تحليل أداء إستراتيجية لاختيار مراكز البيانات الأفضل لمعالجة طلبات المستخدمين في شبكات الحوسبة السحابية

م. مرح ناجي مفلح*

□ الملخّص □

يشير مصطلح الحوسبة السحابية إلى نقل المعالجة ومساحة التخزين الخاصة بالحاسوب إلى السحابة، والتي هي عبارة عن مخدم يتم الوصول إليه عبر شبكة الإنترنت، وبالتالي تقدم نموذجَ لتحقيق وصول دائم للشبكة وتحويل برامج تقنية المعلومات من منتجات إلى خدمات.

تعتمد الحوسبة السحابية على مراكز البيانات في بنيتها الأساسية، ويتم توجيه طلبات المستخدمين إلى هذه المراكز الموزعة في مناطق جغرافية مختلفة، من خلال إستراتيجيات توجيه متعددة؛ كل منها يعتمد على مجموعة من بارمترات قياس الأداء لتحسين اختيار المراكز التي ستعالج طلبات المستخدمين، وبشكل يحقق أفضل استجابة ممكنة.

يتضمن هذا البحث تحليلاً لأداء إستراتيجية لاختيار مراكز البيانات الأفضل لتوجيه طلبات المستخدمين إليها، وبيان نقاط الضعف والقوة؛ من خلال تنفيذ عدة سيناريوهات لتوزع المراكز واختبارها، باستخدام برنامجي CloudSim و CloudAnalyst.

الكلمات المفتاحية: الحوسبة السحابية، مراكز البيانات، إستراتيجية التوجيه، استهلاك الطاقة، محاكي CloudSim، برنامج .CloudAnalyst

---

*مهندسة في كلية هندسة تكنولوجيا المعلومات والاتصالات– جامعة طرطوس.

Analysis of the Performance of a Strategy to Select the Best Data Centers to Handle User Requests in Cloud Computing Networks

Eng. Marah Naji Mfleh

□ ABSTRACT □

Cloud computing refers to the transfer of processing and storage space of a computer to the cloud, which is a server accessed via the Internet, and thus provides a model for permanent network access and the transformation of IT programs from products to services.

Cloud computing relies on data centers in their infrastructure, and user requests are directed to these centers in different geographic areas through multiple routing strategies, each adopting a set of performance measurement parameters to optimize the selection of centers that will handle user requests in the best possible response.

This research includes an analysis of the performance of a strategy to select the best data centers to target users' requests and to demonstrate weaknesses and strengths through the implementation and testing of several scenarios for the distribution of centers using CloudSim and CloudAnalyst tools.

Key Words: Cloud Computing, Data Centers, Routing Strategy, Energy Consumption, CloudSim and CloudAnalyst.

*Engineer at the Faculty of Information and Communication Technology- Tartous University.

1-مقدمة

يشير مصطلح الحوسبة السحابية Cloud Computing إلى التكنولوجيا التي تعتمد على نقل المعالجة، ومساحة التخزين الخاصة بالحاسوب إلى ما يسمى بالسحابة Cloud؛ وهي جهاز مخدم يتم الوصول إليه عبر شبكة الإنترنت، وبهذا تتحول برامج تكنولوجيا المعلومات من منتجات إلى خدمات، بالتالي تساهم هذه التكنولوجيا في إبعاد مشاكل صيانة وتطوير برامج تقنية المعلومات عن الشركات المستخدمة لها، أي يتركز مجهود الجهات المستفيدة على استخدام هذه الخدمات فقط. [1]

تعتمد البنية التحتية للحوسبة السحابية على مراكز البيانات Data Centers (DCs) المتطورة، والتي تقدم مساحات تخزين كبيرة للمستخدمين، كما أنها توفر بعض البرامج كخدمات، وتعتمد في ذلك على الإمكانيات التي وفرتها تقنيات ويب 2,0. [2] [1]

تتصل مراكز البيانات مع بعضها على شكل شبكة Data Center Networks (DCNs) تزود بمجال واسع من خدمات الحوسبة السحابية، وبشكل مشابه لاكتشاف البنية التحتية لمراكز البيانات فإننا بحاجة لإستراتيجية توجيه فعالة لتسيير طلبات المستخدمين فيما بين المراكز والمخدمات التي تتكون منها هذه المراكز. [3]

ينبغي على هذه الإستراتيجية أن تأخذ بالحسبان العديد من البارامترات المؤثرة في جودة الخدمة المقدمة؛ مثل التأخير الزمني (زمن الاستجابة وزمن المعالجة)، الموثوقية، التدفق، استهلاك الطاقة وغيرها. كذلك قد تظهر لدينا حاجة للإرسال المجموعاتي Multicast لعدة مخدمات مركزية في الشبكة أو القيام بالتوجيه على عدة مسارات من أجل تحقيق ما يسمى بموازنة الحملLoad Balancing. بناءً على ما سبق يجب على إستراتيجية التوجيه أن تختار مركز البيانات الذي سيعالج طلبات المستخدم بأفضل استجابة ممكنة. [4]

