

LeMagIT.com

LHC : le CERN teste un data warehouse autonome

par Alain Clapaud

C'est sans conteste le joyau de la recherche européenne sur la physique des particules. À cheval sur la France et la Suisse, l'anneau de 27 km du LHC est le plus puissant accélérateur de particules au monde. Sur sa circonférence, l'accélérateur dispose de quatre grands détecteurs de natures différentes : ATLAS, CMS, ALICE et LHCb. Rien que le CMS et ses 150 millions de capteurs traitent jusqu'à 40 millions de collisions par seconde.

Lorsqu'il est en service, le LHC génère chaque année une trentaine de Pétaoctets de données scientifiques. Pour stocker et les mettre à disposition des chercheurs et des centres de calcul de tous les pays, le CERN a créé une énorme infrastructure de stockage, le CASTOR (pour CERN Advanced STORAGE Manager). Depuis 2005, le volume de CASTOR grimpe en flèche : au début de l'année 2019, il proposait 617 millions de fichiers, soit un volume de 342,4 Po sur disques et cartouches.

CASTOR peut globalement se résumer à une architecture de données hiérarchique dotée d'API pour récupérer les fichiers correspondants aux expériences qui intéressent les chercheurs.

Mais outre ces données, le LHC – qui est actuellement le plus grand instrument scientifique de la planète – produit aussi énormément de données de fonctionnement.

Des clusters Oracle RAC pour stocker les données au plus près de l'anneau

Le LHC est en effet une énorme installation industrielle que les techniciens et les ingénieurs doivent piloter et entretenir en permanence.

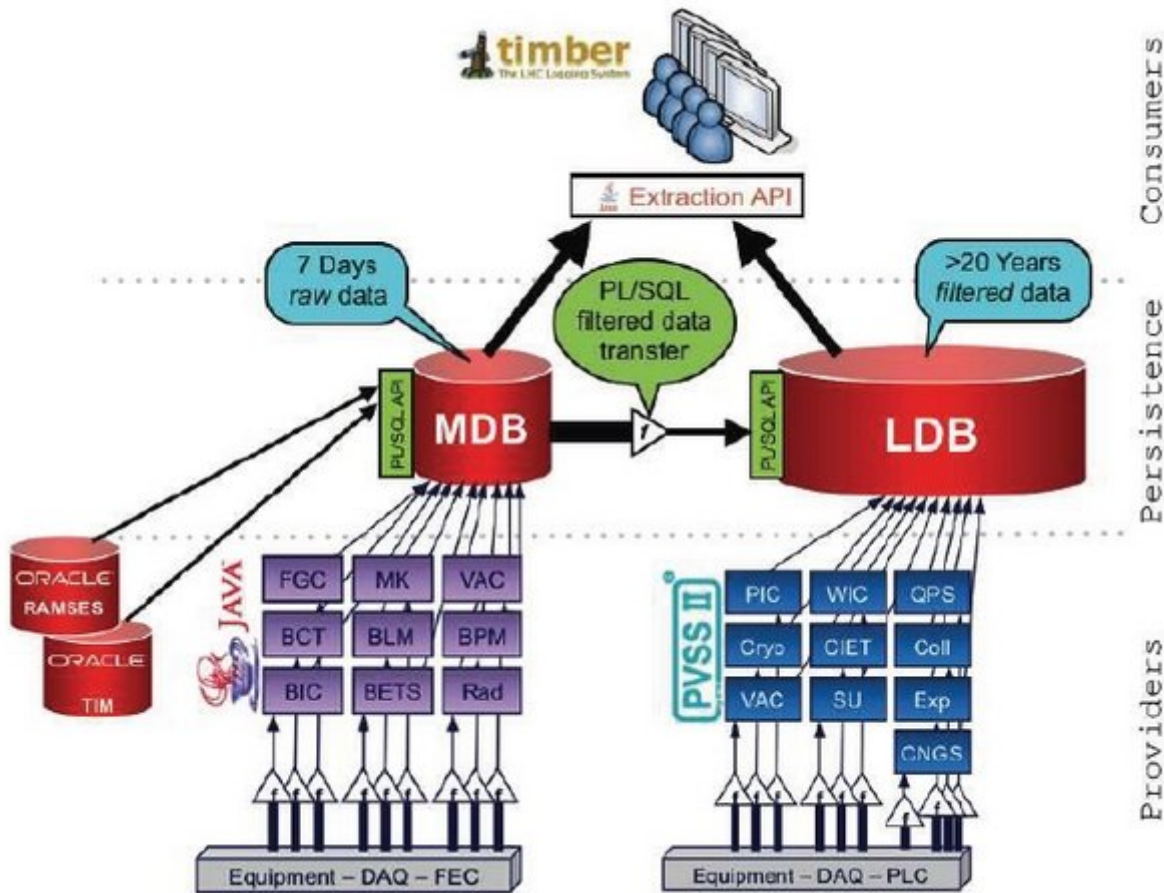
En plus « des 30 Po de données générées chaque année par les collisions, nous avons 2,5 To de données de contrôle pour chaque jour de fonctionnement » résume Sébastien Masson, DBA Oracle au CERN.

