الجمهورية الشعبية الديمقراطية الجزائرية
People's Democratic Republic of Algeria
وزارة التعليم العالي و البحث العلمي
Ministry of Higher Education and Scientific Research
المدرسة العليا للإعلام الآلي 8 ماي 1945 - سيدي بلعباس
Higher School of Computer Science
8 Mai 1945 - Sidi Bel Abbes

Graduation Thesis

To obtain the diploma of **Engineering Degree**

Field of Study: **Computer Science**

Specialization: **Computer Systems Engineering**

Theme

Design, Implementation, and AI-Driven Optimization of an E-Commerce Platform using Microservices Architecture and DevOps Principles

---

Presented by
Harche Samir

Defended on: **15 September, 2025**
In front of the jury composed of

Dr. MALKI Abdelhamid	President of the Jury
Dr. Abdelhak-Djamel SERIAI	Thesis Supervisor
Dr. Bensenane Hamdan	Co-Supervisor
Dr. BEZZAOUCHA Fatima Souad	Examiner

Academic Year: 2024/2025

Abstract

The widespread adoption of microservices architecture, driven by the need for scalability and agility, has introduced significant operational complexities in managing dynamic, distributed systems. Traditional load balancing mechanisms, often reliant on static rules, are ill-equipped to handle the fluctuating workloads and real-time performance variations inherent in these environments, leading to suboptimal resource utilization, increased latency, and a degraded user experience.

This thesis addresses these challenges through the design, implementation, and evaluation of an intelligent, self-optimizing e-commerce platform. The work integrates a full-stack microservices architecture with modern DevOps principles, including Infrastructure as Code (IaC) and Continuous Integration/Continuous Delivery (CI/CD). The system is containerized with Docker, orchestrated on Kubernetes, and leverages an Istio service mesh to manage inter-service communication, security, and observability.

The primary contribution is the development of an AI-driven optimization controller that utilizes Reinforcement Learning, specifically Q-learning, to intelligently manage traffic. This controller interfaces directly with the service mesh, dynamically adjusting load balancing policies based on real-time performance metrics (P99 latency, RPS, CPU usage, error rates) collected from a Prometheus monitoring stack.

Through empirical evaluation comparing the AI-controlled system against a baseline scenario with traditional load balancing, the proposed approach demonstrates significant improvements across key performance indicators. The findings show a marked reduction in P99 latency, an increase in request throughput, more balanced CPU utilization across service instances, and a substantial decrease in server-side error rates. This work serves as a comprehensive reference architecture and validates the efficacy of integrating artificial intelligence into the service mesh layer for creating resilient, efficient, and autonomous cloud-native applications.

Keywords: Microservices Architecture, Reinforcement Learning, DevOps, Service Mesh, Istio, Dynamic Load Balancing, Performance Optimization, Kubernetes, E-Commerce Platform, CI/CD.

Résumé

L'adoption généralisée de l'architecture microservices, motivée par le besoin de scalabilité et d'agilité, a introduit des complexités opérationnelles significatives dans la gestion des systèmes distribués et dynamiques. Les mécanismes traditionnels d'équilibrage de charge, souvent basés sur des règles statiques, sont mal adaptés pour gérer les charges de travail fluctuantes et les variations de performance en temps réel inhérentes à ces environnements, ce qui entraîne une utilisation sous-optimale des ressources, une latence accrue et une expérience utilisateur dégradée.

Cette thèse aborde ces défis à travers la conception, la mise en œuvre et l'évaluation d'une plateforme de commerce électronique intelligente et auto-optimisée. Le travail intègre une architecture microservices full-stack avec les principes DevOps modernes, y compris l'Infrastructure en tant que Code (IaC) et l'Intégration Continue/Livraison Continue (CI/CD). Le système est conteneurisé avec Docker, orchestré sur Kubernetes et s'appuie sur un maillage de services (service mesh) Istio pour gérer la communication inter-services, la sécurité et l'observabilité.

La contribution principale est le développement d'un contrôleur d'optimisation piloté par l'IA qui utilise l'Apprentissage par Renforcement, spécifiquement le Q-learning, pour gérer intelligemment le trafic. Ce contrôleur s'interface directement avec le maillage de services, ajustant dynamiquement les politiques d'équilibrage de charge en fonction des métriques de performance en temps réel (latence P99, RPS, utilisation du CPU, taux d'erreur) collectées à partir d'une pile de surveillance Prometheus.

