



---

## Les points forts

- Assurer un accès sécurisé aux applications stratégiques de l'entreprise.
  - Charger, analyser et visualiser très rapidement vos données.
  - Télécharger des données à partir de différentes sources et les intégrer grâce au langage R.
  - Obtenir de meilleures informations utiles grâce à des algorithmes d'analyse prédictive en base de données.
  - Élargir au Cloud les environnements d'entrepôts de données « on-premise ».
  - Analyser les données JSON grâce à l'intégration native avec IBM Cloudant.
- 

# IBM dashDB

*Entrepôts de données « as-a-service »,  
conçus pour l'analytique*

IBM® dashDB est un entrepôt de données dans le Cloud, rapide et entièrement géré, qui utilise l'analytique intégrée pour apporter rapidement des réponses. Ses capacités inégalées (analytique en base de données, modélisation prédictive avec le langage R, outils de business intelligence) vous donnent toute liberté pour analyser vos données et générer plus rapidement des connaissances précises.

IBM dashDB est très simple à mettre en œuvre grâce aux capacités de provisionnement rapide de la plateforme IBM Bluemix. Vous pouvez tester la solution ou commencer à utiliser dashDB gratuitement, jusqu'à 1 Go de données, puis pour seulement 50 dollars par mois pour un espace de stockage de données de 20 Go. Des configurations plus étendues d'une capacité de plusieurs téraoctets sont possibles en fonction de l'évolution de vos besoins en stockage de données et de la nécessité pour vos utilisateurs de disposer d'un environnement dédié. En outre, les traitements massivement parallèles (MPP) permettent à la fois d'accélérer l'exécution des requêtes et d'étendre les volumes de données traitées.

IBM dashDB concilie la simplicité d'une appliance d'entrepôt de données en tant que service, et l'agilité et l'évolutivité du Cloud, pour répondre aux besoins des entreprises, toutes tailles confondues. Vous pouvez ainsi composer rapidement des applications d'analyse grâce à un riche éventail de services pour développeurs et d'offres complémentaires disponibles dans IBM Bluemix ou avec les outils de votre choix « on premise ».

## Les entrepôts de données avant dashDB

La construction d'un entrepôt de données a souvent été un travail laborieux. Jusqu'ici, vous aviez à choisir un logiciel d'entrepôt de données spécifique, puis à déterminer et assurer l'équilibre adéquat entre matériels et ressources de stockage nécessaires. Une fois le côté physique de l'entrepôt de données constitué, il était nécessaire à la fois de bâtir le système et les modèles de données logiques nécessaires pour accompagner votre démarche. Lorsqu'il était nécessaire d'étendre l'entrepôt de données (et c'était forcément le cas puisque le volume des données collectées progressait en permanence), il était inévitable d'acquiescer davantage de puissance de traitement, de ressources de stockage et de logiciels.



L'ensemble du processus introduisait des risques dès que vous aviez à modifier l'entrepôt de données. Aviez-vous maintenu le bon équilibre pour le traitement des données ? Aviez-vous acheté le bon matériel ? Les données étaient-elles allouées correctement à l'échelle du système ? Le matériel était-il dépassé et incompatible avec les technologies les plus récentes du marché ? Le logiciel nécessitait-il des mises à jour ? À tout moment du processus d'extension, vous étiez exposé à tout un ensemble de problèmes potentiels. À cela s'ajoutait le fait que vous aviez à payer pour des capacités même si vous ne les utilisiez pas. Et c'est vous et votre entreprise qui preniez tous les risques.

### L'arrivée des appliances d'entrepôt de données

Les appliances d'entrepôt de données ont contribué à alléger l'essentiel des problèmes liés à la construction de votre propre entrepôt de données. Ces systèmes étaient préconfigurés et intégraient l'entrepôt de données pour fournir de bonnes performances analytiques. Il ne vous restait plus qu'à choisir votre modèle et la taille de la configuration requise, et une fois le système branché et opérationnel, à y charger vos données.

