



---

**Aniket Kulkarni, IBM**

Littleton, Massachusetts, États-Unis  
Adresse e-mail : aniket.kulkarni@us.ibm.com

**Venkatesh Gopal, IBM**

Leawood, Kansas, États-Unis  
Adresse e-mail : gopalv@us.ibm.com

---

# Haute disponibilité et reprise dans IBM Db2 Warehouse on Cloud

## Résumé

L'un des indicateurs clés pour évaluer les services Cloud est la disponibilité. Les fournisseurs doivent concevoir leurs services Cloud pour réduire le risque de défaillance des composants sous-jacents, un risque qui augmente à mesure que les services gagnent en taille et en complexité.

La haute disponibilité (HA) correspond à la capacité d'un service à gérer ces risques. Au niveau le plus basique, elle permet de (a) détecter les défaillances de composant et de (b) redevenir opérationnel de manière automatisée. Si les systèmes de gestion et d'orchestration de conteneurs comme Kubernetes proposent des fonctionnalités similaires, leurs modalités d'utilisation varient pour garantir des niveaux de fiabilité supérieurs dans certains types de services.

Dans cet article, nous allons examiner comment IBM® Db2 Warehouse on Cloud, l'entrepôt de données Cloud d'IBM, permet d'obtenir une haute disponibilité et une haute fiabilité. Nous allons décrire plus précisément son modèle de service conteneurisé, sa fiabilité multiniveau et sa capacité de reprise en cas de défaillances.

## Fiabilité multiniveau

Db2 Warehouse on Cloud offre aux clients un entrepôt de données Cloud massivement parallèle, à mémoire distribuée (« Shared Nothing ») et déployé sous la forme d'un modèle de service conteneurisé<sup>1</sup> basé sur IBM Cloud Container. Le service IBM Cloud Container se superpose à Kubernetes<sup>2</sup>, un système d'orchestration de conteneurs Open Source, qui automatise le déploiement, dimensionne et gère les applications conteneurisées.



Pour assurer une haute disponibilité et une haute fiabilité dans Db2 Warehouse on Cloud, IBM s'appuie sur Kubernetes avec une couche de fiabilité multiniveau qui comprend :

- Niveau 1 : aucun point de défaillance dans le système, chaque composant a une ou plusieurs copies redondantes, et la reprise automatique intervient en cas de défaillance de plusieurs composants redondants
- Niveau 2 : reprise intra-conteneur automatique en cas de défaillance de composants à l'intérieur d'un conteneur
- Niveau 3 : reprise de conteneur automatique en cas de défaillance d'un conteneur
- Niveau 4 : reprise automatique du service en cas de défaillance système du ou des serveurs qui hébergent le ou les conteneurs

De plus, Db2 Warehouse on Cloud intègre une fonctionnalité de sauvegarde et de restauration rapides, contrôlée et définie par l'utilisateur. Ce service assure également une prise en charge continue de DevOps qui surveille les défaillances et l'intégrité du système en cas de défaillances multiples ou irrécupérables que même la couche de fiabilité multiniveau n'est pas capable de gérer.

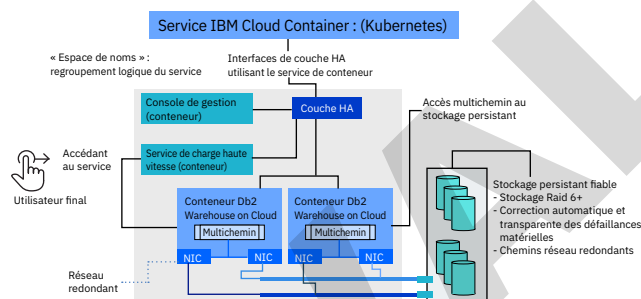


Figure 1 : Représentation simplifiée de l'architecture qui sous-tend la fiabilité du service Db2 Warehouse on Cloud

- Chaque composant du service est un conteneur.
- La couche HA étend et complète Kubernetes pour offrir une haute disponibilité plus rapidement.
- Les connexions réseau au stockage persistant sont redondantes et regroupées logiquement dans un multichemin.
- La couche HA peut rétablir des conteneurs en cas de défaillance de composants à l'intérieur de ceux-ci, sans redémarrer le conteneur, pour accélérer la reprise.
- Le stockage persistant a plusieurs redondances au niveau du réseau et du stockage.

Examinons plus en détail chacun de ces niveaux, ainsi que les fonctionnalités de sauvegarde et de restauration.