Ces données de fonctionnement viennent des milliers d'équipements des quatre grands détecteurs, mais aussi des systèmes de refroidissement, de ventilation et d'alimentation électrique. Elles sont indispensables au pilotage en temps réel de l'installation, mais aussi à la maintenance de ce gigantesque mécano industriel.

Si beaucoup de ces données sont filtrées (jusqu'à 95 % d'entre elles sont éliminées), beaucoup sont conservées dans une architecture dédiée – le CALS (CERN Accelerator Logging Service) – qui se compose de deux bases de données principales : MDB et LDB.

MDB (pour Measurement DataBase) joue le rôle de buffer. Ce cluster Oracle RAC (Real Application Cluster) récolte toutes les données de fonctionnement émises par les équipements à haute fréquence. Le cluster les traite, les filtre puis les envoie dans la LDB (Logging DataBase), un autre cluster RAC qui va historiser ces données. LDB peut aussi recevoir en direct les données qui ne nécessitent pas de filtrage.

MDB stocke environ une semaine de fonctionnement du LHC. LDB, lui, héberge actuellement 1 Po, soit un historique de plus de 20 années de fonctionnement.



CERN

accelerator logging service architecture – Les clusters Oracle RAC sont toujours utilisés pour la collecte et l’analyse temps-réel des données issues des équipements techniques du LHC. Hadoop intervient pour les analyses plus complexes nécessitant le traitement d’historiques de données plus longs. ©CERN

« L’atout de cette architecture est sa relative simplicité », résume le DBA. « LDB et MDB reposent toutes deux sur des clusters Oracle RAC, une technologie robuste qui nous permet de mener des analyses temps réel sur les données de quelques jours ou quelques semaines de fonctionnement. Les performances obtenues sont satisfaisantes, ce qui permet d’utiliser ce système pour le contrôle en temps réel de l’accélérateur. »

Mais l’inconvénient d’une architecture qui s’appuie sur une base relationnelle reste sa faible capacité à pouvoir scaler de manière horizontale. Dit autrement, l’architecture est performante pour accéder à des volumes de données portant sur quelques jours de fonctionnement, mais il montre ses limites lorsqu’il s’agit d’analyser des tendances sur plusieurs mois, voire plusieurs années.

Un ingénieur du CERN rapporte qu’il faut 12 heures de traitement pour extraire 24 heures de données de fonctionnement de CALS. « C’est une architecture qui a été conçue il y a de nombreuses années maintenant ; elle stocke aujourd’hui pratiquement 100 fois plus de données que ce pour quoi elle avait été conçue au départ. »

Après le relationnel : Hadoop et un Data Warehouse Cloud

Pour mener des analyses impliquant de plus gros volumes de données, le CERN s’est tout naturellement tourné vers Hadoop. Le but était de retrouver une plus grande liberté quant aux analyses de données et d’aller jusqu’au Machine Learning.

Un projet de nouvelle plateforme – baptisé NXCALS (pour Next CALS) – a été lancé afin de remplacer l’infrastructure Oracle par une plateforme Big Data s’appuyant sur différentes briques bien connues : Kafka, Hadoop, HBase, Spark, Parquet. En outre, au moment de l’étude de NXCALS, le CERN disposait d’un [cloud privé OpenStack](#) comptant 250 000 cœurs – une infrastructure idéale pour porter le cluster Hadoop du LHC.

Mais NEXCALs n'a pas remplacé CALS et le complète pour les analyses nécessitant de brasser beaucoup de données.

Une autre voie d'avenir semble aujourd'hui se dessiner au travers de l'[Openlab du CERN](#), un programme via lequel les chercheurs ont accès aux technologies de multiples partenaires technologiques tels que Google, Micron, Intel, Siemens et Oracle.

Ce dispositif a permis au CERN d'être l'un des tout premiers à tester le service Oracle Autonomous [Data Warehouse](#) (ADW), une infrastructure d'entrepôt de données Exadata dans le cloud.