يوجد العديد من إستراتيجيات اختيار مراكز البيانات، منها يعتمد على اختيار المركز الأقرب للمستخدم service proximity based routing ويتم الاختيار بشكل عشوائي في حال وجود أكثر من مركز، ومنها يعتمد على تأخير أو استجابة الشبكة لتحديد المركز الأفضل Performance optimized routing، ومنها يحدد نشر التطبيق بناءً على الحمل الذي يواجهه المركز حالياً Dynamically reconfigure routing. [7] [6] [5]

2-هدف البحث

يهدف هذا البحث إلى تحليل أداء إستراتيجية تعمل على اختيار مركز البيانات الأكثر ملاءمة لمعالجة طلبات المستخدم بناءً على عدة خطوات، بحيث يتم توزيع جميع الطلبات على مراكز البيانات باستخدام مسارات الشبكة الأقل ازدحاماً وتنفيذها بأزمنة استجابة أصغرية؛ لتحقيق جودة الخدمة المطلوبة لكل تطبيق، ودراسة بعض القيم

الإحصائية لنتائج الاختبار (زمن المعالجة وزمن الاستجابة) والاستفادة منها في إيجاد إمكانية تنبؤ لإحداها انطلاقاً من معرفة الأخرى.

3-طرائق البحث ومواده

جرى التزويد بالعديد من تقنيات المحاكاة لبيئات الحوسبة السحابية؛ لكونه من الصعب والمكلف. كذلك قياس الأداء باستخدام منصات الحوسبة الحقيقية [8].

يعد محاكي CloudSim من أهم هذه التقنيات وهو عبارة عن أداة لنمذجة ومحاكاة أنظمة الحوسبة السحابية بالاعتماد على لغة البرمجة جافا، يؤمن مجموعة المكونات الأساسية لبيئة الحوسبة السحابية مثل الآلات الافتراضية (المخدمات) (Virtual Machines (VMs، طلبات المستخدمين Cloudlets، مراكز البيانات (DCs)، وغيرها. من تقنيات المحاكاة كذلك محاكي CloudAnalyst وهو عبارة عن أداة مفتوحة المصدر مبني على أداة CloudSim بحيث يقدم مجموعة من الواجهات التفاعلية والأدوات لمحاكاة بيئة الحوسبة السحابية. [9] [10]

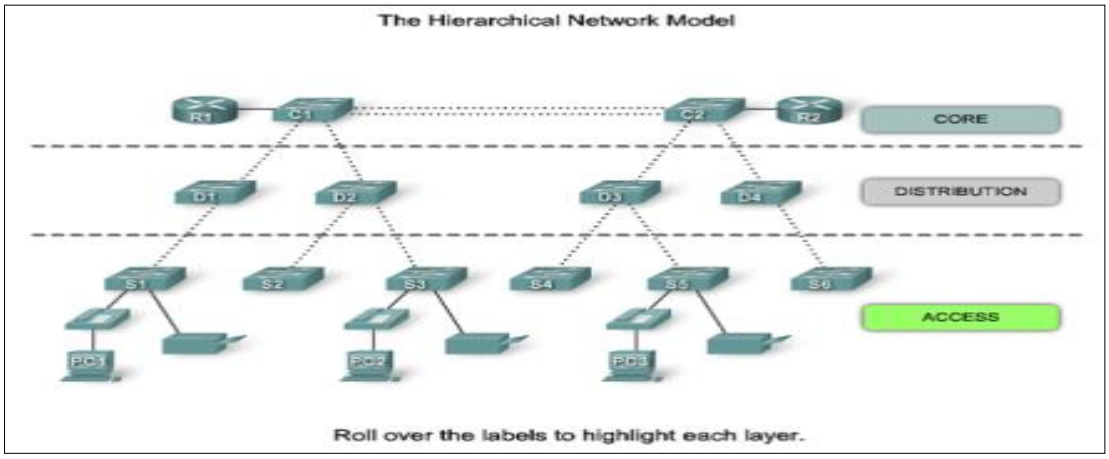
جرى إنجاز هذا البحث باستخدام كل من محاكي CloudSim وCloudAnalyst وبالاعتماد على مراجع علمية عديدة في هذا المجال.

