

Calcul intensif de données pour les grands ensembles de données (Big Data) avec IBM Spectrum Symphony



Améliore la gestion des données structurées et non structurées

Points forts

- Promeut une architecture à faible latence ;
 - Permet une abondance des ressources ;
 - Prend en charge plusieurs applications et différents types de charges de travail mélangés exécutés dans le même cluster
 - Planifie les charges de travail selon la priorité.
-

IBM® Spectrum Symphony Advanced Edition est un moteur d'exécution distribué aux entreprises pour les applications MapReduce. Il est conçu pour fournir des fonctionnalités prêtes à l'emploi en production telles qu'une disponibilité élevée des ressources et une grande prévisibilité, une prise en charge de plusieurs applications et systèmes de fichier, une maturité opérationnelle, une politique de contrôle des accords de niveau de service (SLA) et une utilisation intense des ressources pour les applications MapReduce. Fort de nombreuses années de planification des charges de travail distribuées et d'expérience de gestion développées par les professionnels IBM Spectrum Computing, IBM Spectrum Symphony offre des services modernes visant à exécuter des charges de travail distribuées pour vos applications MapReduce.

Avec l'explosion du nombre de données disponibles, les entreprises luttent contre la gestion et le traitement des données structurées et non structurées. En effet, environ 80 % des données de certaines entreprises sont des données non structurées. Pour extraire des valeurs à partir de ces données non structurées, il faut de nouveaux systèmes et de nouveaux processus. Utiliser le Hadoop File System (HDFS) et le système Hadoop MapReduce est une structure logicielle populaire pour répondre à ce besoin croissant. IBM fournit un système, tel que celui-ci avec IBM InfoSphere BigInsights, qui est adapté à l'utilisation des entreprises. BigInsights fournit de nombreuses améliorations supplémentaires dans la gestion et les analyses que les entreprises trouvent utiles dans le cadre d'une stratégie Smarter Analytics.



IBM Spectrum Computing

Description des solutions

Pour certaines entreprises, le besoin d'acquies ces systèmes va bien au-delà du traitement, en exécutant les applications Hadoop dans leur infrastructure informatique distribuée. Ces entreprises requièrent un système pouvant traiter les requêtes ou les tâches avec un temps de réponse extrêmement rapide (inférieur à la milliseconde (ms)). Elles souhaitent exploiter le cluster distribué pour plusieurs types d'applications, et pas uniquement pour les applications Hadoop. De plus, de nombreuses entreprises veulent partager l'application avec plusieurs sous-équipes, chacune ayant son propre accord de niveau de service (SLA) et sa propre affectation des ressources au sein de l'environnement partagé. IBM Spectrum Symphony offre une plateforme informatique distribuée qui fournit un support d'application hétérogène dans un seul cluster. Les applications exécutées dans cet environnement sont pourvues d'une latence extrêmement faible et l'architecture mutualisée fournie permet une meilleure affectation des ressources et des accords de niveau de service (SLA) à différentes sous-équipes partageant l'environnement.

Architecture à faible latence

IBM Spectrum Symphony utilise des techniques d'optimisation qui permettent une mise à disposition des charges de travail en un temps inférieur à la milliseconde. Cela implique que l'exécution des tâches courtes affiche un pourcentage de temps inférieur dans les étapes d'approvisionnement et de désapprovisionnement, procurant ainsi des tâches utiles supérieures pour dépasser le pourcentage. Il a également un taux de production élevé grâce auquel le système peut soumettre plus de 5 000 tâches par seconde. Les systèmes « open source » similaires sont deux ordres de grandeur plus lents que la publication actuelle.