Sébastien Masson, DBA Oracle au CERN

Sébastien Masson raconte : « Avec Oracle RAC d'un côté et Hadoop de l'autre, nous avons deux systèmes asymétriques. L'arrivée d'Autonomous Data Warehouse a attisé notre curiosité. Elle devrait nous permettre de disposer d'un système qui nous permettrait à la fois l'analyse temps réel (comme avec notre plateforme RAC) et l'analyse approfondie de nos données historiques sur des volumes de données bien plus importants. »

Le CERN a donc testé la solution pour voir si ce [Data Warehouse dans le cloud](#) pouvait répondre à la fois à son besoin de visualisation de données en temps réel et à son exigence d'analyse Big Data. Résultat, « au niveau performances temps réel, ça fonctionne plutôt bien », constate Sébastien Masson. « Nous avons un niveau de performance comparable à notre ancien système RAC sur les scénarios représentatifs que nous avons pu tester jusqu'à aujourd'hui. L'analyse temps réel sur ADW est validée. »

Le DBA est assez surpris, car ADW est une approche différente d'une base de données classique. Inutile de créer des partitions, des index et d'optimiser finement les paramètres de la base pour en tirer la quintessence. « [En tant que DBA](#), c'est assez bizarre. La première fois que l'on charge des données dans Autonomous Data Warehouse, on doit oublier tout ce qu'on devait faire jusqu'à présent. Vous lancez un Data Pump en excluant tout ce que vous aviez fait jusqu'à présent, c'est assez perturbant, il faut bien le reconnaître ! Cela signifie que l'on peut aujourd'hui économiser tout l'aspect tuning qui est assez difficile à réaliser. En outre, il n'y a pas de changement à apporter au code existant. »

Côté restitution, les utilisateurs disposent de l'outil Oracle Analytics Cloud (OAC) pour analyser leurs données. « Nous mettons à disposition les données, puis les utilisateurs se connectent au cloud, lancent une instance OAC et ils sont autonomes dans la génération de leurs tableaux de bord. »

Pour l'analyse des données historiques, le CERN n'a pas chargé directement dans l'entrepôt ses 600 To de fichiers Parquet. Ces données sont copiées en l'état dans un Object Store. Du point de vue de la base de données, celles-ci sont alors vues comme une table externe dont la localisation des données se trouve dans l'Object Store, avec l'avantage de pouvoir encore y accéder via une requête SQL de manière transparente.

« Ces tables peuvent même être partitionnées et ne lire que les fichiers Parquet pertinents pour la requête », souligne le DBA qui ajoute qu'il « travaille actuellement en lien direct avec les développeurs d'Oracle Data

Warehouse à San Francisco [sur] la notion de partitionnement hybride, apparu en version 19. Dans cette approche, des tables peuvent avoir des données qui sont à la fois dans la base et dans l'Object Store, le tout accessible de manière transparente. C'est une caractéristique qui nous intéresse, car elle nous permettrait de réunifier nos deux approches temps réel/Big Data. »

La piste du Data Warehouse dans le Cloud tient la route, mais...

Pour le DBA, il est encore trop tôt pour affirmer que le CERN va décommissionner ses clusters RAC et Hadoop au profit d'une solution 100 % Open Source ou pour Oracle Autonomous Data Warehouse.

Le remplacement des vieux clusters Oracle RAC pose en effet une problématique typique du [Edge Computing](#). « Nous ne pourrions pas supprimer ces clusters facilement, car se pose le problème de la proximité de tous les équipements qui sont sur l'accélérateur et qui envoient leurs données à haute fréquence sur le réseau. C'est une contrainte technique forte. Donc je pense que, quelle que soit la solution choisie, nous aurons de toute façon besoin d'une base locale avant d'envoyer les données dans le cloud ; mais il est encore trop tôt pour l'affirmer totalement. »

Actuellement, le LHC est à l'arrêt pour sa seconde phase de maintenance et de modernisation (LS2 pour Long Shutdown 2). La reprise des expériences est prévue en 2021. Après l'arrêt suivant – le LS3 de 2024 à 2026 – le LHC devrait connaître une transformation majeure avec la mise en place du High Luminosity Large Hadron Collider (HL-LHC). Ce HL-LHC devrait faire encore s'envoler les volumes de données à traiter. L'équipe IT du CERN a jusqu'à 2026 pour s'y préparer.

19 mai 2020

All Rights Reserved, [Copyright 2007 - 2020](#), TechTarget | [Read our Privacy Statement](#)