3-1 الحوسبة السحابية ومراكز البيانات

يمكن استعراض الحوسبة السحابية على أنها مجموعة من الخدمات التي يمكن تقديمها كبنية طبقات. والخدمات المُقدمة من خلال الحوسبة السحابية تشتمل على:

- ✓ خدمات تكنولوجيا المعلومات والتي يُشار إليها بـ SaaS (برنامج كخدمة)، تظهر في أعلى الحزمة. وتسمح SaaS للمستخدمين بتشغيل التطبيقات عن بعد من خلال الخدمات السحابية.
- ✓ البنية التحتية كخدمة (IaaS) وتشير إلى مصادر الحوسبة كخدمة. يشمل الحواسيب الظاهرية ذات قوة المعالجة المضمونة ومعدل نقل البيانات المحجوز للتخزين ودخول الإنترنت.
- ✓ المنصة كخدمة (PaaS) وهي مشابهة لطبقة IaaS، لكنها تتضمن أيضاً أنظمة التشغيل والخدمات المطلوبة لتطبيق معين. بمعنى آخر، PaaS هي IaaS مع مجموعة برمجيات مخصصة لتطبيق معين.
- ✓ تخزين البيانات كخدمة (dSaaS) توفر مساحة تخزين يستخدمها المستهلك بما في ذلك متطلبات معدل نقل البيانات للتخزين. [6] [2]

تعتمد البنية التحتية للحوسبة السحابية على مراكز البيانات، والتي تتصل مع بعضها على شكل شبكة كما ذكرنا في المقدمة. تتكون أية شبكة عالية الأداء لمركز البيانات من ثلاث طبقات: تسمى الأعلى منها طبقة النواة أو القلب للشبكة (core layer)، وتضم الموجهات والمحولات التي تربط الشبكة بالعالم الخارجي، ويمر عبرها كل



البيانات الداخلة والخارجة من الشبكة. يشترط في مكونات هذه الطبقة أن تكون عالية الأداء والسرعة لأنها يجب أن تستوعب كل المرور عبر الشبكة ولمختلف التطبيقات ولأعداد متفاوتة من المستخدمين. أما الطبقة الثانية الوسطية فهي طبقة التوزيع (distribution layer) ووظيفتها ربط طبقة النواة مع طبقة الوصول. أهم خصائص أجهزة هذه الطبقة هو احتوائها بوصفها موجهات أو محولات على عدد كبير من المنافذ لتوسيع الشبكة والسماح بالتوسيع المستقبلي. أخيراً تضم الطبقة القريبة من المستخدمين والمسماة طبقة الوصول (access layer) عدداً من المبدلات التي ترتبط مباشرة بحواسيب المستخدمين كما في الشكل (1) [11] .

الشكل (1): طبولوجيا شبكة مراكز البيانات [11]

3-2 إستراتيجية اختيار مراكز البيانات

ينبغي على إستراتيجية التوجيه القيام باختيار مركز البيانات الأكثر ملاءمةً لمعالجة طلبات المستخدمين، ومن ثم تخصيص الطلبات للآلات الافتراضية الشاغرة والأسرع ضمن المركز المختار؛ من أجل تحقيق أفضل استجابة ممكنة، حيث إن زمن استجابة عالٍ يعني أن مركز البيانات يتعرض للتحميل الزائد، وبالتالي انخفاض في جودة الخدمة. يمكن تقليل زمن الاستجابة باختيار مسار الشبكة الأقل ازدحاماً، ولكن اختيار هذا المسار مراراً وتكراراً سيجعله مزدحماً مع الوقت، وأيضاً يجب ألا ننسى أن مركز البيانات المختار سيكون لديه العديد من الطلبات التي تنتظر تخديمها؛ ويجب أخذ هذا الأمر بالحسبان عند اختيار المركز [14] [13] [12].

تعمل إستراتيجية التوجيه كما يأتي:

1. إعداد مصفوفة التأخير الزمني لكل مركز من مراكز البيانات.
2. إعداد مصفوفة عرض الحزمة المتاحة لكل مركز من مراكز البيانات.
3. إعداد مصفوفة سرعة المعالجة لكل آلة افتراضية (مخدم) VM ضمن كل مركز من مراكز البيانات، وترتيبها تنازلياً بالاعتماد على قيمة MIPS.

