

A BIG DATA ARCHITECTURE FOR PRODUCT REPUTATION MONITORING AND TRACKING.

SUPERVISEUR :

FAHD KALLOUBI



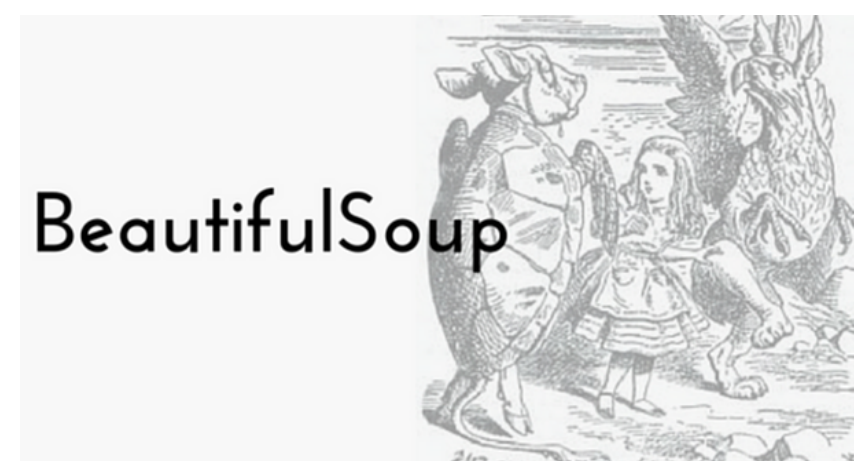
PROBLEMATIQUE



- Comment exploiter une architecture Big Data pour surveiller en temps réel la réputation des produits d'Amazon
- En quoi l'analyse des commentaires des vidéos de youtube relatifs à un produit peut-elle contribuer à une évaluation plus précise de la réputation de ce dernier
- Comment les Large Language Models (LLMs) peuvent apporter une précision dans l'analyse des commentaires relatifs aux produits?