L'attribution de matériels et logiciels était effectuée pour vous, mais l'appliance imposait tout de même d'acquiescer un nouveau système dès que les limites de ressources étaient atteintes. Compte tenu du volume croissant des données, des nouvelles applications et de l'augmentation des effectifs d'utilisateurs, la mise en œuvre d'un nouvel entrepôt de données pouvait s'avérer coûteuse en cas d'extension en dynamique. Ajoutez à ceci les mises à jour, les correctifs de logiciels, la maintenance et la dépréciation globale du matériel et vous vous retrouvez avec l'essentiel des opérations de maintenance nécessaires aux entrepôts de données traditionnels.

Une appliance d'entrepôt de données reste une partie essentielle du système décisionnel d'une entreprise. Cependant, n'existerait-il pas une technologie complémentaire pouvant alléger certaines de ces problématiques croissantes ? Examinons la solution IBM dashDB.

### Les avantages du Cloud

Grâce aux caractéristiques d'évolutivité et d'accès à la demande du Cloud computing, vous pouvez assurer très rapidement l'installation et le fonctionnement d'un entrepôt de données grâce au provisionnement accéléré que permet le Cloud. En l'absence d'investissements dans une infrastructure, vous bénéficiez d'une véritable agilité opérationnelle. En effet, vous achetez ce que vous souhaitez quand vous le souhaitez. C'est vous qui décidez.

Dans le cadre d'un service Cloud entièrement géré, les aspects quotidiens de la maintenance de dashDB sont gérés pour vous. Ces opérations sont limitées à l'application ou non de tel ou tel ensemble de correctifs (vous n'êtes mis au courant de leur mise en œuvre que lorsque vous recevez la documentation correspondante), mais concernent aussi les versions. Avec dashDB, nos ingénieurs et développeurs dédiés ajoutent en permanence des fonctionnalités et des éléments de compatibilité et d'intégration au produit. Pour vous, ces nouveautés sont intégrées automatiquement.

### Protéger vos données avec dashDB

Depuis la conception jusqu'au déploiement, dashDB est optimisé pour assurer la sécurité de vos données. La solution est dotée de plusieurs couches de sécurité, en particulier le chiffrement automatique des données statiques et en transit, la surveillance des activités avec IBM InfoSphere Guardium et une protection renforcée du contrôle des accès et du déploiement de la base de données.

La sécurité s'appuie d'abord sur la conception et le développement de meilleures pratiques. dashDB a été développé en intégrant notamment les pratiques d'évaluation du risque, de modélisation des menaces et d'analyse du code statique et dynamique avec IBM AppScan.

### Développer rapidement vos activités de cœur de métier plutôt que de consacrer du temps à vos entrepôts de données

La simplicité du Cloud est un aspect essentiel. dashDB est un service entièrement géré et vous connaissez déjà certains des avantages inhérents au Cloud, notamment l'installation automatique des ensembles de correctifs et des versions, le provisionnement rapide des ressources dans le Cloud et l'agilité des activités. En bénéficiant de dashDB sous la forme d'un service, vous payez pour le service que vous utilisez. Ce qui diffère considérablement de l'achat d'un entrepôt de données complet. Avec dashDB, si le volume des données à traiter augmente, vous payez pour les capacités supplémentaires nécessaires, de manière simple et maîtrisée.

Conjointement à la simplicité de sa tarification, le fait que cette solution soit un service et non un cluster matériel ou une appliance vous aide à vous développer en fonction des demandes de votre activité. En résumé, ni achat, ni installation, ni test de logiciels et de matériels nouveaux. Vous vous développez tout simplement en fonction de vos besoins, grâce au Cloud.

## Des performances intégrées avec la technologie in-memory pour produire des réponses rapides

Qu'est-ce qui différencie dashDB des autres entrepôts de données basés sur l'environnement Cloud ? Bien des éléments. D'abord la technologie BLU Acceleration d'IBM, qui constitue le cœur de la solution.

