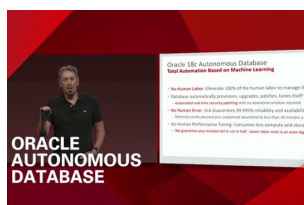


Oracle Autonomous Database tient-il ses promesses ?

Xavier Biseul
JDN

Mis à jour le 09/04/19 13:10



Près d'un an et demi après son annonce, comment le marché a réagi à la promesse d'Oracle d'une base de données autonome, capable de s'administrer par elle-même ?

1^{er} octobre 2017. Larry Ellison introduit la session inaugurale d'Oracle Openworld.

Comme à son habitude, le cofondateur d'Oracle fait le show. Il annonce le lancement d'une base de données capable de se gérer par "elle-même" grâce à ses mécanismes d'apprentissage automatique (voir [la vidéo](#)). Basée sur la version 18c de la base de données maison et sur la plateforme intégrée Exadata dans le cloud, Oracle Autonomous Database assure automatiquement les tâches de provisioning, de sauvegarde, d'application de correctifs, de mises à jour et autres opérations de maintenance.

Larry Ellison garantit aux entreprises clientes une disponibilité à 99,995%, sans aucune administration. Ce qui facilite la définition de tables et le chargement de données tout en fournissant une compression, une mise en cache et une indexation automatiques.

En réduisant les interventions, une base de données autonome limite voire élimine, selon lui, les erreurs humaines tout en permettant aux administrateurs de bases de données (DBA) de se consacrer à des tâches à plus haute valeur ajoutée. L'apprentissage automatique peut également renforcer la cyber-protection en détectant des comportements suspects comme des valeurs aberrantes ou des schémas anormaux au sein d'une base de données. L'application automatique de correctifs permet, enfin, de lutter plus efficacement contre les failles "zero day".

La nécessaire évolution du rôle du DBA

Comment ce concept de base de données autonome a été reçu par le marché ? Près d'un an et demi après l'annonce de Larry Ellison, le JDN a interrogé des partenaires et clients d'Oracle. Premier constat, peu d'entreprises ont testé Autonomous Database et encore moins l'ont déployé. En revanche, la promesse d'une base de données autonome semble répondre aux attentes du marché. "Cela rejoint l'état de réflexion de mes clients, observe Éric Leygonie, directeur technique régional d'Easyteam à Toulouse. S'ils acceptent de basculer leur base de données dans le cloud, c'est pour ne plus s'en occuper. La base doit s'administrer toute seule et s'adapter à leurs besoins en faisant de l'auto-scaling."

Practice manager à Business & Decision de Bordeaux, Renaud Durand partage ce constat. "Sur mon territoire de La Nouvelle Aquitaine, les entreprises clientes qui réfléchissent à migrer dans le cloud sont surtout de belles PME et des sociétés en forte croissance pour qui maintenir et administrer des serveurs dans leurs locaux peut être compliqué. La promesse d'Oracle résonne à leurs oreilles." Il constate, par ailleurs un reflux d'Hadoop et des bases de données NoSQL. "Beaucoup d'entreprises ont cru qu'elles pourraient remplacer leur datawarehouse. Ce n'est pas si simple et on assiste à un regain d'intérêt pour les bases relationnelles, mais dans des versions modernisées."

Et si Oracle est monté dans le cloud en retard par rapport à d'autres fournisseurs, le géant américain reste, à ses yeux, une référence du marché en matière de base de données avec une technologie éprouvée et une base installée conséquente. Avec ce type de solution, "Oracle veut se différencier de ses concurrents frontaux dans le cloud que sont Snowflake, AWS Redshift ou Azure SQL Data Warehouse", estime Renaud Durand. Reste le rôle dévolu au DBA. Est-il appelé à disparaître avec l'arrivée des bases de données autonomes ? Éric Leygonie qui a fait toute sa carrière sur les solutions d'Oracle a déjà vu son métier évoluer. "En tant que prestataire, les missions portent plus sur de l'architecture de base de données que de l'administration en régie."

Sauermann : un POC sur un cas d'usage de BI

En ce qui concerne les entreprises interrogées, elles sont pour l'heure au stade de la découverte des possibilités d'Autonomous Database. Entreprise internationale spécialisée dans la qualité de l'air en milieu confiné (plus de 400 employés), Sauermann a initié un POC d'au moins 3 mois sur un cas d'usage restreint de business intelligence. L'expérimentation porte sur les données d'analyse de performance des ventes à l'échelle du groupe. Provenant de différentes sources et notamment des ERP, les données seront consolidées au niveau d'Autonomous Database d'Oracle. Objectif : construire un dataset dédié au reporting orienté client et au décisionnel avec Microsoft PowerBI pour la couche analytique.

"En restreignant le POC à ce que l'on connaît, il sera plus aisé de comparer la fiabilité des tableaux de bord avec ceux que le marketing construit aujourd'hui de façon plus artisanale, mais aussi les temps de traitement, avance Mejdi Bettaieb, DSI de Sauermann. Si le rendu fonctionnel et les performances techniques sont au rendez-vous, nous étendrons le périmètre à d'autres axes d'analyse de BI."

