Twitter : [@Neuralytix](https://twitter.com/Neuralytix)

Nos bureaux dans le monde :  
Australie  
Brésil  
Hong Kong, SAR  
Allemagne  
Royaume-Uni