

---

---

## **DST Airlines**

Exploration du Trafic Aérien : Utilisation de l'API Lufthansa pour  
le Suivi des Vols

---

---

# Part I : Introduction

# Contexte, objectifs et équipe projet

## Contexte :

Dans le cadre de notre formation en Data Engineering, nous explorons les possibilités de suivi des vols en temps réel grâce aux données disponibles mondialement. Actuellement, il est possible de traquer un avion en temps réel en utilisant des API fournies par différentes compagnies aériennes. Notre projet se concentre spécifiquement sur l'utilisation des données fournies par l'API de Lufthansa.

## Objectifs :

- Utiliser le streaming en temps réel pour la surveillance continue des avions.
- Implémenter un traitement par lots pour ingérer, analyser et présenter les données de l'API Lufthansa.

## Equipe du projet

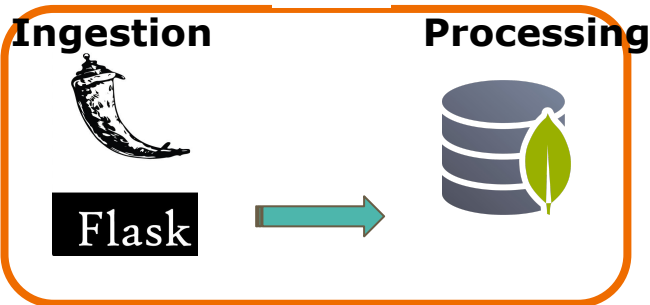
- **Soufiane A.**
- **Badr A.**
- **Nadia A.**