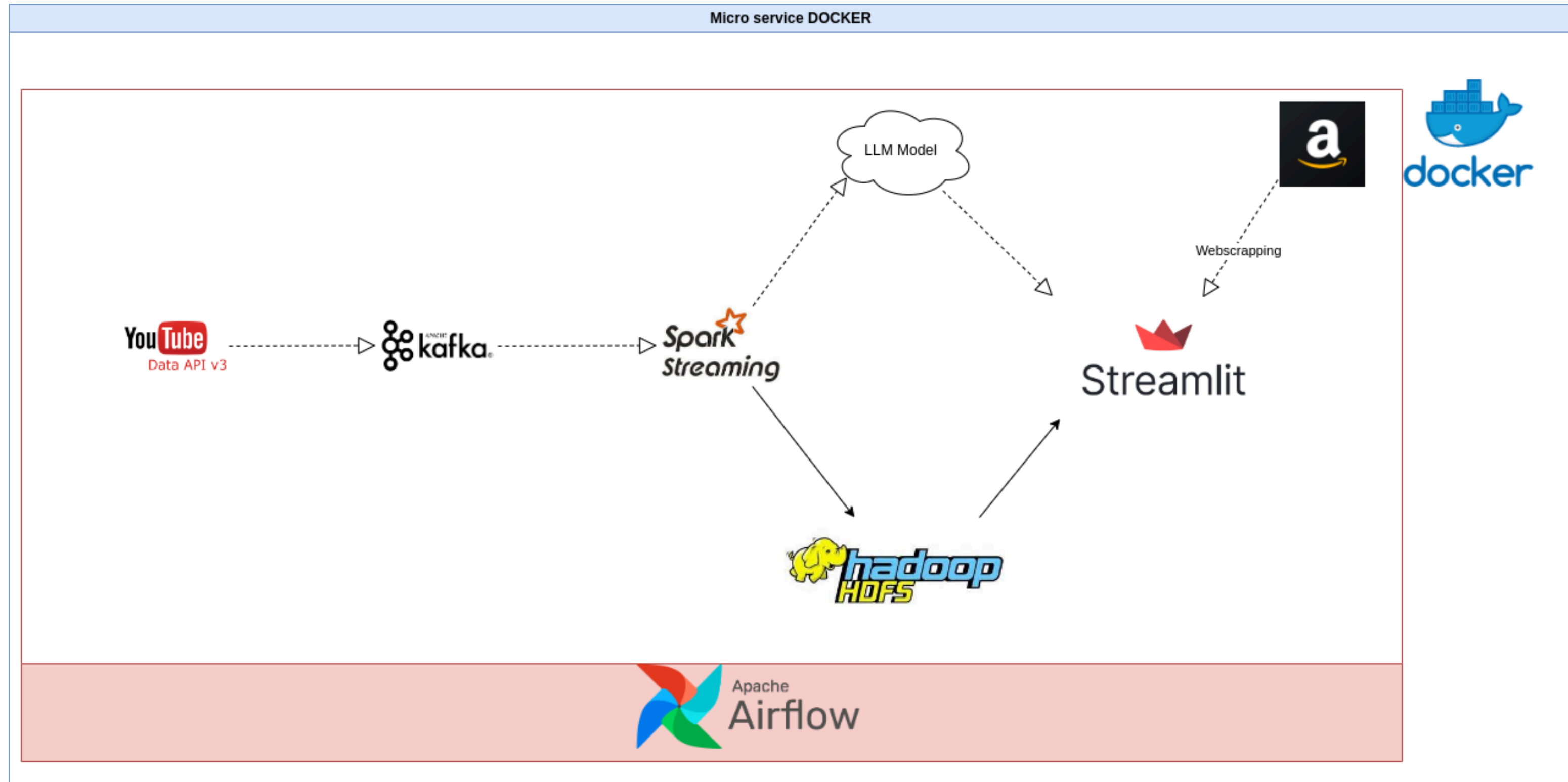


Technologies utilisées



ONNX

ARCHITECTURE



Modèles de Langages

	BERT	RoBERTa	DistilBERT
Size (millions)	Base: 110 Large: 340	Base: 110 Large: 340	Base: 66
Training Time	Base: 8 x V100 x 12 days* Large: 64 TPU Chips x 4 days (or 280 x V100 x 1 days*)	Large: 1024 x V100 x 1 day; 4-5 times more than BERT.	Base: 8 x V100 x 3.5 days; 4 times less than BERT.
Performance	Outperforms state-of-the-art in Oct 2018	2-20% improvement over BERT	3% degradation from BERT
Data	16 GB BERT data (Books Corpus + Wikipedia). 3.3 Billion words.	160 GB (16 GB BERT data + 144 GB additional)	16 GB BERT data. 3.3 Billion words.