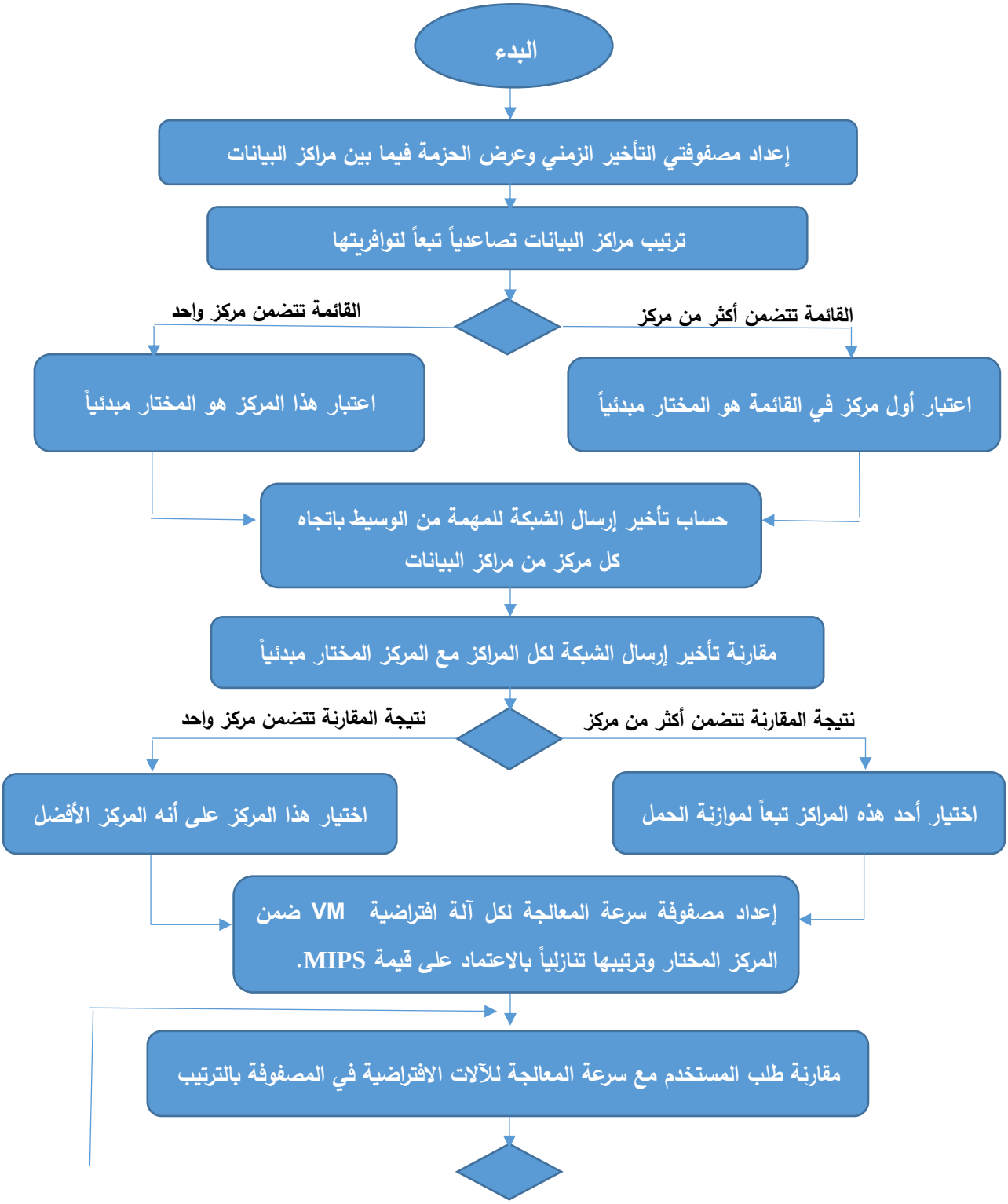
4. حساب نسبة التوافر للمراكز؛ وهي عبارة عن نسبة التأخير الزمني على عرض الحزمة.
5. ترتيب المراكز تصاعدياً ضمن قائمة بناءً على نسبة التوافر.
6. حساب تأخير إرسال الشبكة لطلب المستخدم باتجاه كل مركز من مراكز البيانات باستخدام المعادلة: $T = y + (s/b)$، حيث إن T هي عبارة عن تأخير إرسال الشبكة لطلب مستخدم بحجم S على مسار الشبكة الذي لديه قيمة تأخير مساوية إلى y وعرض الحزمة المتاحة هو b. [15] [12]
7. ترتيب المراكز تصاعدياً ضمن قائمة بناءً على تأخير إرسال الشبكة.
8. مقارنة قائمتي التوافر والتأخير للمراكز بالترتيب لإختيار المركز الأفضل توافرية والأقل تأخيراً في الوقت نفسه.
9. في حال كانت نتيجة المقارنة:
   a. مركز واحد: يتم اعتماده لمعالجة طلب المستخدم.
   b. وإلا: فيتم اعتماد أحد المراكز المتشابهة من حيث التأخير والتوافرية بناءً على سياسة موزانة الحمل المطبقة.
10. بعد اختيار المركز، مقارنة طلب المستخدم مع سرعة المعالجة للآلات الافتراضية الموجودة بدءاً من أول آلة في المصفوفة:
   a. في حال عدم مشغولية الآلة الأسرع وطلب المستخدم أقل من عتبة سرعة المعالجة لديها يتم تخصيص الطلب لها وتحديث حالتها إلى مشغولة.
   b. وإلا في حال طلب المستخدم أكبر من العتبة، ننتقل لمقارنة طلب المستخدم مع الآلة التالية غير المشغولة في المصفوفة.
   c. في حال مشغولية الآلة الأسرع ننتقل للآلة التالية في المصفوفة ونقارن معها وهكذا.
   d. في حال عدم مشغولية أية آلة يتم انتظار فراغ أحدها.