# Architecture générale

Data Source



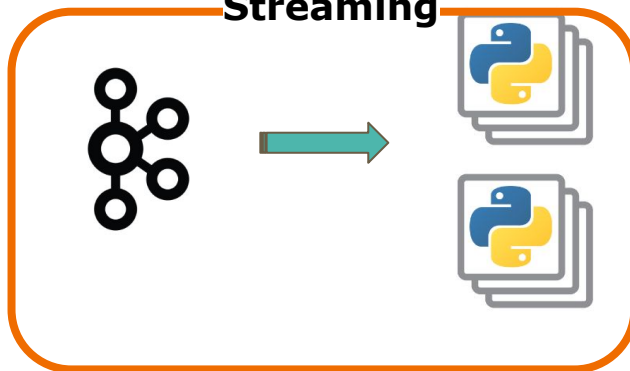
Batch



Consommation

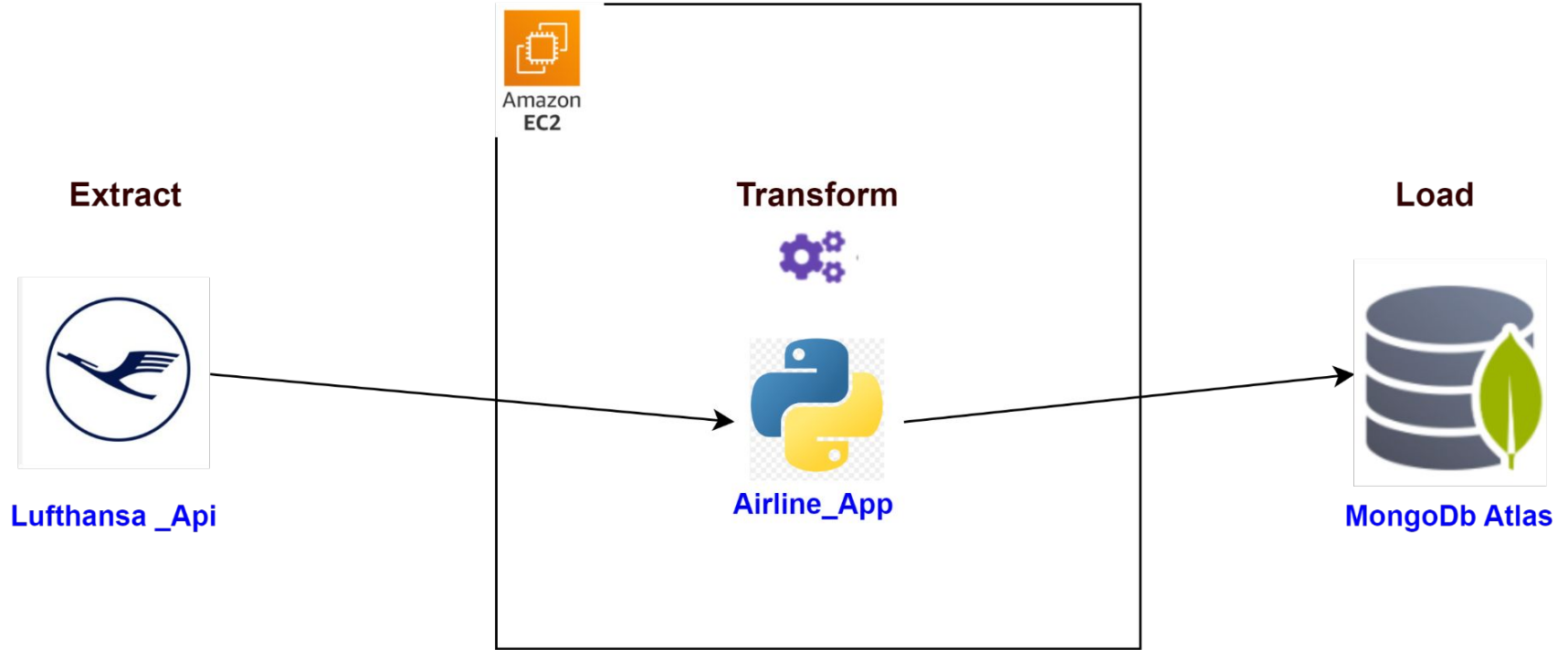


Streaming

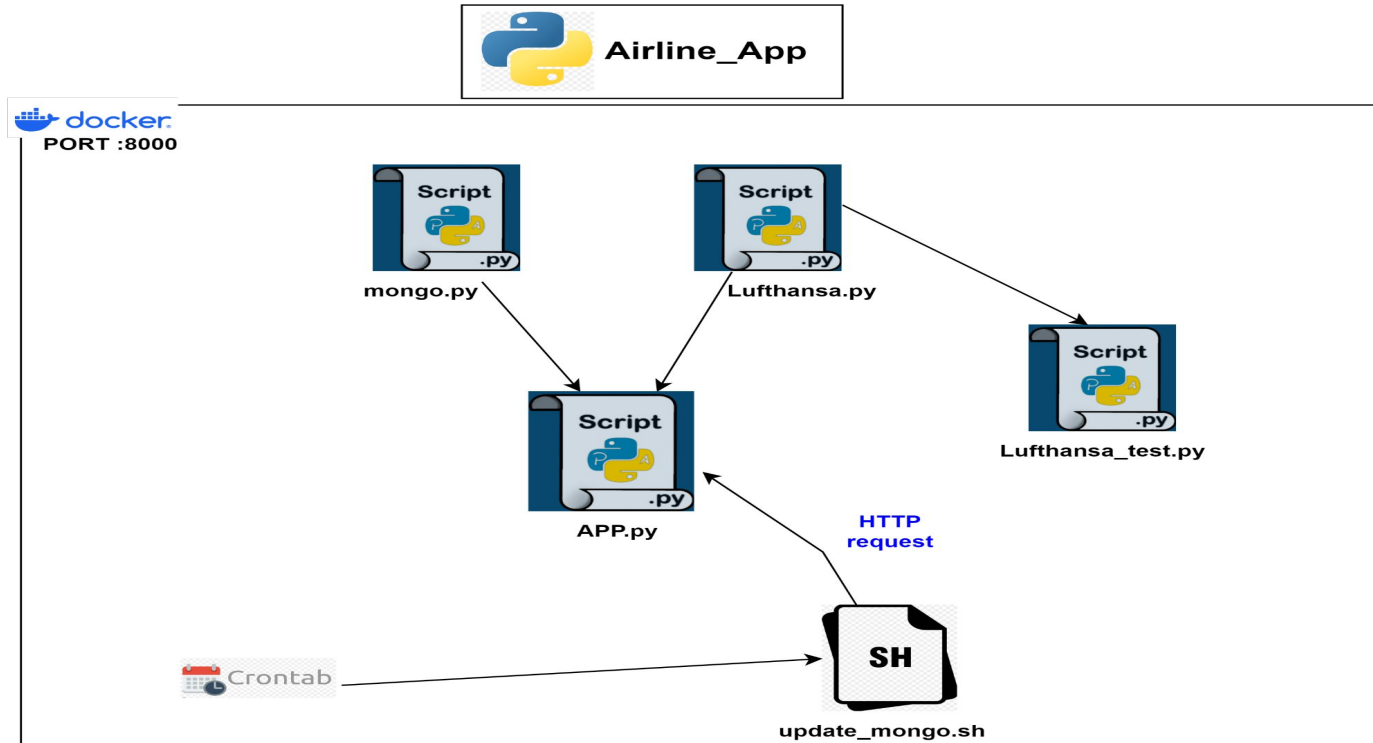


## Part II : Batch

# Processus ETL



# Batch : architecture globale



# Exemple de document MongoDB

## Flights:

```
_id: ObjectId('6675f800be51b73272174142')
▶ TotalJourney : Object
▼ Flight : Array (2)
  ▼ 0: Object
    ▼ Departure : Object
      AirportCode : "JFK"
      ▼ ScheduledTimeLocal : Object
        DateTime : "2024-06-22T15:45"
      ▶ Terminal : Object
    ▼ Arrival : Object
      AirportCode : "FRA"
      ▼ ScheduledTimeLocal : Object
        DateTime : "2024-06-23T05:25"
      ▶ Terminal : Object
    ▶ MarketingCarrier : Object
    ▶ Equipment : Object
    ▶ Details : Object
  ▶ 1: Object
insert_at : "2024-06-21 22:00"
```



# Dash : consommation des données batch

Airlines, des milliers de vols au bout de votre souris.

