

**Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Computer Science**



Amelioration Real-time Containerization Big data Streaming for Online E-learning System in Cloud Environment

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Doctor of Philosophy in
Computer Science**

**By
Wael Wadullah Mahmood Hadeed**

**Supervised by
Prof. Dr. Dhuha Basheer Abdullah**

2023 A.D.

1444 A.H.

ABSTRACT

A new trend of advanced applications with high demands has emerged in recent years. The integration of cloud computing, big data and e-learning has become a critical need in the education sector. This integration makes it more feasible when it comes to providing infrastructure in terms of location, cost and maintenance. Moreover, cloud-computing-based e-learning approaches have the ability to provide better response time, high computing power, heterogeneous operating systems usage, reliable data security, document transmission, and facilitating collaborative teamwork. Big data is considered a remarkable aspect of developing communication and information systems. From the other side, edge computing is a new type of computing that brings cloud services closer to customers. In addition to that, edge computing reduces client/server latency significantly.

In this thesis, two e-learning platforms, Django Website Platform (DWP) and Flask Website Platform (FWP), were designed using different frameworks, Django and Flask, respectively. Furthermore, each platform is linked to the PostgreSQL database and the Apache-Kafka big data framework using the Debezium connector. In addition to deploying the platforms on the Heroku cloud service that supports website applications and containerization. In e-learning, the choice of platform-building tools plays an important role in having better response times. Therefore, response time is considered one of the important factors used for evaluating e-learning platforms' work, which is proven in this thesis.

A novel uncentralized Urgent Container Migration (UCM) algorithm is proposed for urgent status. In case of an event occurs (i.e., resource allocation, maintenance, and energy conservation) on one of the edge nodes, the execution of the current task in an unstabled node is interrupted and searching for a suitable replacement node. The algorithm is based on finding the best/closest suitable node to the current node. Using the container live migration technology that relies on migrating the container between two nodes, the migration takes place from the point it was held in the parent node and resumes execution from the same point in the host node. The principle of migration is the pre-copy mechanism, and the tool used is Checkpoint Restore In User Space (CRIU). Yet, choosing the appropriate node is based on time constraints and priority. On the other hand, in another scenario, a controller node is chosen to balance system loads between edge nodes and the cloud. The controller node has a list of the status of each node in the network, through which it identifies the required node upon receiving a

request for implementation through the centralized proposed algorithm called Load Balancing in Edge and Cloud (LBEC).

The system is implemented in a virtual and realistic environment. The response time between the edge node is compared with the two proposed e-learning platforms DWP and FWP. In the experiments, the response time of the FWP platform built using the Flask framework outperformed the DWP platform built using Django. Therefore, choosing the appropriate tools for building e-learning platforms has a significant impact on the whole performance of the platform. In the case of benchmarking the proposed platforms with other platforms in the literature, such as Moodle, Coursera, and Edmodo. The benchmarking shows that the rate response time of the proposed FWP platform is superior to the other platforms.



وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم علوم الحاسوب

تحسين تدفق البيانات الكبيرة للحاويات في الوقت الحقيقي لنظام التعلم الإلكتروني عبر الانترنت في بيئة السحابة

اطروحة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة دكتوراه فلسفة في
علوم الحاسوب