## Fiabilité de niveau 1 : redondance des composants

Db2 Warehouse on Cloud est un entrepôt de données élastique, qui permet aux utilisateurs de dimensionner les cœurs d'UC ainsi que les capacités de stockage indépendamment et à la demande en fonction des besoins métier. Cette élasticité de l'architecture provient du découplage entre calcul et stockage. La capacité de calcul est gérée par le service IBM Cloud Container qui s'appuie sur Kubernetes et le stockage est placé dans le stockage par blocs IBM Cloud hautes performances en réseau<sup>3</sup>.

Le réseau étant un composant essentiel pour la fiabilité et la performance, IBM a également vérifié que le réseau est exempt de tout point de défaillance. Chacun des serveurs sous-jacents a une paire de contrôleurs d'interface réseau redondants, qui sont connectés à différents chemins conduisant au fond de panier de stockage et qui sont reliés ensemble pour optimiser les performances (en équilibrant la charge) dans la mesure du possible. De plus, le fond de panier de stockage a une paire de points d'accès (ou « cibles ») redondants, qui permettent à l'infrastructure de stockage d'accéder au réseau. S'agissant du calcul (conteneur), une configuration multichemin active/passive garantit un « basculement » sans interruption. Ainsi, en cas de défaillance d'un des chemins de stockage, le système bascule vers l'autre automatiquement.

Par ailleurs, le système de stockage utilise plusieurs disques SSD (Solid State Drive) redondants et échangeables à chaud dans une configuration RAID (Redundant Array of Independent Disks) 6+. Plus performante que la RAID 6 traditionnelle, cette configuration permet de gérer et de réparer tout problème concernant les disques SSD sans aucune interruption.

## Fiabilité de niveau 2 : reprise intra-conteneur

Dans un modèle de services Web ou Kubernetes traditionnel, c'est le redémarrage du conteneur qui résout la défaillance d'un composant à l'intérieur d'un conteneur. Si simple et efficace qu'elle soit, cette méthode ne répond pas aux besoins des utilisateurs qui exigent une haute disponibilité pour leurs applications stratégiques.

Compte tenu des charges de travail dans un entrepôt de données Cloud, où les ensembles de données sont volumineux, il est plus rapide de restaurer un composant défaillant du conteneur que de redémarrer l'ensemble du conteneur.

Un redémarrage complet obligerait à dissocier le conteneur du stockage, à arrêter le calcul, à sélectionner une nouvelle ressource de calcul puis à lui réassocier un stockage. En revanche, gérer directement le composant concerné permet de gagner un temps précieux et de minimiser l'interruption maximale admissible ou RTO (Recovery Time Object).

Cette limite du modèle traditionnel est contournée en divisant ces reprises en deux niveaux : le niveau 2 et le niveau 3. Dans Db2 Warehouse on Cloud, la « couche HA » s'étend à chacun des conteneurs, d'où son nom « Couche HA de conteneur ». Cette couche surveille l'intégrité de chacun des composants à l'intérieur du conteneur et exécute les actions correctives nécessaires en cas de défaillance.

Prenons l'exemple d'un conteneur Db2 qui connaît une défaillance dans l'un de ses processus de partition de données. La couche HA de ce conteneur va détecter cette défaillance et tenter de redémarrer ces partitions de données, puis de restaurer la base de données pour garantir la cohérence des données. De plus, en cas de défaillances en cascade (par exemple, défaillance d'une partition de données pendant la reprise d'une autre partition), la couche HA de conteneur va annuler la reprise en cours et procéder à une reprise complète des deux partitions.