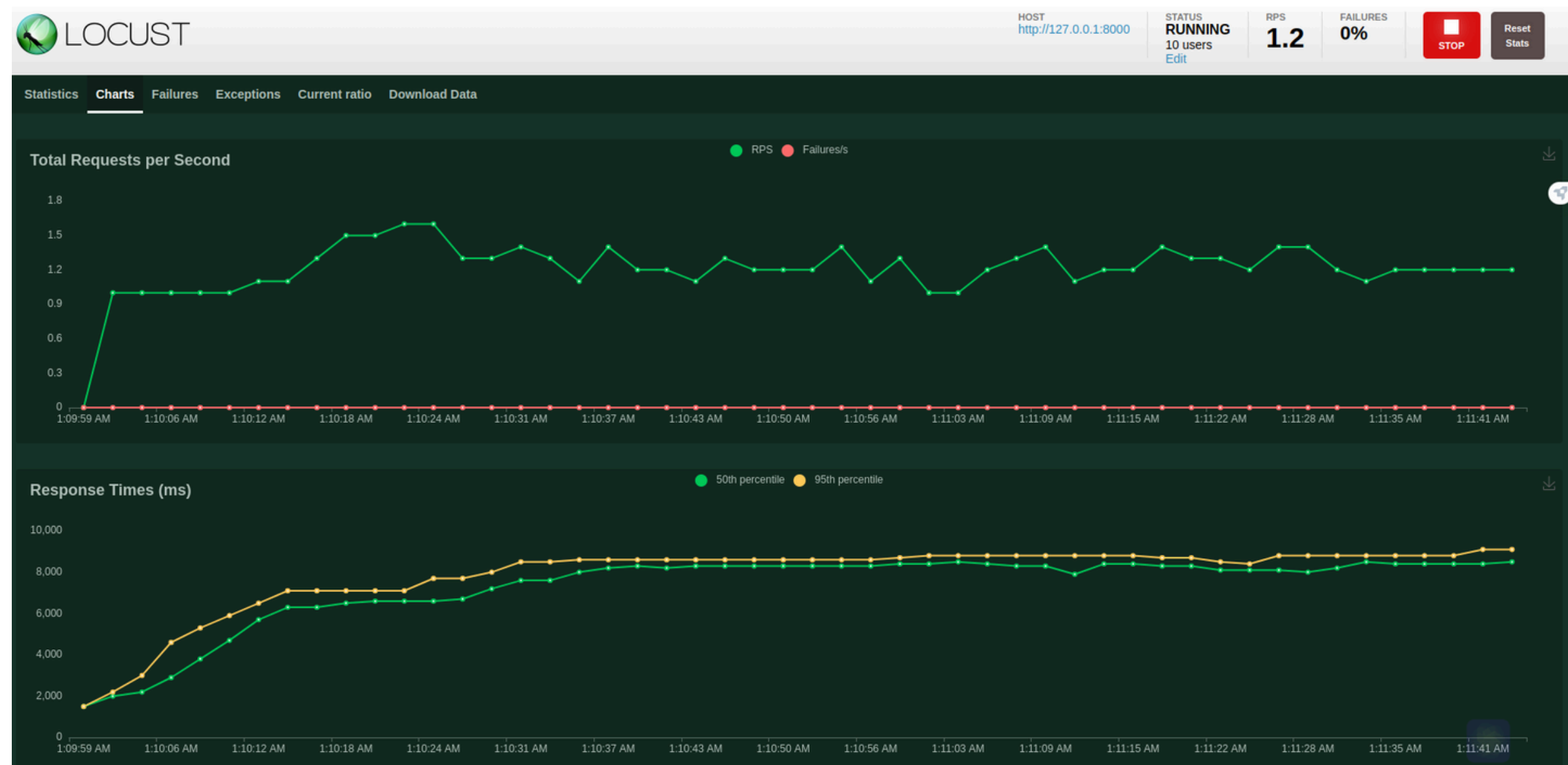
RESULTATS

Comparaison des modèles : Performance globale

Modèle	Accuracy (%)
BERT	98
RoBERTa	98
DistilBERT	97

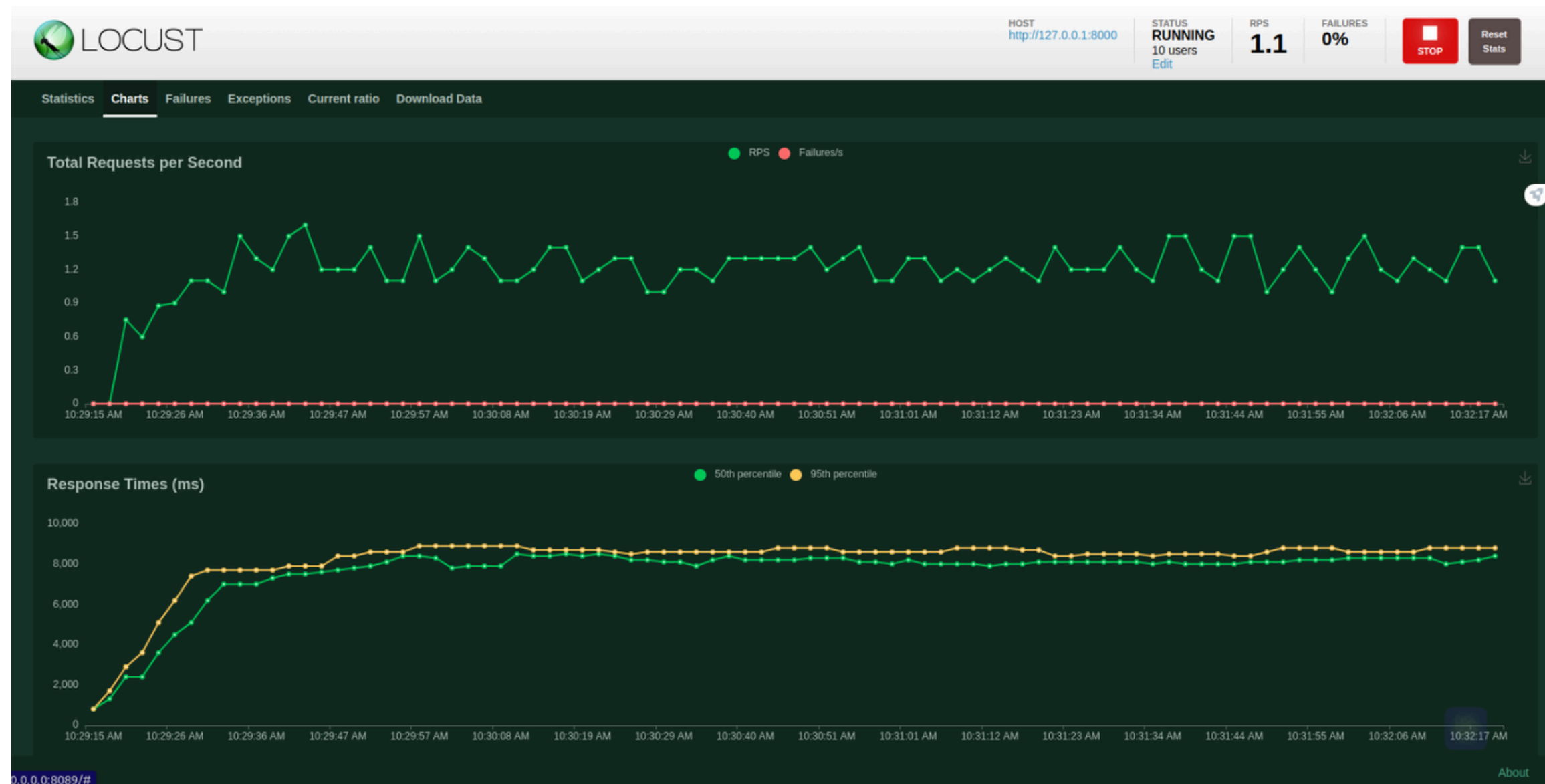
RESULTATS