من قبل

وائل وعدا الله محمود حديد

بإشراف

أ.د. ضحى بشير عبدالله

الخلاصة

ظهر اتجاه جديد للتطبيقات المتقدمة ذات المتطلبات العالية في السنوات الأخيرة. أصبح تكامل الحوسبة السحابية والبيانات الضخمة والتعليم الإلكتروني حاجة ماسة في قطاع التعليم. هذا التكامل يعطي تكامل مطلق عندما يتعلق الأمر بتوفير البنية التحتية من حيث الموقع والتكلفة والصيانة. تتمتع مناهج التعلم الإلكتروني القائمة على الحوسبة السحابية بالقدرة على توفير وقت استجابة أفضل، وحوسبة عالية، ويعطي امكانية استخدام أنظمة تشغيل غير متجانسة، وأمنية بيانات جيدة، ومرونة في نقل المستندات، مع تسهيل العمل الجماعي التعاوني. تعتبر البيانات الضخمة جانبًا ممتازًا لتطوير أنظمة الاتصالات والمعلومات. علاوة على ذلك، تعد حوسبة الحافة نوعًا جديدًا من الحوسبة التي تجعل الخدمات السحابية أقرب إلى المستخدم. بالإضافة إلى ذلك، تعمل حوسبة الحافة على تقليل زمن انتقال العميل / الخادم بشكل كبير.

في هذه الأطروحة، تم تصميم منصتين للتعلم الإلكتروني، منصة (DWP) و (FWP) باستخدام أطر عمل مختلفة، Django و Flask على التوالي. علاوة على ذلك، ترتبط كل منصة بقاعدة بيانات PostgreSQL وإطار عمل البيانات الضخمة Apache-Kafka باستخدام موصل Debezium. بالإضافة إلى نشر الأنظمة الأساسية على الخدمة السحابية Heroku التي تدعم تطبيقات مواقع الويب والحاويات. في التعلم الإلكتروني، يلعب اختيار أدوات بناء المنصة دورًا مهمًا في الحصول على أوقات استجابة أفضل. لذلك يعتبر وقت الاستجابة من العوامل المهمة المستخدمة في تقييم عمل منصات التعلم الإلكتروني، وهو ما تم إثباته في هذه الرسالة.

تم اقتراح خوارزمية جديدة غير مركزية لهجرة الحاويات العاجلة (UCM) عند الحالات الحرجة. في حالة وقوع حدث (مثل تخصيص الموارد، الصيانة، الحفاظ على الطاقة) على إحدى العقد الطرفية، يتم قطع تنفيذ المهمة الحالية في العقدة الغير مستقرة والبحث عن عقدة بديلة مناسبة. تعتمد الخوارزمية على إيجاد أفضل / أقرب عقدة مناسبة للعقدة الحالية. باستخدام تقنية الترحيل المباشر للحاوية التي تعتمد على ترحيل الحاوية بين عقدتين، يتم الترحيل عند النقطة التي كانت توقف فيها التنفيذ في العقدة الأصلية ومن ثم يستأنف التنفيذ من نفس النقطة في العقدة المضيفة. مبدأ الترحيل هو تقنية (Pre-copy)، والأداة المستخدمة هي Checkpoint Restore In User Space (CRIU). ومع ذلك، فإن اختيار العقدة المناسبة يعتمد على قيود الوقت والأولية. من ناحية أخرى، في سيناريو آخر، يتم اختيار عقدة تسيطر على موازنة أحمال النظام بين عقد الحافة والسحابة. تحتوي عقدة السيطرة على قائمة بحالة كل العقد في الشبكة، والتي من خلالها تحدد العقدة المطلوبة عند تلقي طلب التنفيذ من خلال الخوارزمية المركزية المقترحة التي تسمى موازنة التحميل في الحافة والسحابة (LBEC).

يتم تنفيذ النظام في بيئة افتراضية وواقعية. تتم مقارنة وقت الاستجابة بين عقدة الحافة بمنصتي التعلم الإلكتروني المقترحتين DWP و FWP. في التجارب، تم اعتماد وقت استجابة منصة FWP باستخدام إطار عمل Flask بسبب تفوقه على منصة DWP التي تم إنشاؤها باستخدام Django. لذلك، فإن اختيار الأدوات المناسبة لبناء منصات التعلم الإلكتروني له تأثير كبير على أداء النظام الأساسي بالكامل. في حالة مقارنة المنصات المقترحة مع منصات أخرى في الأدبيات مثل Moodle و Coursera و Edmodo. تُظهر النتائج أن معدل وقت استجابة منصة FWP المقترحة أفضل من الأنظمة الأساسية الأخرى.