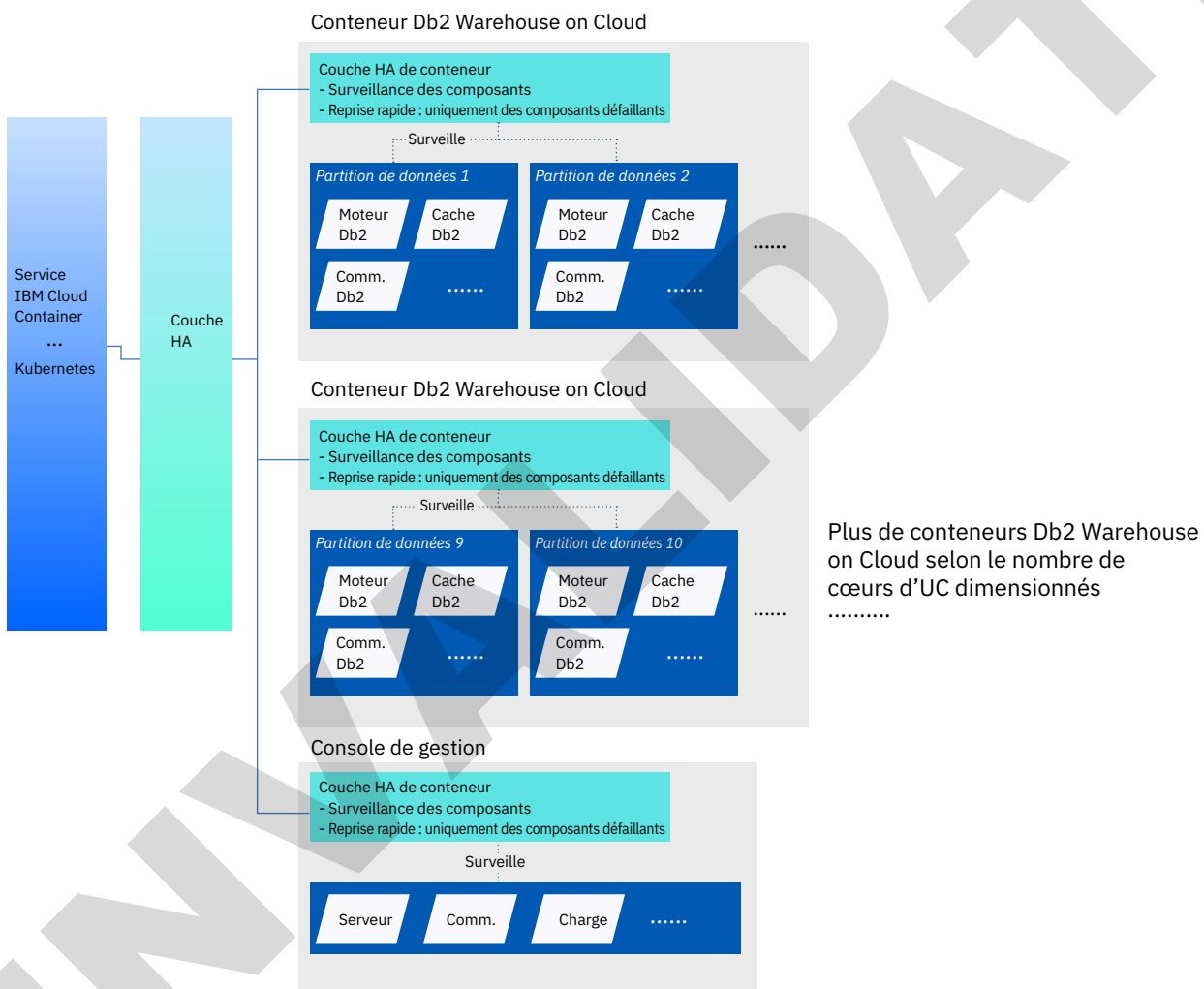


Figure 2 : Haute disponibilité intra-conteneur de niveau 2 :

- Chaque conteneur intègre une couche HA qui tente d'assurer une haute disponibilité rapidement en cas de défaillance d'un sous-ensemble de composants à l'intérieur de ce conteneur.
- Dans les conteneurs Db2 Warehouse on Cloud, la couche HA surveille l'ensemble des partitions et des composants de chaque partition. Seules les partitions défaillantes font l'objet d'un redémarrage et d'une reprise automatiques.
- Dans les autres conteneurs, ce sont tous les composants qui font l'objet d'une surveillance et d'une reprise dans la mesure du possible.
- La couche HA joue le rôle d'interface avec la couche HA du service de conteneur, et la complète. Ce qui permet de passer à une reprise de niveau 3, si une reprise de niveau 2 n'est pas possible.

### Fiabilité de niveau 3 : reprise de conteneur

Si l'un des conteneurs connaît plusieurs erreurs ou si le conteneur s'est arrêté, une reprise de niveau 3 est lancée. Dans ce cas, la couche HA revient au modèle HA Kubernetes traditionnel de redémarrage de conteneur et :

1. Le ou les conteneurs concernés sont arrêtés correctement et nettoyés.
2. Le stockage persistant associé est dissocié et le service de conteneur réattribue le conteneur à d'autres cœurs de calcul.
3. Ensuite, le stockage est automatiquement réassocié et le conteneur démarre.

Notez que pour un service comme Db2 Warehouse on Cloud qui comprend plusieurs conteneurs, la couche HA pilote une reprise simultanée de la base de données dans tous les conteneurs.