De

Francfort

×

▼

Départ

22/06/2024

Vers

Argyle

×

▼

Retour

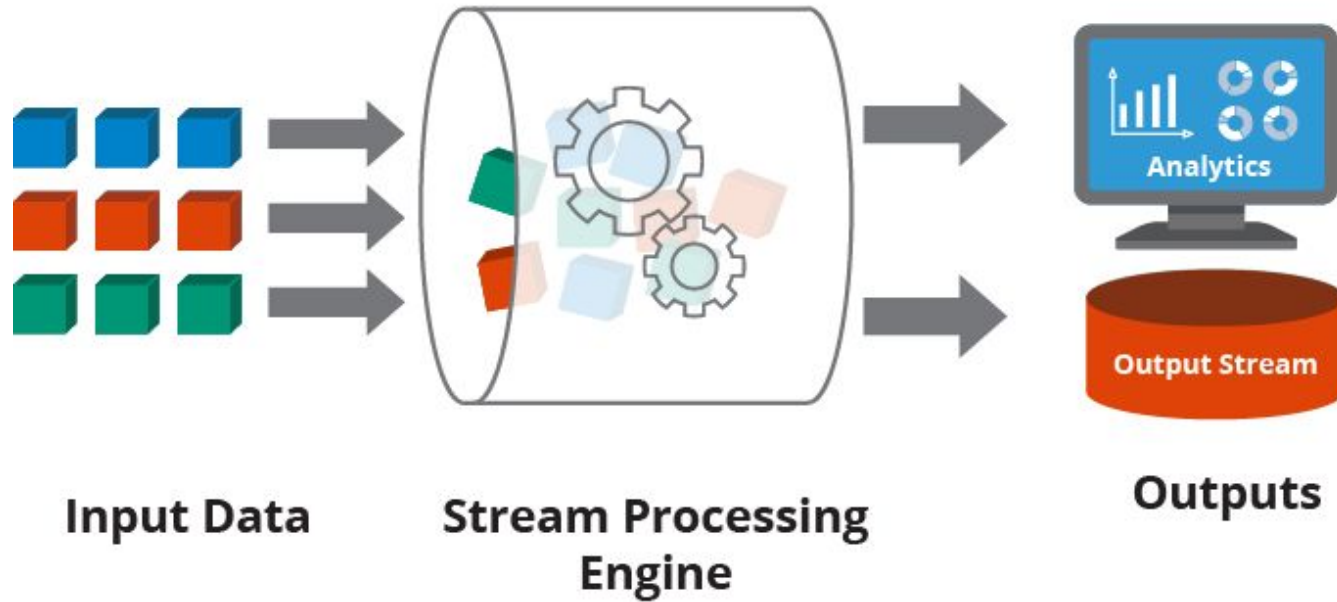
02/07/2024

Rechercher

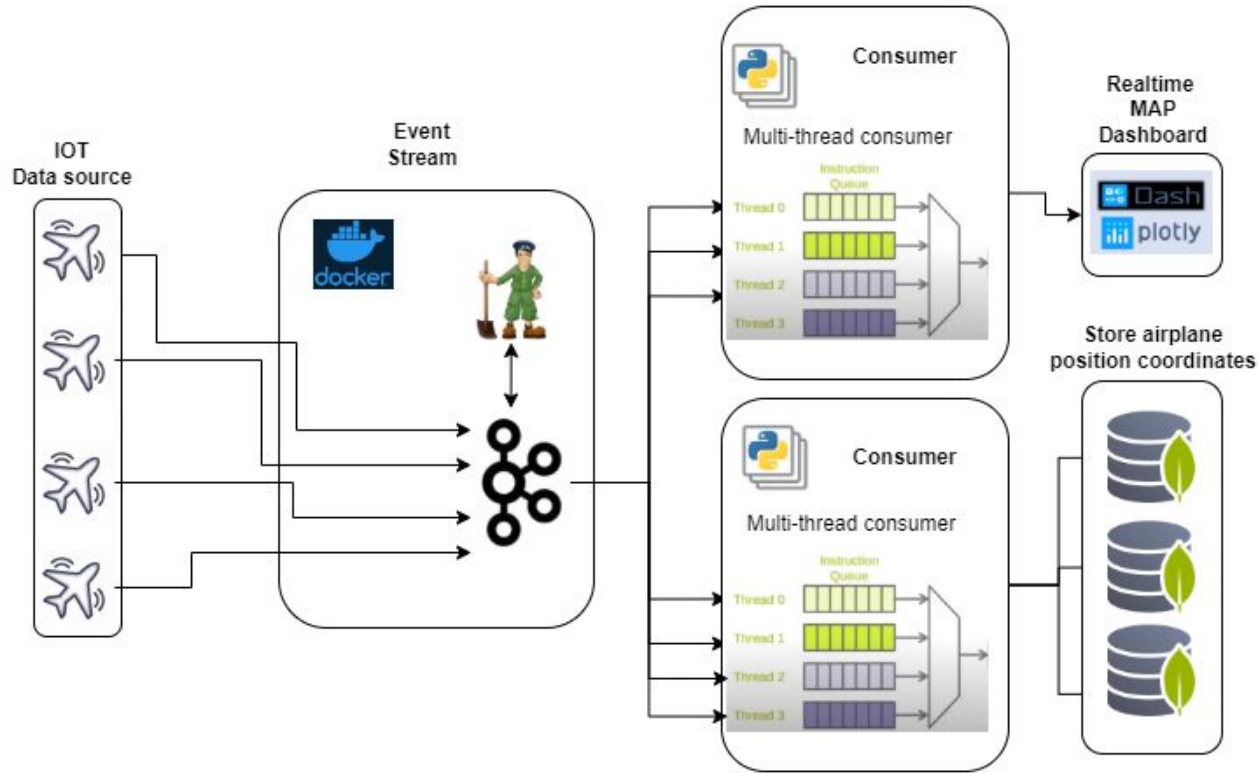
Durée	De	Vers	Date de départ	Heure de départ	Date d'arrivée	Heure d'arrivée
9 h 54 min	Francfort	Argyle	22/06/2024	06h51	22/06/2024	09h53
11 h 14 min	Francfort	Argyle	22/06/2024	06h51	22/06/2024	09h53

