

Saint-Gobain bâtit une usine Big Data en self-service

Afin d'accélérer les initiatives Big Data pour les différents métiers du groupe, la DSI de Saint-Gobain a lancé un "Big Data Central Service" destiné à toutes les entités du groupe. Cette offre de services mutualisée vise à proposer une "Fabric Big Data" ou "Usine Big Data", pour que les sociétés du groupe s'essayent à ces technologies, le plus simplement possible afin d'en percevoir concrètement les bénéfices.

En s'appuyant sur une expertise Big Data, il s'agit d'accompagner les projets métiers du tout le Groupe et de former les contributeurs. Fondée en 1665 en tant que Manufacture Royale de Glaces, Saint-Gobain est aujourd'hui le leader mondial de l'habitat et de la construction comptant plus de 190 000 dans 65 pays, et gérant des dizaines de filiales spécialisées. D'où l'intérêt d'une démarche à la fois centralisée et mutualisée, sous forme de self-service.

La quadrature du cercle Big Data

L'approche repose sur quatre piliers : **l'agilité de la solution** pour rapidement concrétiser l'idée à travers la réalisation d'un POC (proof of concept ou prototype évolué). Dispenser de **la formation** sur les concepts Big Data basiques, **l'intégration** pour éviter la propagation des données et faciliter le partage et l'apprentissage, et la **maîtrise des coûts**.

« Nous avons très vite choisi Hadoop, car ces technologies peuvent couvrir la plupart des situations et ne sont pas restreintes à un cas d'usage spécifique, » explique Eric Merlateau, responsable IT Infrastructure Services Design chez Saint-Gobain. *« Nous avons estimé que la distribution open source Cloudera convenait le mieux à nos attentes: c'est l'une des offres leaders mature et éprouvée, son logiciel Manager facilite l'administration des clusters, la solution est complète au-delà de Hadoop (Impala, Spark...), et l'éditeur propose un support professionnel. »*

Des services accessibles à chacun selon ses besoins

La plateforme installée dans le datacenter de Saint-Gobain a été conçue pour être partagée en mode multitenant, sur un modèle hébergé à la demande. La sécurité est assurée par un mécanisme Kerberos (protocole d'authentification réseau basé sur un chiffrement symétrique utilisant des clés secrètes), le tout en interne chez Saint-Gobain. Enfin, une feuille de route a été définie pour intégrer au fil de l'eau les nouvelles versions et les outils externes.

Trois services sont proposés. Avec **Discover**, l'utilisateur découvre les concepts du Big Data et peut imaginer les diverses opportunités pour son entreprise. **Le service Explore** offre la possibilité de lancer des expérimentations et de concevoir des POC. Enfin, **Travel** est destiné auxancements des projets Big Data en production et à leur supervision.

« Plusieurs types d'expérimentations ont été lancées. Parmi nos activités industrielles, Sekurit (vitrage automobile) Glass (verres fenêtres et baies) ou encore Certaineed (matériaux de construction) ont misé sur

l'optimisation de la qualité de la production. Côté BI, les projets ont tenté de se mesurer aux solutions BI en place, et de s'essayer à l'analyse prédictive aussi bien chez Point-P (matériel BTP), que chez KparK (fenêtres sur mesure de groupe Lapeyre/Saint-Gobain) ou la filiale brésilienne. Sur le Web, la segmentation des clients, le machine learning ou l'expérience utilisateur ont séduit KparK et Weber (préparation et finition des sols),» se félicite Eric Merlateau.

Un succès au-delà des attentes

La décision de lancer cette initiative a été prise en mars 2015 (lors du salon Big Data), et le premier POC a été initié en septembre 2015. *«Un processus itératif a permis d'affiner sans cesse notre plateforme et nos pratiques,»* raconte Stéphane Saretta, responsable Architecture & Engineering chez Saint-Gobain. *«Au final, les retours des toutes équipes métier sont unanimement positifs, et particulièrement sur le service de découverte du Big data, Discover. Un outil qui a donc logiquement été intégré au catalogue des formations internes de Saint-Gobain. Sur les 10 projets initiés, plusieurs n'ont pas abouti. Cependant, ils nous ont permis de comprendre différents points essentiels et d'apprendre en marchant. Au final, même ces derniers ont été bénéfiques, pour nous, et pour les entreprises concernées, et pour le Groupe.»*

Parmi les leçons tirées de ces expériences, les deux responsables ne se contentent pas de constatations positives. *« Même assistée, la définition des chargements de données reste une tâche complexe et lourde à mettre en place,»* explique Stéphane Saretta. *«Cependant, l'intégration s'est révélée une des clés du succès, car 75% des données utilisées utilisaient des données internes.»*

« Le choix de Kerberos est un frein au déploiement en production de ce type d'environnement,» relate Eric Merlateau. *«En revanche, la ségrégation des données a été très utile pour déployer très rapidement des Proofs of concept (Prototypes).»*

Un avenir tout tracé, ou presque.

L'entreprise planifie déjà l'intégration la mise en production d'environnements en clusters dans un cloud partagé, intégré au réseau SG (Saint-Gobain), l'organisation de fourniture de services de production, l'élargissement de la palette d'outils (bases NoSQL et outils de recherche, notamment), et la définition d'une politique de mise à jour pour assurer la cohérence avec Cloudera.

En un an, la fabrique Big Data initiée pour tester le concept a donc bien donné lieu à des projets concrets et à une nouvelle dynamique dans le système d'information.

Qui a dit que les grands groupes traditionnels avaient plus de mal à suivre et intégrer l'innovation?

A lire aussi :

[Linky, c'est aussi un projet de Big Data](#)

[Big Data : les entreprises ne sont pas au niveau](#)

Crédit photo : agsandrew / Shutterstock