

الجزائرية الديمقراطية الشعبية الجمهورية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

المدرسة العليا للإعلام الآلي - 08 ماي 1945 - بسبدي بلعباس
Ecole Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès



MÉMOIRE

En Vue de l'obtention du diplôme de **Master**

Filière : **Informatique**

Spécialité : **Ingénierie des Systèmes Informatiques (ISI)**

Thème

Visibility Graph-based Prefetching for
Time Series Data Workload

Présenté par :

- Mr BENSALÉM Akram

Soutenu le : **26/09/2023**

Devant le jury composé de :

- | | |
|---------------------------|--------------|
| - M KHALDI Belkacem | Président |
| - M BENSENANE Hamdan | Examinateur |
| - M BELFEDHAL Alaa Eddine | Encadreur |
| - M LALLET Julien | Co-Encadreur |

Année Universitaire : 2022 / 2023

Abstract

The emergence of edge cloud systems has revolutionized the accessibility and convenience of applications and services across multiple sectors. One particular challenge presented by these systems is the effective management of data workloads, especially when it comes to connected cars.

Effective management of time series data workloads is essential to ensure optimal performance of edge cloud systems. Cache management is critical in achieving this objective by reducing retrieval time. Visibility graphs are a modern innovation in this domain that facilitates real-time workload management. Visibility graphs have demonstrated remarkable efficacy across multiple applications. When we integrate them into cache management, it helps the system run smoothly and reduces latency.

Recent studies on real-time data processing have shown significant advancements in cache management and forecasting data-based visibility graphs. The combination of these two approaches is a promising solution that opens the door for more enhancement in the future.

Keywords: Big Data, Caching, Visibility Graph, Edge Cloud, Time Series

Résumé

L'émergence des systèmes de cloud en bordure (edge cloud systems) a révolutionné l'accessibilité et la commodité des applications et services à travers de nombreux secteurs. Un défi particulier posé par ces systèmes est la gestion efficace des charges de travail des données (data workloads), en particulier en ce qui concerne les voitures connectées (connected cars).

La gestion efficace des charges de travail des données en séries temporelles (time series data workloads) est essentielle pour garantir une performance optimale des systèmes de cloud en bordure (edge cloud systems). La gestion du cache (cache management) est cruciale pour atteindre cet objectif en réduisant le temps de récupération. Les graphes de visibilité (visibility graphs) sont une innovation moderne dans ce domaine qui facilite la gestion de la charge de travail en temps réel. Les graphes de visibilité ont démontré une efficacité remarquable à travers plusieurs applications. Lorsque nous les intégrons à la gestion du cache, cela aide le système à fonctionner sans heurts et réduit la latence.

Des études récentes sur le traitement des données en temps réel ont montré des avancées significatives dans la gestion du cache et les graphes de visibilité basés sur la prévision des données. La combinaison de ces deux approches est une solution prometteuse qui ouvre la porte à davantage d'améliorations à l'avenir.

Mots-clés: Big Data, Caching, Graphe de Visibilité, Cloud de Bord, Série Temporelle

الملخص

بدأ استخدام أنظمة الحوسبة السحابية في تعزيز الخدمات الذكية وفي مجالات عدة. وقد واجهت منصة السيارات الذكية تحدياً كبيراً في هذا الجانب. خصوصاً أن إدارة بيانات المتتالية زمنياً تحتاج إلى طريقة فعالة لإدارة ذاكرة التخزين. يمكن أن يساعد تخزين البيانات في الحوسبة السحابية على تحسين أداء النظام.

تعتبر إدارة البيانات المتسلسلة زمنياً بفعالية أمر أساسي لضمان الأداء الأمثل لأنظمة السحابة. وقد شهدت إدارة تخزين الذاكرة المؤقتة نتائج جيدة لتحقيق هذا الهدف وذلك من خلال تقليل الوقت الضائع في نقل البيانات. والشبكة للرؤية هي ابتكار حديث في هذا المجال قد يسهل إدارة البيانات الزمنية المتتالية. أظهرت شبكة الرؤية فعالية ملحوظة في عدة استعمالات عند دمجها مع الذاكرة التخزين المؤقتة مما يساهم في المحافظة على سلامة النظام.

أظهرت الدراسات الحديثة حول إدارة الذاكرة المؤقتة الخاصة في معالجة البيانات المتتاليات الزمنية كما الدراسات حول التنبؤ بالبيانات من خلال استعمال تقنية الشبكة الرؤية تقدماً ملحوظاً. ويتبين أن الجمع بين هذين المنهجين قد يكون حلاً واعداً يفتح الباب لمزيد من التحسينات في المستقبل.

كلمات مفتاحية : البيانات الضخمة، التخزين، شبكة الرؤية، الحوسبة السحابية، السلسلة الزمنية

List of Abbreviations

EC *Edge Computing*

DBMs *Database management systems*

ITS *Intelligent Transportation Systems*

TSDB *time series database*

QoS *Quality of service*

ACF *Autocorrelation Function*

NSM *Node Similarity Matrix*

MVA *Maximum Visibility Approach*

FPGA *Field-Programmable Gate Arrays*