Comparaison des modèles : Test de charge



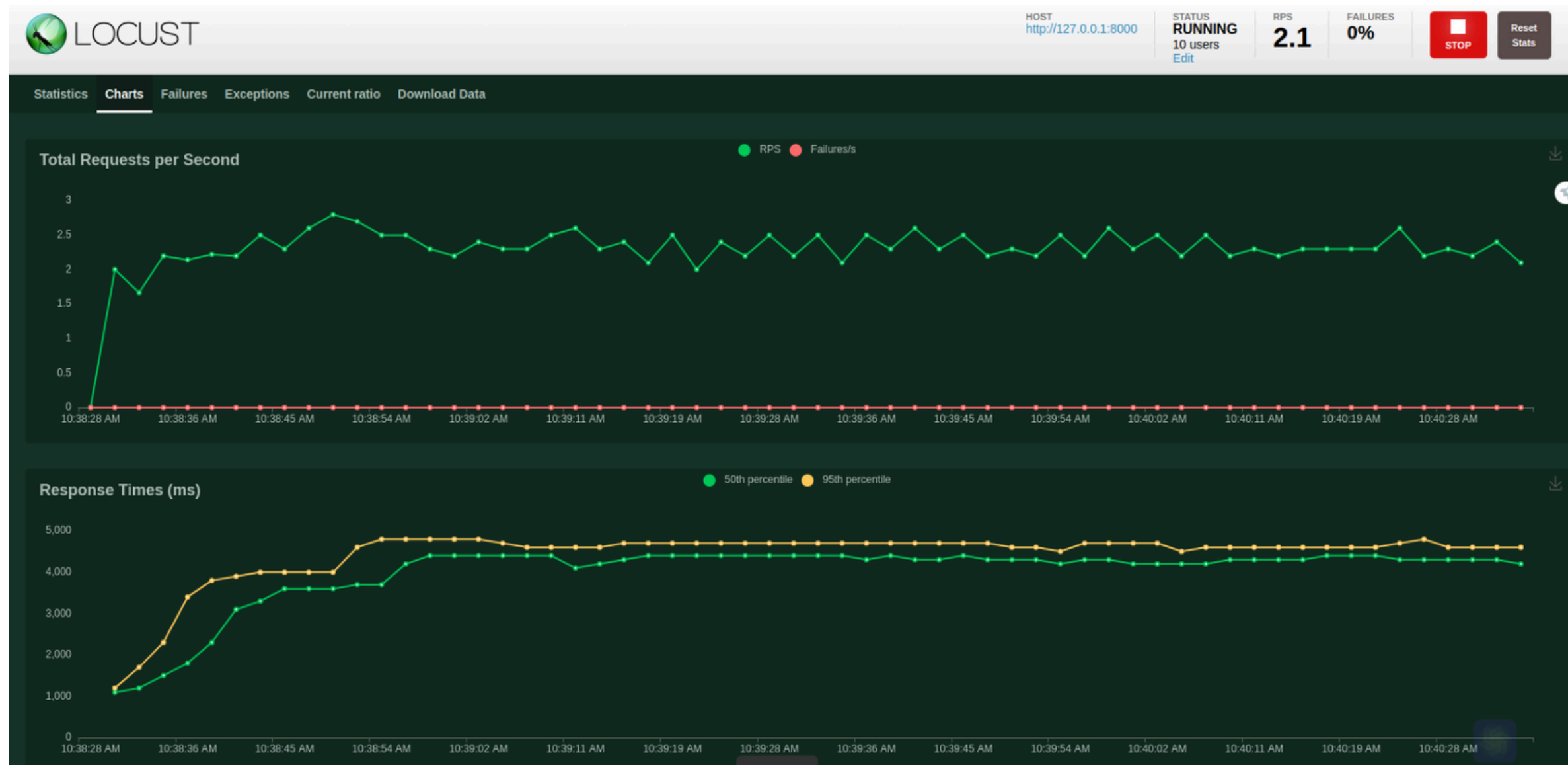
RESULTATS

Comparaison des modèles : Test de charge



RESULTATS

Comparaison des modèles : Test de charge



RESULTATS

Streamlit App