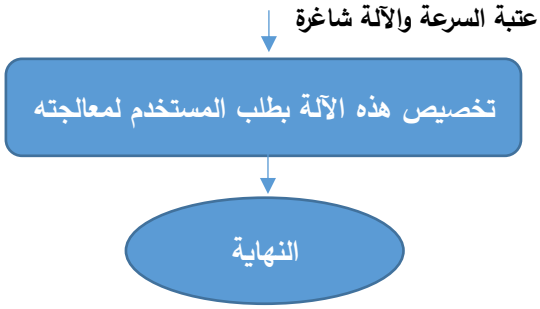
الهدف الرئيس من الإستراتيجية المتبعة هو تحقيق زمن الاستجابة الأقل مع الحد الأدنى من الكلفة وزمن المعالجة لتحقيق الأداء الفعال لعملية التوجيه لمراكز البيانات في الشبكات السحابية، مع الأخذ بالاعتبار حجم طلب المستخدم، عرض الحزمة وتأخير الشبكة.

تقاس سرعة المعالجة للآلات الافتراضية باستخدام بارامتر MIPS (Million Instructions per Second) وبالتالي كلما كانت قيمة هذا البارامتر أكبر كان أفضل [16].

يتم تحقيق جزء الإستراتيجية المرتبط باختيار الآلات الافتراضية المناسبة لمعالجة الطلبات من خلال تطبيق خوارزمية Throttled Load Balancer Algorithm (TLB) لموازنة الحمل ضمن مراكز البيانات، والتي تعمل على تخصيص عدد محدد مسبقاً من الطلبات لكل VM ولزمن معين. في حال كان عدد المهام أكبر من قدرة المعالجة للآلة الافتراضية المتاحة، تضع الخوارزمية الطلبات القادمة ضمن رتل بحيث تنتظر فراغ الآلة

الافتراضية التالية. مهمة موازن الحمل هذا هو تحديد جدول للآلات الافتراضية وبيان حالة كل منها متاح أو مشغول. بحيث عندما يقوم المستخدم بتقديم طلب إلى مركز البيانات لتحديد مدى توفر آلة افتراضية مناسبة لتنفيذ مهمة معينة، يقوم موازن الحمل بفحص جدول VMs للعثور على آلة افتراضية مناسبة لمعالجة البيانات [18]. تخصيص طلبات المستخدم للآلة الأسرع المتاحة يقلل من زمن المعالجة وينهي الطلب بسرعة أكبر مما يحسن من زمن الاستجابة أيضاً. يبين المخطط الصندوقي الآتي توضيحاً لخطوات هذه الإستراتيجية.





3-3 استهلاك الطاقة

ينتشر أكثر من 500 مركز بيانات حول العالم، وتشغيل هذه المراكز يتطلب استهلاك كمية كبيرة من الطاقة، والتي تفسر شريحة كبيرة من كلفة التشغيل الكلية، وكذلك تسبب انبعاث كمية هائلة من إشعاعات غاز ثاني أوكسيد الكربون CO_2 سنوياً.

ينبغي ضمن مراكز البيانات المثالية وصول كل الطاقة إلى تجهيزات الحوسبة من أجل تنفيذ طلبات المستخدمين، ومن ثم يتم تقسيم هذه الطاقة فيما بين عملية الاتصال والمكونات الصلبة.

تعتمد معظم الحلول من أجل تحقيق فعالية الطاقة على تخفيض تردد التشغيل (أو معدل الإرسال)، أو إطفاء الجهاز بشكل كامل أو مكوناته من أجل حفظ الطاقة بالنسبة إلى مخدمات الحوسبة. وقد بين عدد من الدراسات أن تحسيناً بسيطاً في هيكلية مركز البيانات والجدولة المدركة للطاقة energy-aware scheduling يمكن أن تؤدي وبشكل ملحوظ إلى حفظ طاقة بحوالي 75%، اعتماداً على تقنيات إدارة الحركة والحمل.

من ناحية مستخدم السحابة، تصبح الطاقة القلق الأعظم. بحيث يستخدم الكثير من مستخدمي السحابة التجهيزات النقالة (smart phones, laptops, tablet PCs) للوصول إلى خدمات السحابة. والطريقة الوحيدة الفعالة لحفظ طاقة هذه الأجهزة المزودة بالبطاريات هي إطفاء معظم المكونات الرئيسة، بما فيها المعالج المركزي، المرسل/المستقبل وكذلك الذاكرة. والهدف هو تخفيض زمن معالجة الطلب وبالتالي استهلاك طاقة أقل من البطارية [18].

