

# Conception et développement de benchmarks pour évaluer l'impact des choix de modèles de stockage relationnels et orientés documents sur la performance de la recherche dans une collection de documents

## Contexte

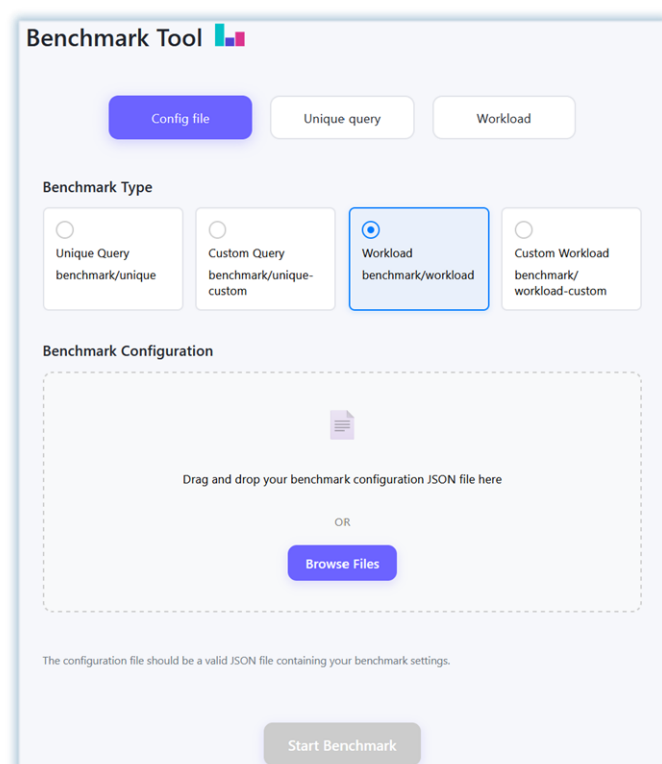
Avec l'essor des applications manipulant de grands volumes de données semi-structurées, les bases de données NoSQL orientées documents sont devenues une alternative populaire aux modèles relationnels traditionnels. Cependant, des SGBD relationnels matures comme PostgreSQL intègrent désormais des fonctionnalités natives pour gérer des données documentaires, notamment via le type JSONB, brouillant ainsi les frontières entre les deux paradigmes. Ce travail évalue objectivement ces approches dans un contexte de recherche documentaire.

## Objectifs

L'objectif principal était de concevoir et de développer une plateforme de benchmark pour comparer l'impact de trois modèles de stockage sur la performance de la recherche documentaire:

- PostgreSQL classique (mode relationnel)
- PostgreSQL avec le type JSONB (mode document).
- Couchbase (mode document natif).

La plateforme, développée avec Java, Javalin, Vue.js et déployée via Docker, permet de configurer des workloads (type/volume de requêtes, indexation, etc.), de les exécuter et de collecter des métriques détaillées.

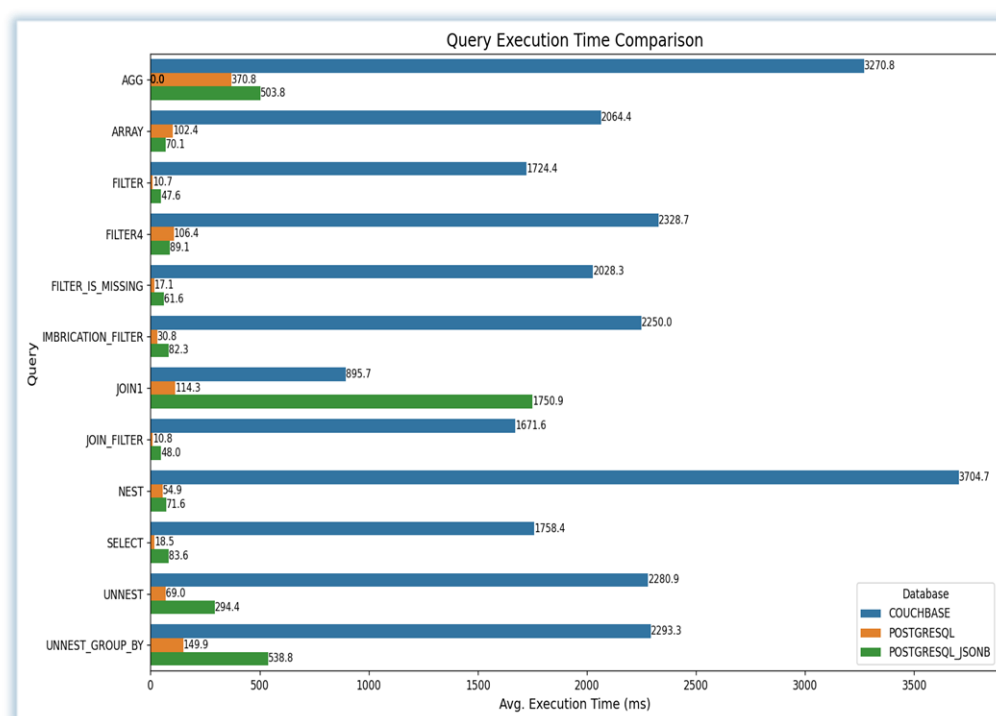


Auteur: Eva Ray  
Prof. responsable: Nastaran Fatemi  
Sujet proposé par: Eva Ray

## Résultats

L'analyse des benchmarks a révélé une hiérarchie de performance nette et a démontré que les optimisations fines sont un facteur de performance critique.

- **Hiérarchie claire** : PostgreSQL relationnel a les meilleures performances, suivi de l'approche JSONB, qui surpasse elle-même largement Couchbase sur la majorité des requêtes testées.
- **Compromis JSONB** : JSONB se révèle être une alternative très compétitive, combinant flexibilité et performance, surpassant parfois le relationnel (ex. : recherches dans des tableaux) et presque toujours Couchbase.
- **Rôle clé des optimisations** : Les performances s'améliorent considérablement grâce à une syntaxe optimisée, à l'ajout d'index adaptés et à la génération de statistiques pour les champs JSONB.



## Conclusion

PostgreSQL (Relationnel) reste une solution extrêmement robuste et performante, garantissant l'intégrité des données, au prix d'une modélisation initiale plus stricte. PostgreSQL (JSONB) est une alternative performante et sérieuse à Couchbase mais impose des requêtes plus verbeuses et nécessite l'ajout manuel de statistiques, sans lesquelles l'optimiseur peut ne pas fournir le plan d'exécution le plus performant. Couchbase reste en retrait, mais son manque de maturité face à PostgreSQL peut expliquer cet écart.