En plus d'être une solution de repli à une fiabilité de niveau 2, cette reprise gère une autre utilisation importante pour l'intégrité des données en cas de défaillances multiples au niveau du réseau. Dans l'éventualité – rare – de plusieurs défaillances réseau sur chacun des chemins réseau redondants, Db2 Warehouse on Cloud met automatiquement le stockage en lecture seule pour protéger la cohérence des données.

Si la couche HA détecte ce cas de figure, elle force le conteneur à s'arrêter, le dissocie, le réattribue, puis l'associe et le redémarre. Le conteneur est alors transféré à un autre serveur sous-jacent, avec un réseau sain et déprotège le stockage, pour préserver l'intégrité des données et rétablir le système.

### Fiabilité de niveau 4 : reprise du service

La reprise de conteneur est le modèle HA Kubernetes le plus traditionnel. La couche HA permet à Kubernetes de détecter les défaillances totales des serveurs sous-jacents qui fournissent les cœurs de calcul aux conteneurs.

Cette reprise est similaire à la reprise de niveau 3 : tous les conteneurs sont arrêtés, le stockage concerné est dissocié, tous les conteneurs affectés sont réattribués, le stockage entier est associé, puis tous les conteneurs sont redémarrés.

La couche HA détecte un serveur qui tombe en panne ou ne répond plus, par l'absence de plusieurs signaux. Après un délai d'attente, les actions de reprise (mentionnées ci-dessus) sont exécutées et le serveur, identifié comme « non réattribuable », ne sera plus utilisé pour démarrer d'autres conteneurs. Cela permet aux équipes des opérations de réparer le serveur concerné puis de le remettre en ligne, sans perturbation supplémentaire pour l'utilisateur final.

### Sauvegarde et restauration

La possibilité pour les utilisateurs de sauvegarder et restaurer la base de données est essentielle afin de protéger le service Cloud contre les erreurs et permet, en dernier ressort et dans de rares cas, de rétablir le service en cas de défaillance totale. Db2 Warehouse on Cloud offre aux utilisateurs une interface simple pour gérer cette fonctionnalité eux-mêmes :

- L'utilisateur peut programmer des sauvegardes au moment qui convient le mieux à son entreprise. Les sauvegardes s'exécutent une fois par 24 heures à l'heure définie.
- Les sept dernières sauvegardes sont conservées et, au besoin, les utilisateurs peuvent restaurer la base de données rapidement à partir de l'une de ces sauvegardes, par un simple clic.

### Instantanés ROW (Redirect-on-Write) : la clé pour des sauvegardes et des restaurations ultra-rapides

Fiable, la couche de stockage persistant prend également en charge les instantanés ROW (Redirect-on-Write) quasi immédiats pour les sauvegardes. Chaque partition de données dans la base de données bénéficie d'un volume de stockage persistant, fiable et en réseau. La figure 3 montre (en minutes) qu'une sauvegarde par instantané est beaucoup plus rapide que les technologies de sauvegarde traditionnelles pour une base de données de 3 à 4,2 To. Les sauvegardes par instantané ne durent que de 1 à 3 minutes, alors que les sauvegardes de base de données traditionnelles prennent entre 2 et 3,5 heures.

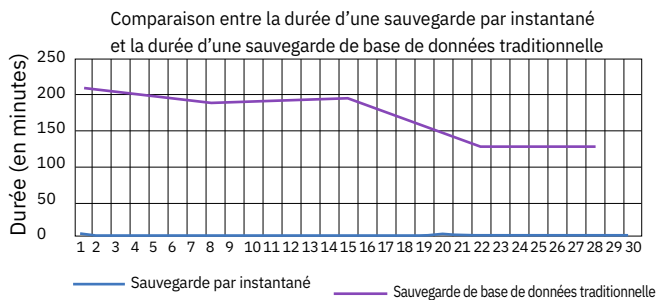


Figure 3 : Ce tracé compare la durée (en minutes) de sauvegarde d'une base de données de 3 à 4,2 To sur un mois, entre la sauvegarde par instantané et les sauvegardes de base de données traditionnelles. Ces données ont été mesurées sur différents systèmes exécutant des charges de travail de production.

Variante optimisée de la technologie d'instantané COW (Copy-on-Write), la technologie ROW (Redirect-on-Write) s'utilise pour le stockage persistant fiable. Les blocs de données sont reliés par des références. L'instantané n'est qu'un simple groupe de références qui permettent au lecteur de lire le bon ensemble de blocs.

Par exemple, si l'utilisateur prend un instantané puis modifie un bloc dans le volume, le volume de stockage inscrit les modifications à un nouvel emplacement et fait pointer les références vers le nouvel emplacement correspondant à la version actuelle de ce bloc.

