

REPONSES



REPONSE AUX QUESTIONS POSEES DANS LE CADRE DU MARCHE N° AOO N° 008/ARCEP/PRMP/2024 du 13 août 2024

Question 1 : Au vu des recommandations pour assurer la haute disponibilité et les enjeux critiques de ce projet, seriez-vous prêts à adopter cette nouvelle architecture ? Sinon avez-vous déjà pris en compte les problèmes mentionnés et souhaitez-vous maintenir la configuration actuelle ? Nous souhaitons clarifier votre position concernant cette évolution.

Réponse : *Nous avons pris en compte avec rigueur les recommandations concernant la haute disponibilité et comprenons les arguments avancés par Cloudera, notamment la mise en HA des composants critiques comme NIFI et Cloudera Manager. Cependant, après une analyse approfondie des besoins spécifiques de notre environnement et des ressources disponibles, nous avons décidé de maintenir l'architecture actuelle.*

En effet, notre configuration actuelle est conforme aux exigences de haute disponibilité, tout en intégrant les bonnes pratiques de Cloudera CDP. De plus, la réplication et la redondance de données, telles que spécifiées dans notre architecture, répondent aux normes requises sans nécessiter d'ajout de nœuds supplémentaires. Cette approche nous permet de rester conforme aux exigences de l'ARCEP (99,99 % d'uptime) sans surcharger notre infrastructure avec des licences et des composants supplémentaires.

Question 2 : Quel est la cible de l'architecture ?

Réponse : *L'architecture cible met l'accent sur la sécurisation et la haute disponibilité des nœuds Master, notamment les NameNodes, afin de garantir l'intégrité des métadonnées et la continuité des opérations en cas de défaillance. En assurant une redondance adéquate des Master et en appliquant les meilleures pratiques de Cloudera CDP Private, nous nous prémunissons contre toute perte de données critiques. Les nœuds Worker évolueront progressivement pour répondre aux futurs besoins en capacité, tout en maintenant l'efficacité actuelle. Par ailleurs, certains services clés, comme le ResourceManager et HDFS, seront redondants sans nécessiter l'ajout d'infrastructure supplémentaire, assurant ainsi une haute disponibilité optimale sans surcharge.*

Question 3 : La mise en place de la Haute Disponibilité du Data Lake mentionné dans le cahier de charge ne concerne que l'environnement de production. Qu'en est-il des environnements de développements et de tests ?

Réponse : *Concernant les environnements de développement et de tests, il est important de préciser que ceux-ci ne sont actuellement pas prévus dans notre infrastructure. La mise en place de la haute disponibilité telle que définie dans le cahier des charges s'applique exclusivement à l'environnement de production.*

Question 4 : Quel est la stratégie et la vision long terme de l'ARCEP dans la mise en haute disponibilité de ses données et ressources ?

Réponse : *La stratégie et la vision à long terme de l'ARCEP en matière de haute disponibilité de ses données et ressources reposent sur une approche progressive et soutenue, centrée sur la maîtrise graduelle de chaque étape du processus. Cette démarche vise à garantir une résilience optimale tout en minimisant les risques à chaque phase de l'implémentation. L'ARCEP privilégie une montée en puissance contrôlée, en renforçant d'abord les éléments critiques tels que les nœuds Master et les services essentiels, tout en prévoyant une évolution continue de l'infrastructure pour répondre aux besoins futurs.*

Question 5 : Pouvez-vous nous fournir la liste des instances de données et des composants concernés par la Haute Disponibilité ?

Réponse : *Les instances et composants critiques pour la haute disponibilité incluent HDFS (avec NameNode, JournalNodes), YARN (ResourceManager), Hive Metastore, ZooKeeper, HBase (Master et RegionServers), Impala (Catalog Server et StateStore), Solr, et Oozie. Toutes les bases de données associées doivent également être configurées en haute disponibilité. Il est important de noter que cette liste n'est pas exhaustive. Le prestataire est tenu de fournir des conseils et une expertise pour maîtriser ces composants critiques et garantir leur configuration en HA, conformément aux exigences de l'ARCEP.*