Il s'agit d'une technologie de base de données en mémoire qui assure des performances de pointe, mais sans les contraintes habituelles des solutions en mémoire. dashDB bénéficie ainsi de l'ensemble des avantages de la technologie IBM BLU :

- **Traitements avancés** : dashDB ne nécessite pas le stockage d'un ensemble de données complet en mémoire. Tout en assurant un traitement ultrarapide, la solution utilise une série d'algorithmes brevetés qui permettent de gérer de manière agile la gestion des données en mémoire.
- **Prélecture des données** : dashDB est conçu pour anticiper et effectuer une prélecture des données avant qu'elles soient requises. La solution s'adapte automatiquement pour stocker les données nécessaires dans l'unité centrale ou à proximité.
- **Pas de décompression** : dashDB préserve l'ordre des données et procède à un large éventail d'opérations, en particulier les jointures et les évaluations de prédicats sur les données compressées, et ce sans avoir besoin d'une décompression, ce qui accélère de manière considérable le traitement.
- **Qualification des données** : Pour le traitement du Big Data, il n'est généralement pas nécessaire, pour répondre à une requête particulière, de stocker *toutes les informations* dans l'entrepôt de données. À cet effet, la technologie BLU Acceleration de dashDB est conçue pour déterminer automatiquement les données non retenues pour une analyse avec une requête particulière. Il est donc possible d'ignorer d'importants volumes de données non concernées pour le traitement d'une requête, ce qui permet de réduire les délais et les ressources nécessaires.

## Les capacités MPP (traitements massivement parallèles) permettent d'accélérer les requêtes et de traiter des ensembles de données massifs

Les capacités MPP de dashDB s'appuient sur les avantages du service dashDB standard en apportant encore plus de rapidité et d'évolutivité, ce qui permet de traiter des ensembles de données encore plus volumineux. En effet, l'architecture MPP s'appuie sur un cluster de serveurs reliés en réseau et travaillant en parallèle pour accélérer l'exécution des requêtes. Dans le cluster MPP de dashDB, plusieurs serveurs travaillent simultanément sur la même requête, et le traitement d'une requête est en outre parallélisé sur l'ensemble des processeurs de chaque serveur.

Avec une architecture standard, la parallélisation n'est appliquée qu'au niveau du processeur. En revanche, dans une architecture MPP, une requête est scindée en différents éléments et les serveurs, qui disposent chacun de leurs propres ressources de stockage et capacités de traitement locales, traitent des éléments de données distincts. Cette action collective accélère de manière considérable le processus de traitement des requêtes et réduit les besoins d'entrées/sorties. À cet effet, chaque serveur individuel chargé de traiter une requête en mode MPP bénéficie de la technologie IBM BLU de stockage de colonnes en mémoire dynamique, ce qui permet encore de réduire les entrées/sorties et d'obtenir une accélération considérable des traitements par comparaison aux bases de données classiques de stockage de lignes.

Avec les traitements massivement parallèles, les gains de performances augmentent pour chaque nouveau serveur ajouté au cluster. Pour prendre un exemple, si une requête nécessite une heure avec une architecture standard équipée d'un serveur unique, elle ne prend qu'environ 15 minutes avec un cluster MPP équipé de seulement quatre serveurs. Si vous ajoutez un serveur supplémentaire, ce qui porte le total à cinq, le délai de traitement d'une requête passe à 12 minutes, puis à 10 minutes avec six serveurs, etc. Avec dashDB MPP, l'extension d'une configuration consiste tout simplement à ajouter des serveurs supplémentaires à votre cluster.

## Conçu pour l'analyse, afin de vous aider à comprendre vos données et votre activité

L'évolution des traitements en mémoire progresse de manière plus rapide que jamais du fait de l'augmentation substantielle des volumes de données auxquels les entreprises ont à faire face. Le matériel informatique et la mémoire s'étant banalisés, il est possible de stocker davantage de données en mémoire et de les y traiter.

Les entreprises avaient pris l'habitude d'attendre pour recevoir les rapports d'analyse produits par les entrepôts de données ; désormais, grâce aux nouvelles avancées des traitements en mémoire, ce n'est plus le cas. Les résultats sont aujourd'hui disponibles en temps réel pour permettre la prise de décision. Cette accélération permet également d'analyser de manière plus approfondie les résultats pour ceux qui souhaitent obtenir des détails supplémentaires.

En effet, un entrepôt de données est aujourd'hui utilisé comme un magasin de données principal pour l'analytique. Il est donc logique d'intégrer les outils d'analyse dans un entrepôt de données plutôt que de transférer les données pour les analyser ailleurs.