A plus long terme, Sauermann envisage de croiser les données brutes collectées par sa prochaine génération d'appareils de mesure de la qualité de l'air avec ses données de gestion. "Nous miserons sur Autonomous Database pour absorber les différentes couches de données et faire des analyses croisées, poursuit Mejdi Bettaieb. A partir du numéro série de l'appareil intelligent, on pourrait rapprocher les paramètres d'étalonnage et d'usage à l'historique des réclamations clients enregistrées au niveau de la base installée gérée dans Oracle Service Cloud, ceci nous permettrait de remonter à l'origine des déviations de mesures et proposer proactivement des services de maintenance préventive."

La DSI de Sauermann ayant des ressources limitées, Mejdi Bettaieb préfère que ses équipes se concentrent sur les besoins métiers plutôt que passer leur temps sur des tâches d'administration entre sizing, patching, monitoring et sécurité. "La mission classique d'un DBA n'apporte pas de valeur ajoutée au business, c'est perçu comme un mal nécessaire. La finalité ce n'est pas de stocker et maintenir les données mais de les exploiter correctement pour permettre de prises de décisions rapides et éclairées."

En ce qui concerne la gestion des règles de sécurité, le DSI estime qu'il faut faire appel au machine learning, déjà intégré au niveau d'Autonomous Database, pour assurer un niveau de sécurité très élevé et protéger des données clients sensibles, surtout depuis l'entrée en vigueur du RGPD. "La vigilance de l'humain est dépassée par la variété des failles potentielles. Aujourd'hui, c'est machine contre machine."

Enfin, sur le plan financier, l'approche d'Autonomous Database permet, selon lui, d'avoir une maîtrise de l'investissement à consentir. "Pas besoin de surdimensionner la plateforme, elle monte en puissance en fonction des



besoins. Cette flexibilité est un point clé car il est difficile d'estimer à l'avance la volumétrie des ressources à provisionner pour ce type de projet de BI. Une fois en production, il sera possible d'ajuster à notre rythme le seuil de nos crédits mensuels en fonction de notre consommation réelle des ressources cloud."

Le Cern : un stockage réduit dans un facteur dix

Autre retour d'expérience, celui du Cern. L'organisation européenne pour la recherche nucléaire, connu pour disposer du plus grand accélérateur de particules au monde (LHC), a été un des premiers bêta-testeurs d'Autonomous Data Warehouse (ADW), une déclinaison de l'offre de base de données autonome d'Oracle. La gestion des données est un enjeu clé pour le Cern. Les plus de deux millions de capteurs de contrôle et d'ingénierie répartis sur le LHC génèrent 250 To de données par jour. Sans parler des données provenant d'autres expériences menées par l'organisme de recherche dont Atlas qui vise à "explorer les secrets de l'univers."

Le système actuel repose sur le SGBDR Oracle on-premise et le Cern déploie des efforts importants pour faire évoluer ses infrastructures afin de répondre à une volumétrie de données en perpétuelle progression. "Dans ce contexte, nous voulions challenger ADW sur ses capacités à intégrer différents jeux de données venant de l'IoT, à améliorer les performances en termes de récupération et d'analyse de données et à débloquer des cas d'usage impossibles aujourd'hui", explique Manuel Martin Marquez ingénieur big data senior et data scientist au Cern.

Pour cette évaluation, le Cern a décidé de prendre Oracle au mot. Il n'a pas facilité la tâche à la solution autonome en ne la renseignant pas sur des éléments essentiels de structuration de ses bases de données locales, tels que la table IoT, les partitions et les index. Première découverte : grâce à l'utilisation automatisée de fonctionnalités avancées telles que la compression hybride Oracle (Hybrid Columnar Compression, HCC), le stockage requis a été réduit dans un facteur de dix environ. "Ce qui s'est aussi traduit par une réduction du temps passé à récupérer ou à analyser certaines données", complète Manuel Martin Marquez.

Le Cern a également évalué d'autres aspects liés à l'automatisation de services tels que le provisionnement, les mises à jour, les correctifs et les sauvegardes. Dans ce domaine, "Oracle a réalisé un excellent travail, tranche notre expert. Toutes ces tâches nécessitent généralement des connaissances approfondies. En mode autonome, vous pouvez les oublier ou y consacrer très peu de temps. En plus de l'interface web, Oracle fournit des outils de lignes de commande et des API qui permettent d'aller encore plus loin dans l'automatisation de ces tâches."

Manuel Martin Marquez se dit convaincu, qu'avec ce type de technologies, le DBA sera définitivement libéré de certaines de ses tâches actuelles. A l'image de ce qui s'est passé pour l'ingénieur systèmes, son rôle va évoluer, notamment vers l'analyse de données. Pour autant, "il y aura toujours des bases de données critiques ou aux exigences particulières nécessitant une gestion très minutieuse de la part de DBA experts". Enfin, le spécialiste de la data salue l'annonce d'Oracle que le service de base de données autonome serait disponible sur site. Une option qu'il juge intéressante quand l'hébergement en cloud public n'est pas possible pour des raisons de latence ou de réglementation.