

La supervision : brique essentielle d'un projet Big Data

L'explosion des capacités d'analyse de gros volumes de données a induit la mise en place d'infrastructures de plus en plus complexes pour pouvoir les traiter. On a ainsi tendance à se focaliser uniquement sur la mise à disposition de ces moyens en occultant l'importance de les surveiller efficacement. La majorité des projets big data comporte des dizaines de serveurs surpuissants, imbriqués dans une architecture complexe et comportant de très nombreuses dépendances.

Mais quel est vraiment l'intérêt de lancer de tels projets, souvent très consommateurs de temps et de ressources, si on ne garantit pas la bonne maîtrise de leurs fonctionnements en permettant de comprendre très rapidement ce qui bloque ? Le big data n'a, en effet, de la valeur que si on parle également de Fast Data (temps réel), et donc de **garantie de disponibilité** des moyens considérables mis en œuvre. Déployer, par exemple, une architecture Hadoop peut très vite devenir un cauchemar si tout n'a pas été bien prévu en termes de monitoring.

Cauchemar n'est pas un terme trop fort surtout quand on regarde les résultats de l'enquête menée par Capgemini en novembre 2014 et révélée en début d'année : seuls 27% des responsables IT interviewés ont qualifiés leur projet big data de réussite. Ce chiffre résume à lui seul une situation assez catastrophique.

Un système big data se doit de garantir un respect des Service level agreements (SLA) définis, et ne peut confirmer sa pertinence que si on est en **capacité d'investiguer très rapidement sur des problèmes sources** pour les réparer. Le temps réel est aujourd'hui la clé de voute de toute valeur générée dans l'analyse de données. La réduction du temps d'indisponibilité reste donc au centre des préoccupations de toute équipe responsable d'une infrastructure. C'est lorsque tout va mal que la supervision prend toute sa dimension et toute sa valeur, et qu'on ne peut donc pas sous-investir dans ce domaine.

Ne négligez pas une brique clé qui, si vous la sous-estimez, risque fort de mettre en péril tous vos investissements et, au final, desservir votre stratégie d'analyse en temps réel d'un flux de données très riches. La supervision doit impérativement être perçue comme un prérequis indispensable à toute initiative et surtout comme un investissement essentiel dans une bonne démarche de gouvernance garantissant le bon fonctionnement de votre organisation. La supervision doit être l'affaire de tous.

Anticiper la supervision de son infrastructure big data est ainsi vital pour la bonne continuité du service rendu, sans attendre d'être mis au pied du mur. Être capable de cloisonner ses différents environnements en gardant une vue unifiée du SI, fournir la bonne information au bon interlocuteur et alerter en temps réel le responsable du problème source reste une démarche incontournable à la réussite de ses projets. En profiter pour communiquer sereinement et factuellement sur la maîtrise du SI est également un gage de bonne compréhension et de relation apaisée entre direction IT et direction métier. Enfin, garantir le bon fonctionnement de son infrastructure reste le meilleur moyen d'en justifier le coût.