## L'analyse en base de données pour gagner en efficacité et en performance

Outre les performances de la technologie BLU Acceleration, dashDB bénéficie également de l'intégration du logiciel IBM Netezza Analytics pour des analyses avancées en base de données. Il s'agit de la même technologie que celle utilisée par les appliances IBM Netezza et les systèmes IBM PureData for Analytics.

dashDB dispose ainsi d'innombrables algorithmes de modélisation prédictive *directement intégrés dans la base de données*. Ces algorithmes sont disponibles partout où vous souhaitez les utiliser.

À titre d'exemple, parmi les algorithmes intégrés dans dashDB, figurent :

- Les régressions linéaires,
- La classification des arbres de décision,
- Le partitionnement en k-means,
- Les extensions géospatiales compatibles Esri.

En exécutant les analyses de manière native dans la base de données, là où se trouvent les données, votre entreprise peut gagner considérablement en efficacité. Plutôt que d'extraire les données, les transférer ailleurs, les organiser, puis les traiter, vous les conservez là où elles se trouvent (dans l'entrepôt de données), puis vous leur appliquez directement les analyses.

## Compatibilité avec des outils avancés comme R et IBM Watson Analytics

R est un langage de programmation open source développé pour l'analyse de données avancée et la visualisation graphique. Il est utilisé pour analyser les données issues de différentes sources, notamment les fichiers externes ou les bases de données. dashDB intègre le langage R pour la modélisation prédictive en s'appuyant sur un moteur d'exécution pour le traitement des données. Une Web console est également proposée pour charger les données et procéder à des analyses en quelques minutes. Les analyses de données utilisent les procédures SQL, les outils BI ou les scripts et modèles R. Avec dashDB et le langage R open source, vos options d'analyse sont vastes et variées.

IBM dashDB est proposé avec RStudio, environnement de développement R intégré, conçu pour créer des analyses prédictives rapides basées sur ce langage. RStudio dispose d'une fonction de production de code en langage R, d'une aide intégrée pour les packages R et de capacités de gestion de fichiers, parmi de nombreuses autres fonctionnalités. En outre, l'installation de packages R supplémentaires peut être facilement réalisée depuis RStudio.

dashDB a été développé en prenant en compte l'écosystème le plus large d'outils de business intelligence. Le service dashDB fonctionne de manière native avec les principales technologies d'IBM, comme IBM Watson Analytics, IBM Cognos Analytics, IBM DataWorks et d'autres, mais sans s'y limiter. En effet, dashDB a été conçu pour intégrer les solutions d'innombrables partenaires commerciaux d'IBM et les nombreux outils BI, dont Looker, Aginity Workbench, Tableau et bien d'autres.

L'écosystème des partenaires IBM se développe rapidement et vous pouvez associer de très nombreux outils tiers au service dashDB. Vous pouvez par exemple connecter :

- IBM InfoSphere Data Architect pour concevoir et déployer votre schéma de bases de données,
- Esri ArcGIS pour effectuer des analyses et des cartes géospatiales avec vos données,
- Un serveur IBM Cognos pour créer des rapports Cognos utilisant vos données,
- Des outils SQL tels que Tableau, Microstrategy ou Microsoft® Excel pour manipuler ou analyser vos données,
- Les applications Bluemix qui nécessitent une base de données analytique,
- Aginity Workbench pour assurer la migration des modèles et des données Netezza vers dashDB.

## Cas d'utilisation

Les entreprises peuvent tirer parti de dashDB dans de nombreuses situations. Nous en proposons quatre ici, qui correspondent à des clients actifs aujourd'hui.

### 1. Enrichissement d'un entrepôt de données existant – Hybride

Nous constatons que plus de 90 % des clients prévoient d'étendre leur entrepôt de données existant<sup>1</sup>. dashDB est idéalement structuré pour cette situation. Avec dashDB, les entreprises peuvent étendre leurs environnements d'entrepôts de données on-premise grâce au Cloud. La facturation étant calculée en fonction de la capacité dont vous avez besoin aujourd'hui, la plateforme est élastique et vous n'y accédez que lorsque vous en avez besoin. Avec dashDB, vous n'êtes pas obligé d'acheter des services avant d'en avoir besoin.