يوجد العديد من التقنيات الفعالة في الحفاظ على طاقة المراكز وتقليل الاستهلاك بشكل ملحوظ. بالنسبة إلى إستراتيجية الاختيار السابقة يتم تحقيق ذلك بوضع الآلات الافتراضية غير المشغولة بوضع Sleep Mode الأمر الذي يعتمد على قضية توزيع الحمل واختيار الآلات الافتراضية التي ستخدم الطلبات، بحيث يتم اختيار إحدى هذه الآلات لمعالجة الطلب وبعد الانتهاء منه يتم تفعيل وضع Sleep في انتظار طلب آخر. في حال كانت الطلبات قصيرة فيتم إنجازها بسرعة من قبل أسرع الآلات الافتراضية بناءً على خطوات الإستراتيجية؛ وبذلك تبقى عدة آلات غير سريعة بوضع Sleep لمدَّة زمنية طويلة مما يقلل من استهلاك الطاقة بشكل عام

ضمن المركز. وكذلك استخدام الآلات الأسرع لمعالجة الطلبات سيمكن من إنجاز العمل بزمن أقل واستهلاك طاقة وكلفة تشغيل أقل. [18]

السيناريو الأول	ست مراكز بيانات 6DCs (مركز في كل منطقة) ووحدة مستخدم واحدة 1UB
السيناريو الثاني	ست مراكز بيانات 6DCs(مركز في كل منطقة) وثلاث وحدات مستخدم موزعة 3UBs
السيناريو الثالث	أربع مراكز بيانات 4DCs وست وحدات مستخدم موزعة 6UBs
السيناريو الرابع	ست مراكز بيانات 6DCs وست وحدات مستخدم 6UBs(مركز ووحدة في كل منطقة)

3-4 التطبيق العملي

3-4-1 سيناريوهات الاختبار

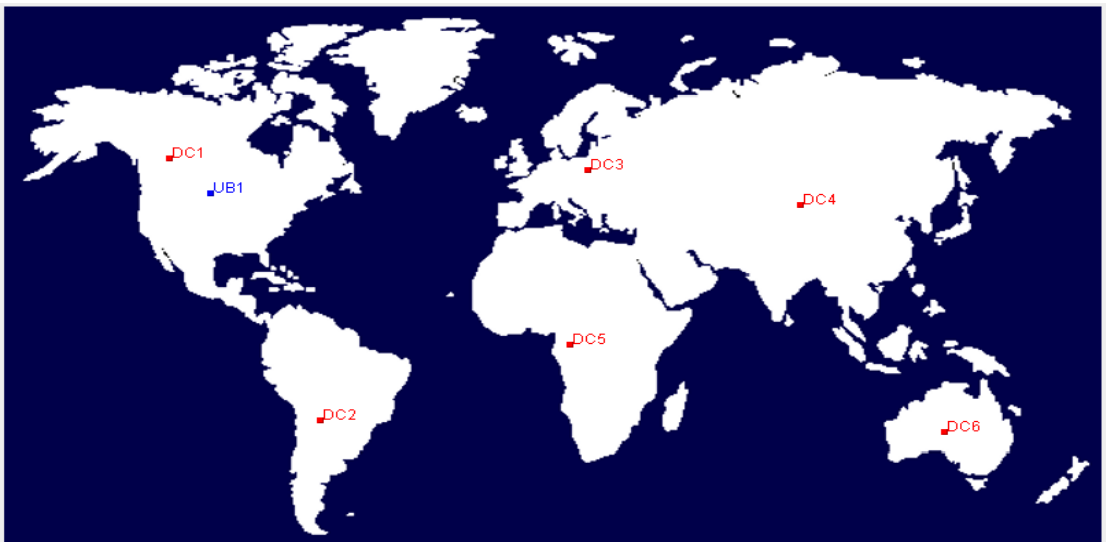
يقسم محاكي CloudAnalyst العالم إلى ست مناطق (القارات الستة)، ويتم توزيع مراكز البيانات ووحدات المستخدمين ضمن هذه المناطق من خلال ضبط إعدادات البرنامج بما يلائم الحالة المدروسة [9]. يشمل الجدول (1) توصيف السيناريوهات التي قمنا بإنشائها ضمن محاكيي CloudSim و CloudAnalyst من أجل القيام بالمحاكاة؛ بحيث أخذنا بالحسبان مواقف مختلفة (مثل توزع المراكز ووحدات المستخدم، تغير الحمل وأحجام طلبات المستخدم). يشمل كذلك كل من الجدولين (2) و (3) أهم الإعدادات التي طبقت على السيناريوهات المدروسة والتي تم اختيارها اعتماداً على عدد من المراجع، وأي إعداد لم يذكر بقي افتراضياً كما هو في البرنامج.