Grâce à une évaluation empirique comparant le système contrôlé par l'IA à un scénario de base avec un équilibrage de charge traditionnel, l'approche proposée démontre des améliorations significatives sur les indicateurs de performance clés. Les résultats montrent une réduction marquée de la latence P99, une augmentation du débit de requêtes, une utilisation plus équilibrée du CPU entre les instances de service et une diminution substantielle des taux d'erreur côté serveur. Ce travail sert d'architecture de référence complète et valide l'efficacité de l'intégration de l'intelligence artificielle dans la couche du maillage de services pour créer des applications natives du cloud résilientes, efficaces et autonomes.

Mots-clés: Architecture Microservices, Apprentissage par Renforcement, DevOps, Maillage de Services, Istio, Équilibrage de Charge Dynamique, Optimisation des Performances, Kubernetes, Plateforme E-Commerce, CI/CD.

ملخص

أدى الاعتماد الواسع لمعمارية الخدمات المصغرة (Microservices)، مدفوعًا بالحاجة إلى قابلية التوسع والمرونة، إلى بروز تعقيدات تشغيلية كبيرة في إدارة الأنظمة الموزعة والديناميكية. إن آليات موازنة الحمل التقليدية، التي تعتمد غالبًا على قواعد ثابتة، غير مهيأة للتعامل مع أعباء العمل المتقلبة وتغيرات الأداء في الوقت الفعلي المتأصلة في هذه البيئات، مما يؤدي إلى استغلال غير أمثل للموارد، وزيادة زمن الاستجابة، وتدهور تجربة المستخدم. تتناول هذه الأطروحة هذه التحديات من خلال تصميم وتنفيذ وتقييم منصة تجارة إلكترونية ذكية وذاتية التحسين. يدمج هذا العمل معمارية خدمات مصغرة متكاملة مع مبادئ DevOps الحديثة، بما في ذلك البنية التحتية كشيفرة برمجية (IaC) والتكامل والتسليم المستمر (CI/CD). تم تحويط النظام باستخدام Docker، وتنسيقه على Kubernetes، كما يعتمد على شبكة الخدمات (Istio Service Mesh) لإدارة الاتصال بين الخدمات والأمن والمراقبة. تتمثل المساهمة الرئيسية في تطوير وحدة تحكم للتحسين مدفوعة بالذكاء الاصطناعي تستخدم التعلم المعزز (Reinforcement Learning)، وتحديدًا خوارزمية Q-learning، لإدارة حركة مرور البيانات بذكاء. تتفاعل وحدة التحكم هذه مباشرة مع شبكة الخدمات، وتضبط سياسات موازنة الحمل ديناميكيًا بناءً على مقاييس الأداء في الوقت الفعلي (زمن الاستجابة للمئين 99، عدد الطلبات في الثانية، استخدام وحدة المعالجة المركزية، معدلات الخطأ) التي يتم جمعها من حزمة المراقبة Prometheus. من خلال تقييم تجريبي يقارن النظام الذي يتم التحكم فيه بواسطة الذكاء الاصطناعي بسيناريو أساسي يعتمد على موازنة الحمل التقليدية، أظهر النهج المقترح تحسينات كبيرة عبر مؤشرات الأداء الرئيسية. وتشير النتائج إلى انخفاض ملحوظ في زمن الاستجابة للمئين 99، وزيادة في إنتاجية الطلبات، واستخدام أكثر توازنًا لوحدة المعالجة المركزية عبر نُسَخ الخدمات، وانخفاض كبير في معدلات الخطأ من جانب الخادم. يمثل هذا العمل معمارية مرجعية شاملة ويثبت فعالية دمج الذكاء الاصطناعي في طبقة شبكة الخدمات لإنشاء تطبيقات سحابية أصلية (Cloud-Native) مرنة وفعالة ومستقلة.

الكلمات المفتاحية: معمارية الخدمات المصغرة، التعلم المعزز، DevOps، شبكة الخدمات، Istio، موازنة الحمل الديناميكية، تحسين الأداء، Kubernetes، منصة التجارة الإلكترونية، CI/CD.