Planificateur politique de charges de travail

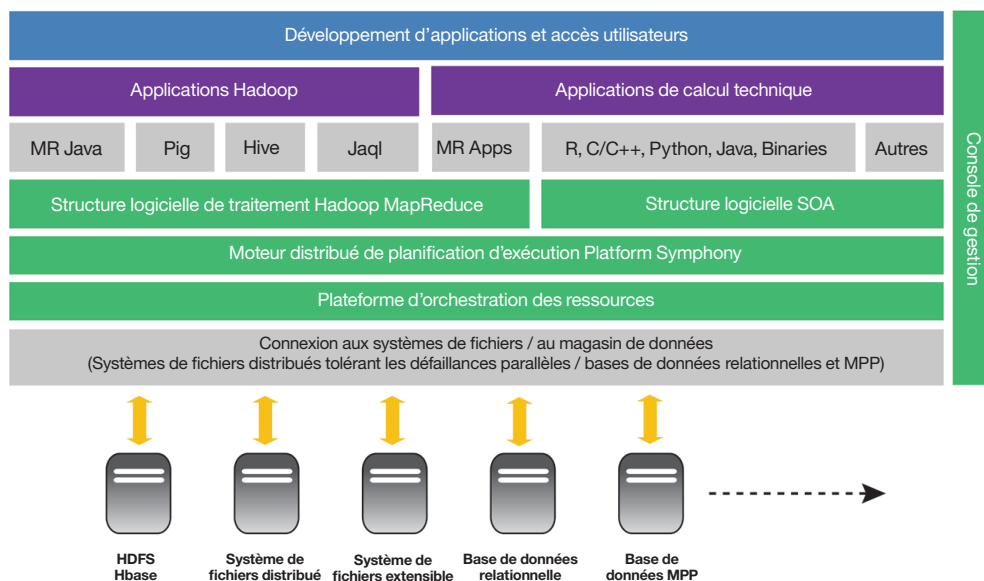
Le planificateur politique de charges de travail dans IBM Spectrum Symphony fournit 10 000 niveaux de priorité et prend en charge plusieurs tâches MapReduce exécutées en parallèle. Ce planificateur politique de charges de travail permet la mise en œuvre des ressources prioritaires pour les tâches préventives, ainsi qu'une planification partagée juste des tâches Mapper et Reducer, toutes effectuées au niveau des tâches pour fournir une meilleure granularité et un meilleur contrôle.

Abondance des ressources

IBM Spectrum Symphony assure une disponibilité dans le moteur d'exécution distribué ; il n'y a donc aucun point de défaillance. Il permet un basculement automatique des suivis de travaux et de tâches, et une reprise de travail sans redémarrage. Pour le système de fichiers Hadoop, IBM Spectrum Symphony offre un basculement automatique du NameNode au sein du système de fichiers Hadoop Distributed et permet une reprise du système de fichiers et une reprise de la tâche associée.

Architecture ouverte pour le développement et le choix des applications

IBM Spectrum Symphony est construit sur une architecture ouverte afin de prendre en charge plusieurs applications MapReduce, compatibles à 100 % avec les applications Hadoop pour les tâches MapReduce basées sur Java™. La technologie d'adaptation des applications intégrée au produit fournit une intégration harmonieuse des applications avec IBM Spectrum Symphony afin que les tâches créées avec la technologie Hadoop MapReduce (Java, Pig, Hive et autres) ne nécessitent aucune modification dans la logique de programmation pour être exécutées avec IBM Spectrum Symphony. Cette architecture ouverte fournit également une méthode pour tirer profit de plusieurs types de systèmes de fichiers, ainsi que des architectures de base de données. IBM Spectrum Symphony prend entièrement en charge HDFS, IBM Spectrum Scale ainsi que d'autres types de systèmes de fichiers distribués et d'autres types de données. De plus, pour les processus MapReduce, le type de système de fichiers de source de données d'entrée peut être différent du système de fichier de source de données de sortie. Cela permet de prendre en charge plusieurs utilisations, y compris une logique de flux d'extraction, de transformation et de chargement (ETL).



Architecture IBM Spectrum Symphony Advanced Edition

Prise en charge de plusieurs applications MapReduce

IBM Spectrum Symphony Advanced Edition intègre l'implémentation MapReduce compatible avec Apache Hadoop et prend en charge jusqu'à 300 applications séparées (suivies de tâches) pour les charges de travail MapReduce, ainsi que d'autres types d'applications distribuées de façon simultanée. Cela permet aux clients de tirer profit à la fois des ressources existantes et nouvelles, et d'optimiser leur infrastructure informatique tout en gardant une interface de gestion unique.

Prise en charge pour les mises à niveau continues

IBM Spectrum Symphony MapReduce prend en charge plusieurs versions des applications MapReduce qui sont exécutées sur les mêmes clusters ; il n'est donc pas nécessaire de supprimer entièrement le cluster pour mettre à niveau le logiciel. Les serveurs exécutant des applications mises à niveau peuvent coexister avec la version précédente du produit sur d'autres nœuds permettant ainsi l'installation progressive des mises à niveau sur un ensemble de serveurs sans supprimer entièrement le cluster.

Fonctions de surveillance et de dépannage améliorées

IBM Spectrum Symphony MapReduce contrôle l'unité centrale (CPU) ainsi que le niveau d'utilisation de la mémoire et affecte les ressources en conséquence. Il est capable de tirer les données collectées des serveurs individuels et de les gérer à partir d'une interface unique.

Affinité des données IBM Spectrum Symphony MapReduce

IBM Spectrum Symphony MapReduce intègre des fonctions puissantes de corrélation des données afin d'améliorer considérablement la performance des applications et l'utilisation des ressources en prenant en compte la localité des données quand il planifie des charges de travail MapReduce. Sa solution de corrélation des données supprime presque le temps passé à accéder à des volumes de données importants requis par les applications MapReduce. Cela augmente considérablement la performance de l'ensemble des applications grâce à un accès plus rapide aux fichiers.