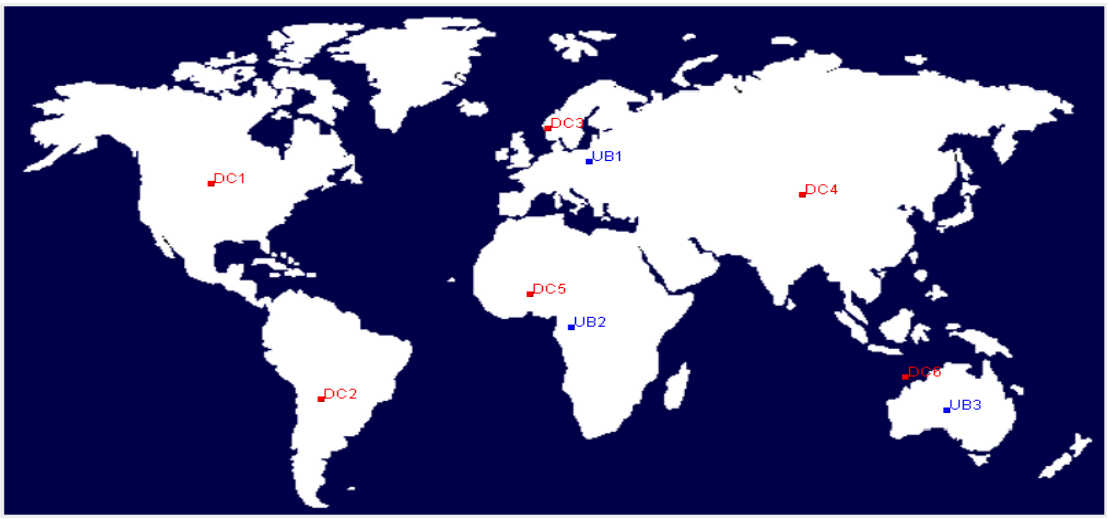
الجدول (1): سيناريوهات المحاكاة

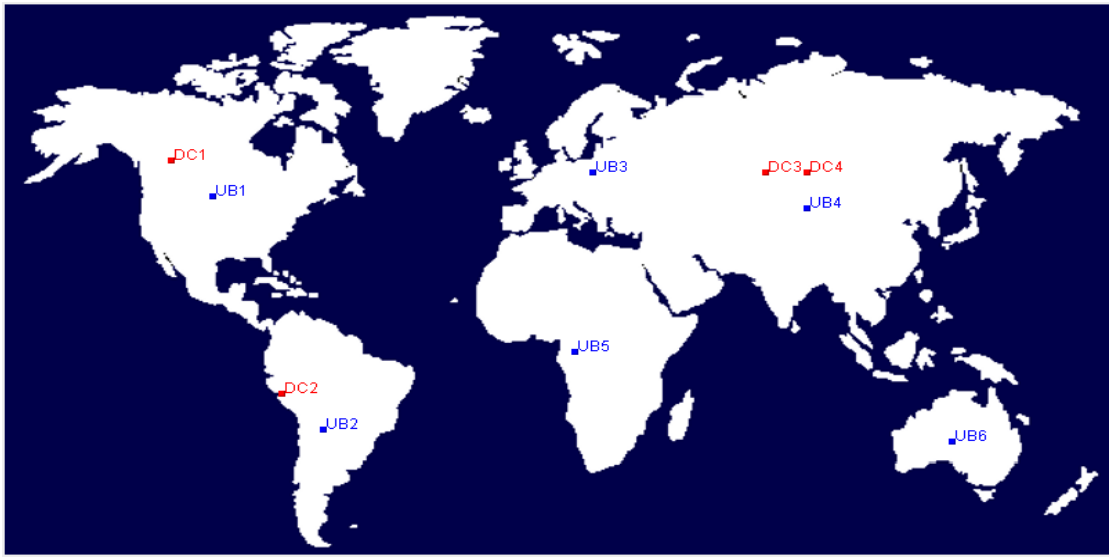
تظهر الأشكال الآتية (2) (3) (4) (5) السيناريوهات كما تبدو في البرنامج.

