



Big Data - Spécifications Techniques

Responsable : **Andreani Quentin**

Veuillez noter que cette architecture technique est une suggestion basée sur les informations fournies dans le document de l'appel d'offres. Il sera nécessaire de l'adapter en fonction des exigences spécifiques du projet.

1. Collecte des données

1.1 IoT Sensors

Les capteurs IoT installés sur les véhicules collecteront une variété de données telles que les informations sur les conditions de conduite (par exemple, vitesse, freinage, direction), les performances du véhicule (par exemple, consommation de carburant, usure des pièces) et les informations environnementales (par exemple, conditions météorologiques, état de la route). Ces données seront transmises en temps réel à un serveur central pour le traitement et l'analyse.

1.2 IoT Gateway

Un passerelle IoT sera utilisée pour collecter les données des capteurs IoT. La passerelle agira comme un intermédiaire, recevant les données des capteurs, les prétraitant si nécessaire (par exemple, filtrage du bruit, mise en forme des données), puis les transmettant au système de traitement des données sur Google Cloud Platform.

2. Stockage des données

2.1 Cloud Data Storage

Les données reçues par l'IoT Gateway seront transmises à un service de stockage de données sur Google Cloud Platform. Le service, Google Cloud Storage, offrira un stockage durable et hautement disponible pour les données, en veillant à ce qu'elles soient protégées et prêtes à être consultées à tout moment.

2.2 Database

Les données seront ensuite chargées dans une base de données Big Data, comme Google BigQuery. BigQuery est une base de données entièrement gérée capable de traiter des pétaoctets de données. Elle permet de mener des analyses complexes en temps réel, ce qui signifie que vous pouvez extraire des informations utiles des données presque immédiatement après leur collecte.

3. Traitement des données

3.1 Data Processing Engine

Le moteur de traitement de données, Google Dataflow, sera utilisé pour traiter et analyser les données. Dataflow permet de traiter à la fois des flux de données en temps réel et des lots de données historiques, ce qui signifie qu'il peut transformer les données en informations utiles quasiment en temps réel. Ces données seront donc ensuite mis à la disposition du service en charge de l'intelligence artificiel via dataflow.

3.2 Machine Learning/AI

Google Cloud AI et des outils de Machine Learning seront utilisés pour analyser les données et faire des prédictions. Par exemple, ces outils pourraient prédire quand un véhicule pourrait avoir besoin d'un entretien en fonction de ses données de performance passées et actuelles. Ces prédictions peuvent aider à réduire les coûts d'entretien et à améliorer la sécurité. Pour plus d'information concernant la partie IA veuillez vous référer au document de spécification technique de l'équipe en charge de l'IA.

4. Visualisation des données

4.1 Dashboard

Les informations générées par le système de Machine Learning seront affichées sur un tableau de bord, probablement via un outil comme Google Looker Studio. Cela permettra aux utilisateurs de surveiller en temps réel les performances et les conditions de chaque véhicule. Les anomalies ou les prédictions importantes seront également mises en évidence, de sorte que des mesures appropriées peuvent être prises rapidement.

5. Sécurité des données

5.1 Data Encryption

Toutes les données, de leur collecte à leur traitement, seront sécurisées grâce à l'encryption de bout en bout offerte par Google Cloud Platform. Cela garantit que même si les données sont interceptées, elles ne pourront pas être lues par des acteurs malveillants.

5.2 Access Control

Google Cloud Platform fournit des outils de gestion des droits d'accès pour contrôler qui peut accéder à quelles données. Cela garantit que seules les personnes autorisées peuvent accéder, modifier ou supprimer les données.

5.3 Security Instructions

Des instructions de sécurité claires seront fournies pour assurer la sécurité de la solution. Ces instructions pourraient inclure des recommandations sur la gestion des mots de passe, la définition des autorisations d'accès, et d'autres meilleures pratiques de sécurité.

6. Architecture du Système de données

6.1 Schéma de l'Architecture

```
graph TD
    A(IoT Sensors) -->|Data Collection| B(IoT Gateway)
    B -->|Data Transmission| C(Cloud Data Storage)
    C -->|Data Storing| D(Database - Hadoop HDFS / Google BigQuery)
    D -->|Real-Time Analysis| E(Data Processing Engine - Apache Spark / Google Dataflow)
    E -->|Data Processing| F(Machine Learning/AI)
    F -->|Insights Generation| G(Dashboard)
    G -->|Unusual Data Alert| H(Alert System)
    H -->|Data Visualization| I[End Users]
    I -->|Alerts| J
    J -->|Data Encryption, Access Control| K(Data Security)
    K --> L[Security Instructions]
```