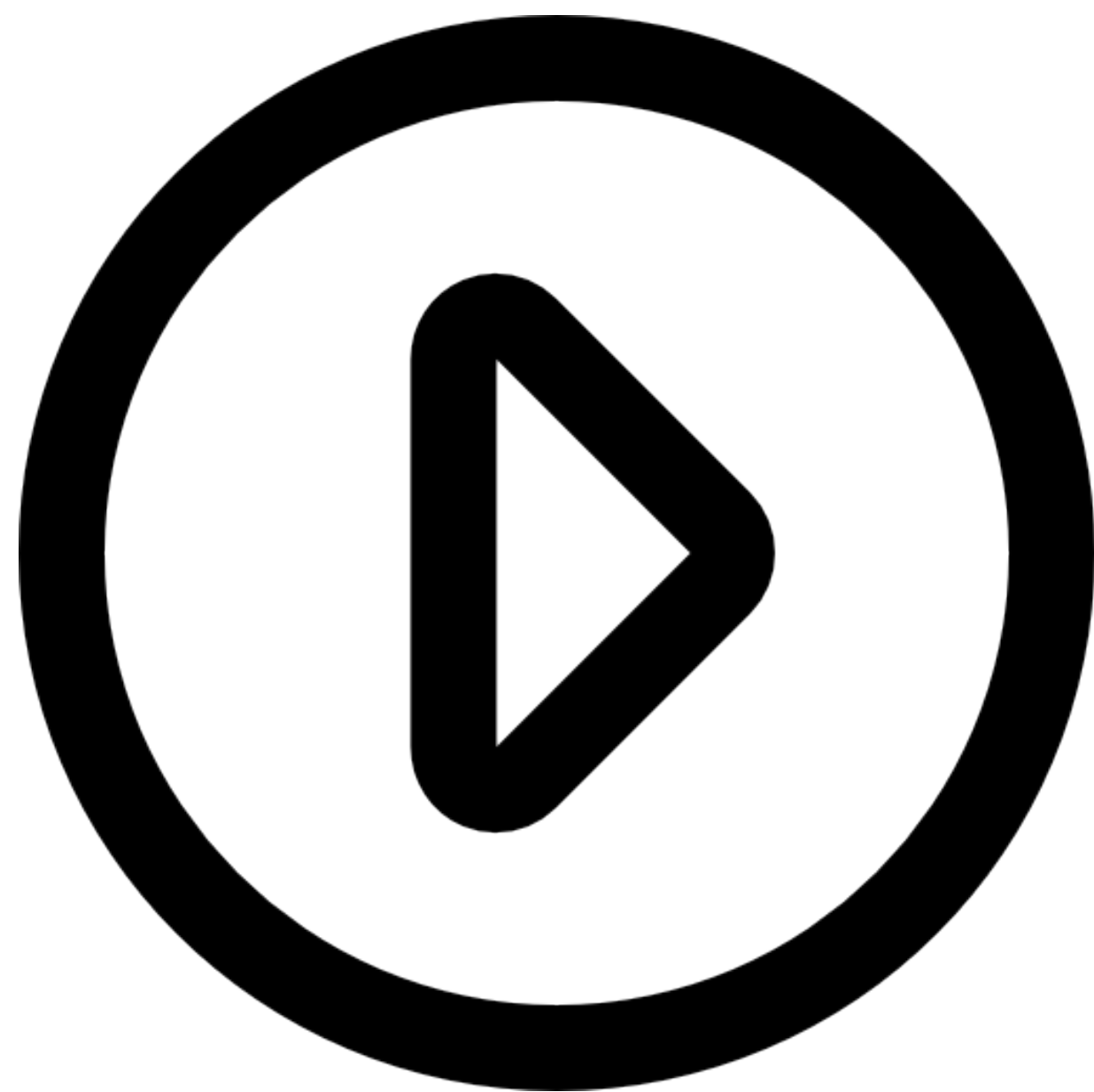


RESULTATS

Streamlit App



DEMO



L'équipe



Alioune Badara
Ndao



EWINSOU D.
Roméo



TINA Djara
Olivier



Merci !

N'hésitez pas à poser des questions si vous en avez.