### 2. Analyses de données NoSQL

Le deuxième cas concerne l'intégration des outils IBM entre Cloudant et dashDB. vous pouvez facilement synchroniser les données JSON traitées par la plateforme Cloudant avec les données structurées gérées par dashDB. Ce processus permet ensuite d'utiliser les outils BI et d'analyse traditionnels dans les entrepôts de données. dashDB disposant d'algorithmes prédictifs en base de données, les clients utilisateurs de la plateforme Cloudant peuvent analyser leurs bases de données de documents JSON sans difficulté.

### 3. Magasin pour l'analyse de données

Illustration spécifique de l'idée d'analyse en base de données, le troisième cas est détaillé ici. Pour les scientifiques spécialisés dans les statistiques ou les données, dashDB offre un ensemble robuste d'algorithmes d'analyse prédictive. Comme indiqué précédemment, dashDB est proposé avec un moteur d'exécution R et le logiciel RStudio. Le langage R est très largement utilisé par les statisticiens et les spécialistes du data mining pour développer des logiciels statistiques et des sorties d'analyse de données.

### 4. Entrepôt de données autonome sous la forme d'un service

dashDB est très souvent utilisé comme entrepôt de données autonome dans le Cloud. Qu'il s'agisse d'un petit « datamart » de démarrage dérivé de données Cloudant, d'un environnement de test ou de développement de base, ou même d'un entrepôt de données d'entreprise fonctionnant dans le Cloud, dashDB est disponible 24 heures sur 24 pour vous et pour votre entreprise.

## Démarrer avec dashDB

IBM dashDB est un service d'entrepôts de données entièrement géré : par conséquent, vous n'avez ni à acheter ni à installer de matériel et de logiciel. Il vous suffit d'ouvrir un compte sur le site [www.dashdb.com](http://www.dashdb.com) pour commencer à l'utiliser gratuitement. Si vous avez besoin d'aide pour vous former à dashDB ou déterminer si la solution est bien adaptée aux besoins de votre entreprise, n'hésitez pas à nous contacter en ligne sur le site [dashdb.com](http://dashdb.com). Nous sommes à votre disposition pour répondre à vos questions ou vous permettre de dialoguer de manière approfondie avec notre équipe technique.

## À propos des solutions IBM dashDB

IBM propose la gamme la plus complète de logiciels, de matériels et de solutions d'entrepôts de données, de gestion de l'information et de logiciels d'analyse métier pour aider ses clients à maximiser la valeur de leurs informations et découvrir de nouvelles connaissances utiles destinées à prendre des décisions plus pertinentes et rapides et à optimiser leurs résultats.

## Pour plus d'informations

Pour en savoir plus sur dashDB, contactez IBM ou un partenaire commercial IBM. Vous pouvez également consulter le site web suivant : [www.dashdb.com](http://www.dashdb.com)

De plus, IBM Global Financing peut vous permettre d'acquies plus facilement la solution dont vous avez besoin. Les clients éligibles peuvent bénéficier d'une solution de financement (prêt, location) adaptée à leurs objectifs et à leur stratégie financière et budgétaire. IBM Global Financing facilite vos investissements IT et contribue au succès de votre entreprise.

Pour en savoir plus, visitez [ibm.com/financing/fr](http://ibm.com/financing/fr)



---

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

IBM, le logo IBM, ibm.com, AppScan, Bluemix, Cloudant, Cognos, dashDB, Guardium, IBM PureData, IBM Watson et InfoSphere 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).

Netezza est une marque ou une marque déposée d'IBM International Group B.V., une entreprise IBM.

Microsoft est une marque de Microsoft Corporation aux États-Unis et/ou dans d'autres pays.

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 tous les pays où IBM est présent.

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 de la garantie décrite dans les conditions générales des contrats dans le cadre desquels ils sont fournis.

Il appartient au client de respecter les lois et réglementations applicables. IBM ne fournit aucun avis juridique et n'assume en aucun cas que ses produits ou ses services garantissent le respect par le client des lois et réglementations en vigueur.

© Copyright IBM Corporation 2015

1 « Predicts 2014: Why You Should Modernise Your Information Infrastructure », 28 novembre 2013. Gartner.



Veillez recycler

---