# Part III : Streaming de données

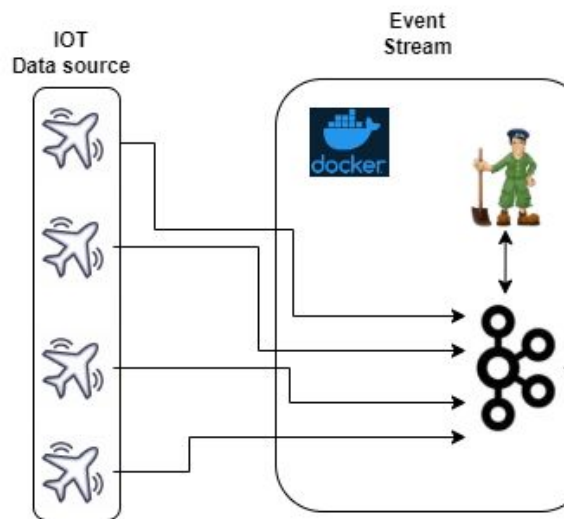
# Streaming de données



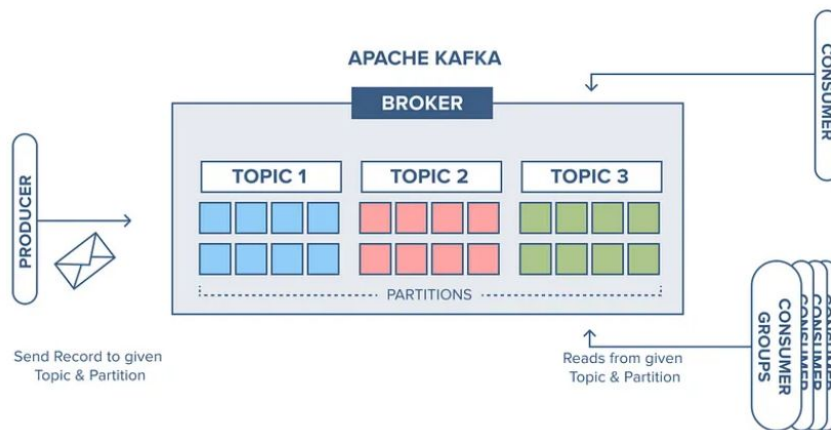
# Real-time streaming : architecture



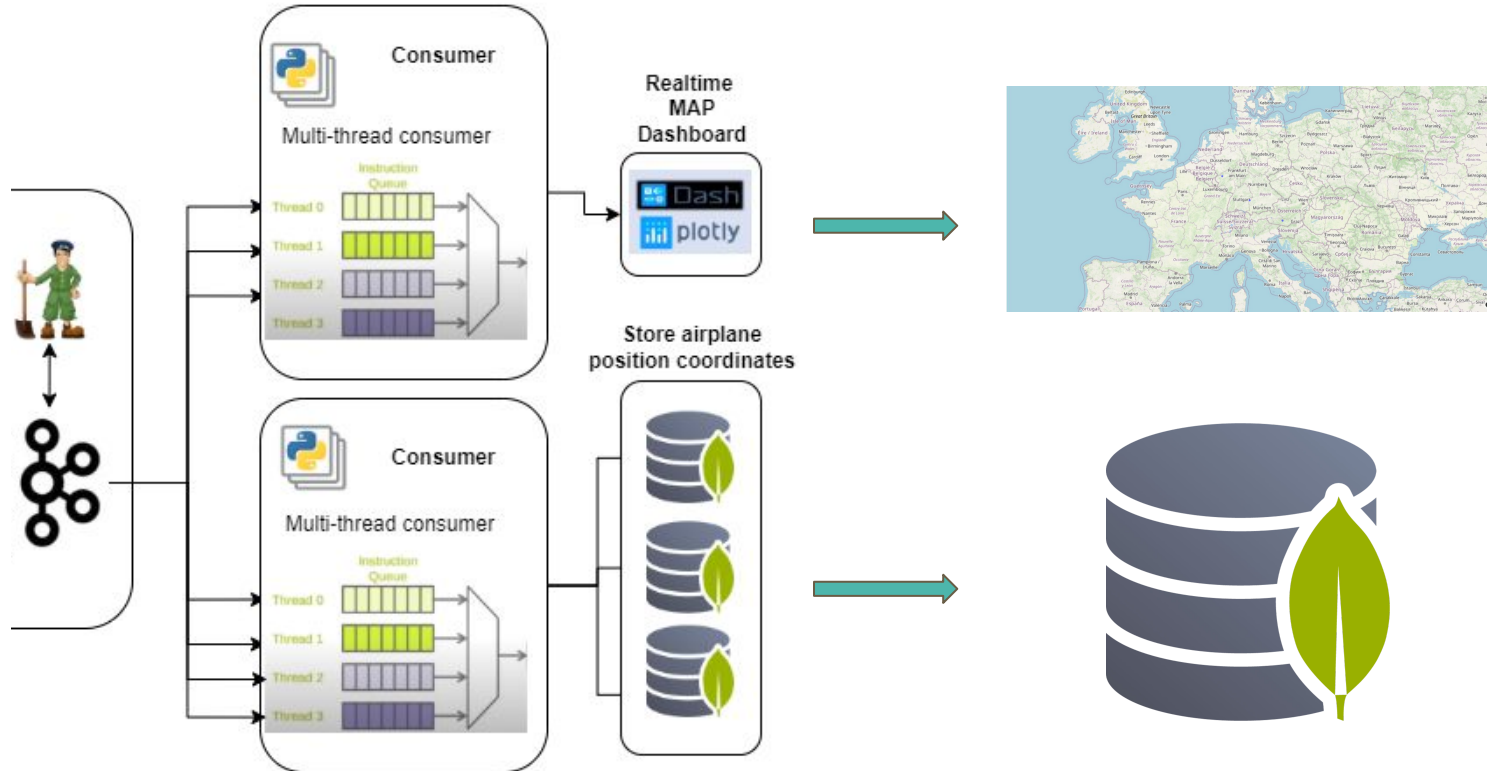
# Real-Time streaming : IOT et Event Stream



```
"flight_id": "LH102",  
"airline_name": "Lufthansa",  
"departure_country": "Germany",  
"departure_airport": "Frankfurt Airport",  
"departure_position": {"latitude": 50.0379, "longitude": 8.5622},  
"arrival_country": "UK",  
"arrival_airport": "Heathrow Airport",  
"arrival_position": {"latitude": 51.4700, "longitude": -0.4543}
```