Pourquoi choisir IBM ?

IBM Spectrum Computing offre un portefeuille complet de solutions d'infrastructure définie par logiciel qui aide votre entreprise à fournir des services informatiques de la manière la plus efficace possible, en optimisant ainsi l'utilisation des ressources pour accélérer la génération des résultats et réduire les coûts. Grâce à ces offres, vous optimisez le potentiel de votre infrastructure pour accélérer les applications d'analyse, de calcul hautes performances (HPC), Apache Hadoop, Spark et natives Cloud à n'importe quelle échelle, extraire des perspectives de vos données et commercialiser plus rapidement des produits de meilleure qualité.

Qu'elles soient déployées dans un centre informatique ou sur le cloud, les solutions IBM Spectrum Computing soutiennent le développement des produits, les décisions commerciales stratégiques et favorisent l'émergence de perspectives inédites dans une multitude de secteurs (services financiers, fabrication, médias numériques, hydrocarbures, sciences de la vie, administration, recherche et éducation). Qu'il s'agisse de concevoir une Formule 1 ou d'analyser les risques liés au crédit, les entreprises d'une multitude de secteurs utilisent IBM Spectrum Computing comme base pour leurs solutions d'infrastructure définie par logiciel pour les grands ensembles de données (Big Data), les calculs hautes performances et le cloud afin d'améliorer leurs résultats.

Pour plus d'informations

Pour en savoir plus sur le calcul intensif de données pour les grands ensembles de données (Big Data) avec IBM Spectrum Symphony, contactez votre interlocuteur IBM ou votre partenaire commercial IBM, ou visitez : ibm.com/fr/systems/spectrum-computing/products/symphony/

En outre, IBM Global Financing propose de nombreuses options de paiement vous permettant d'acquérir la technologie nécessaire à la croissance de votre activité. Nous pouvons gérer l'ensemble du cycle de vie des produits et services informatiques, de leur acquisition à leur élimination. Pour obtenir plus d'informations, consultez : ibm.com/financing



Compagnie IBM France

17 avenue de l'Europe
92275 Bois-Colombes Cedex
France

IBM Ireland Limited enregistré en Irlande sous le n° 16226.
La page d'accueil d'IBM peut être consultée sur ibm.com/fr

IBM, le logo IBM, ibm.com/fr, IBM Spectrum, IBM Spectrum Scale, BigInsights, InfoSphere, Smarter Analytics et Symphony sont des marques commerciales ou déposées d'International Business Machines Corporation aux États-Unis et/ou dans d'autres pays. Les marques d'IBM accompagnées du symbole ® ou ™ présentes au moment de la publication de ces informations sont des marques déposées par IBM au registre des marques commerciales ou déposées appartenant à IBM, conformément aux lois en vigueur aux États-Unis. Ces marques peuvent également être enregistrées aux registres d'autres pays.

La liste actualisée des autres marques IBM est disponible sur Internet à la section « Informations sur le droit d'auteur et les marques » sur ibm.com/fr/legal/copytrade.shtml

Java et l'ensemble des marques et logos Java sont des marques commerciales ou déposées d'Oracle et/ou de ses filiales.

Les autres noms de société, de produits et de services peuvent être des marques déposées ou des marques de service appartenant à des tiers.

Les informations contenues dans cette publication concernent les produits, les programmes ou les services commercialisés par IBM France et n'impliquent nullement l'intention d'IBM de les commercialiser dans tous les pays où la marque est présente.

Toute référence à un produit, programme ou service IBM ne signifie pas pour autant que les produits, programmes ou services peuvent être utilisés. Tout produit, programme ou service fonctionnellement équivalent peut être utilisé à leur place.

Le matériel IBM peut contenir des composants neufs, ou un mélange de pièces neuves et remises en état. Dans certains cas, le matériel peut ne pas être du matériel neuf et peut avoir déjà été installé. Ceci ne modifie en rien le régime des garanties contractuelles IBM applicables.

Cette publication est uniquement donnée à titre indicatif. Ces informations peuvent faire l'objet de modifications sans préavis. Veuillez contacter votre point de vente ou votre revendeur IBM pour connaître les toutes dernières informations au sujet des produits et services IBM.

Cette publication contient des adresses Internet tierces. IBM ne peut être tenu responsable des informations publiées sur ces sites Internet.

IBM ne fournit aucun conseil juridique, comptable ou d'audit, et ne considère et ne garantit pas que ses produits ou services sont conformes aux lois applicables. Les clients sont les seuls responsables de leur conformité avec les lois et réglementations de sécurité en vigueur, y compris les lois et réglementations nationales.

Les photographies présentées dans ce document peuvent représenter des maquettes.

© Copyright IBM Corporation 2016



Veuillez recycler