L'instantané va continuer à faire référence à l'ancien bloc, tandis que les autres blocs restent inchangés et sont identifiés par l'instantané et la vue actuelle.

Db2 Warehouse on Cloud interrompt la charge de travail du système pendant une minute environ et utilise cet état de veille pour prendre un instantané du volume avant de « réveiller » la charge de travail. Les sept sauvegardes les plus récentes sont conservées.

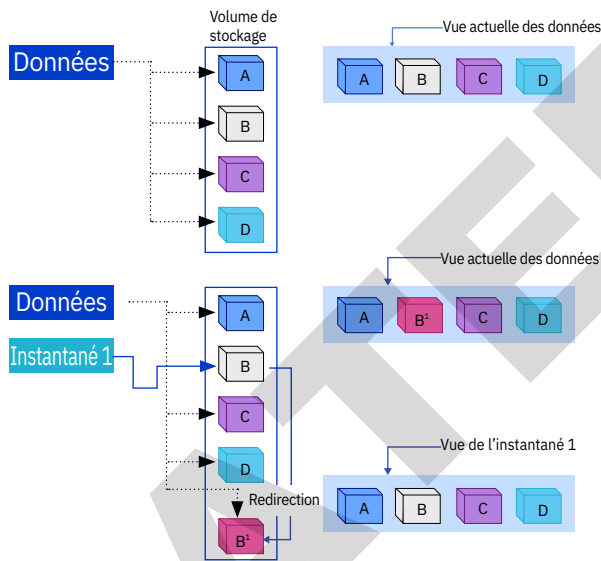


Figure 4 : Cette figure décrit l'opération ROW. Le volume utilise les références pour accéder à la version la plus à jour des blocs et permettre à l'utilisateur final de les lire. Si un instantané (Instantané 1) est créé et que B est remplacé par B', les blocs d'origine restent inchangés et B' est inscrit à un nouvel emplacement.

L'instantané 1 pointe désormais vers le bloc B d'origine, tandis que la version actuelle des données pointe vers B'. Si un instantané (Instantané 1) est créé et que B est remplacé par B', les blocs d'origine restent inchangés et B' est inscrit à un nouvel emplacement.

L'instantané 1 pointe désormais vers le bloc B d'origine, tandis que la vue actuelle des données pointe vers B'.

Cette stratégie couplée à la technologie ROW offre à l'utilisateur la souplesse d'une « restauration temporelle ». C'est-à-dire que l'utilisateur peut restaurer l'instantané 1, tout en conservant les instantanés 2 à 7, ce qui lui permet de restaurer l'instantané dont il a besoin.

### En savoir plus

Une [présentation du produit](#) est également disponible si vous souhaitez interagir avec l'interface de Db2 Warehouse on Cloud.



IBM France  
17 Avenue de l'Europe  
92275 Bois Colombes Cedex

IBM, le logo IBM, **ibm.com**, Db2 et IBM Cloud sont des marques d'International Business Machines Corp., déposées dans de nombreux pays du monde. Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés. Une liste actualisée des marques déposées IBM est accessible sur le web sous la mention « Copyright and trademark information » à l'adresse [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Ce document est considéré comme à jour à sa date initiale de publication et peut être modifié par IBM à tout moment. Toutes les offres ne sont pas disponibles dans chaque pays dans lequel opère IBM.

Les exemples de clients fournis ne sont mentionnés qu'à titre d'information. Les performances réelles peuvent varier selon les configurations et les conditions de fonctionnement spécifiques. Il appartient à l'utilisateur d'évaluer et de vérifier le fonctionnement d'autres produits ou programmes avec les produits et programmes d'IBM. LES INFORMATIONS CONTENUES DANS CE DOCUMENT SONT FOURNIES « EN L'ÉTAT », SANS AUCUNE GARANTIE EXPRESSE OU TACITE, NOTAMMENT SANS AUCUNE GARANTIE DE QUALITÉ MARCHANDE OU D'ADAPTATION À UN EMPLOI SPÉCIFIQUE, ET SANS AUCUNE GARANTIE OU CONDITION DE NON-INFRACTION VIS-À-VIS DES LOIS. Les produits IBM bénéficient d'une garantie, conformément aux conditions générales des contrats dans le cadre desquels ils sont fournis.

© Copyright IBM Corporation 2019

- 1 « Service IBM Cloud Kubernetes. »  
<https://www.ibm.com/cloud/container-service>
- 2 « Orchestration de conteneur de production. »  
<https://kubernetes.io/>
- 3 « Stockage par blocs. »  
<https://www.ibm.com/cloud/block-storage>



Veillez recycler