# Processing et consommation de données



- ▶  admin
- ▶  config
- ▼  **iot\_flights**

 **LH102** ...

 LH453

 LH610

- ▶  local

 **ADD DATA** ▼

 **EXPORT DATA** ▼

 **UPDATE**

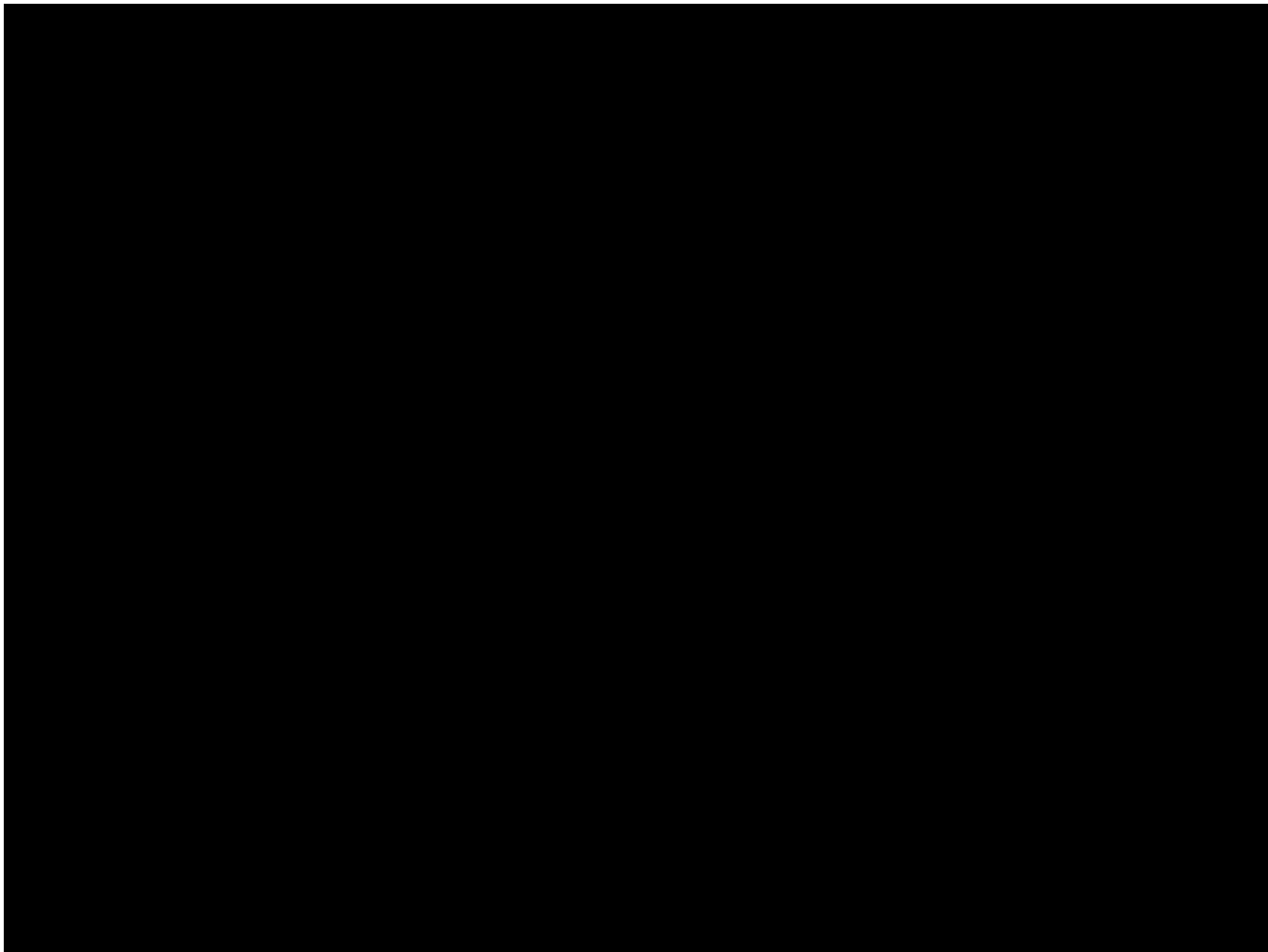
 **DELETE**

```
_id: ObjectId('667a91c81f25577566b4db0a')
flight_id: "LH102"
airline_name: "Lufthansa"
timestamp: "2024-06-25T11:26:42.676072"
departure_country: "Germany"
arrival_country: "UK"
departure_airport: "Frankfurt Airport"
arrival_airport: "Heathrow Airport"
latitude: 50.038348589099726
longitude: 8.562091778215667
speed: 704.3377539420189
altitude: 9278.041430660965
```

```
_id: ObjectId('667a91c81f25577566b4db11')
flight_id: "LH102"
airline_name: "Lufthansa"
timestamp: "2024-06-25T11:27:06.063221"
departure_country: "Germany"
arrival_country: "UK"
departure_airport: "Frankfurt Airport"
arrival_airport: "Heathrow Airport"
latitude: 50.05241578998163
longitude: 8.472502814475817
speed: 414.1742975775032
altitude: 9476.646879087
```

# Dash : consommation des données streaming





# Part V : Conclusion

# Axe d'amélioration

-**Analyse prédictive** : intégrer des modèles d'apprentissage automatique pour prédire les retards de vols ou les changements de trajectoire basés sur les données historiques et en temps réel.

-**Amélioration de l'interface utilisateur** : ajouter des fonctionnalités interactives supplémentaires, comme la possibilité de zoomer sur des régions spécifiques ou de filtrer les vols par compagnie aérienne.

-**Tests automatisés** : développer une suite de tests plus complète, incluant des tests d'intégration et de charge, pour garantir la fiabilité du système lors des mises à jour.

-**Conformité RGPD** : s'assurer que le traitement et le stockage des données sont conformes aux réglementations sur la protection des données, en particulier le RGPD pour les vols européens.

## Part VI : Q&A