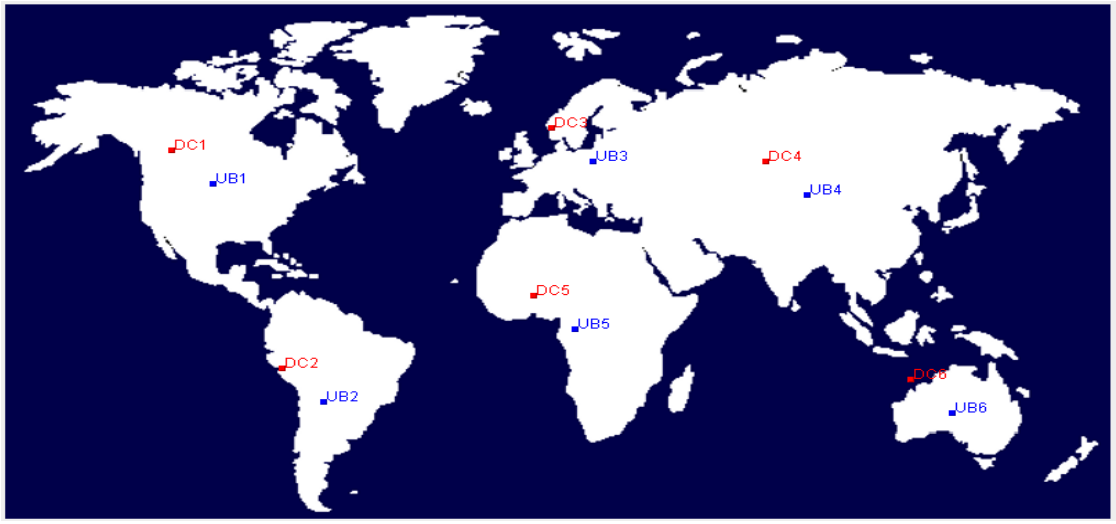
الشكل (2): السيناريو الأول

الشكل (3): السيناريو الثاني









الشكل (4): السيناريو الثالث

الشكل (5): السيناريو الرابع

الجدول (2): أهم الإعدادات المطبقة في مراكز البيانات والعتاد الفيزيائي

إعدادات نشر تطبيق السحابة والعتاد الفيزيائي للمراكز					
No. of VMs	Image size (MB)	Memory (MB)	BW (MB)	Storage (MB)	VM Policy
5	10000	204800	1000000	100000000	TIME_SHARED

الجدول (3): حجم طلبات المستخدم وتوزع وحدات المستخدمين

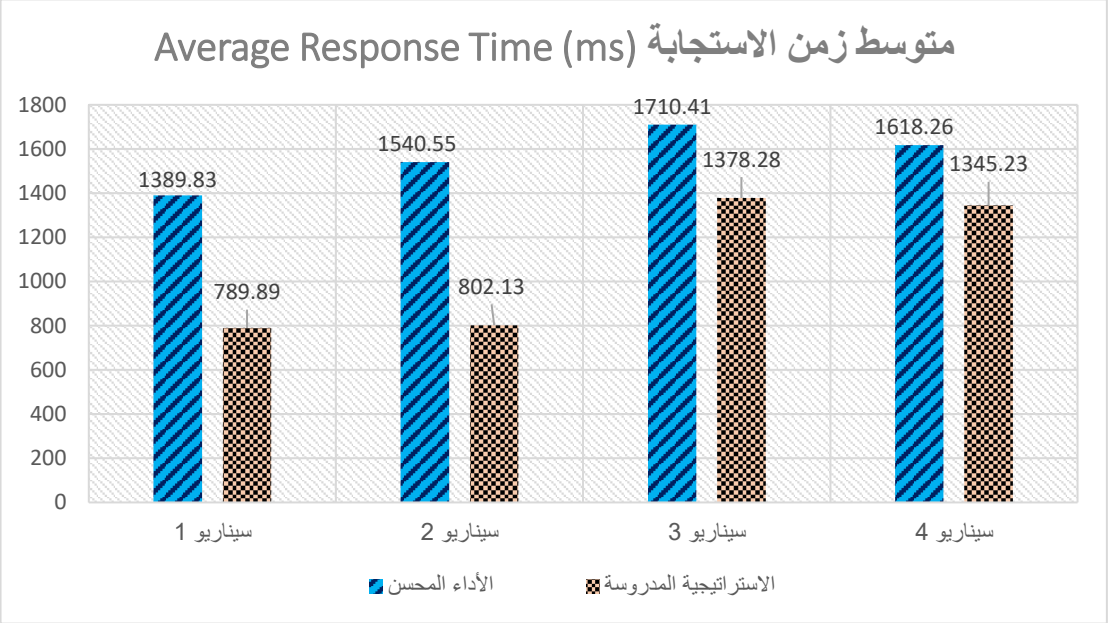
Scenarios	No. of UBs	Region	Request Size (KB)
1	1	0	100
2	3	2, 4, 5	10, 20, 30

3	6	0, 1, 2, 3, 4, 5	1, 100, 50, 60, 40, 10
4	6	0, 1, 2, 3, 4, 5	1, 30, 40, 150, 50, 10

3-4-2 نتائج الاختبار

نفذت السيناريوهات الأربعة لمدة محاكاة يوم كامل لأخذ ساعات الذروة بالحسبان، والتي تبدأ من الساعة 3 GMT وتنتهي في الساعة 9 GMT كما هي افتراضية ضمن البرنامج، وبحيث نفذ كل سيناريو من أجل تقييم زمن الاستجابة، زمن المعالجة، والكلفة الكلية للإستراتيجية المدروسة، مقارنةً مع إستراتيجية الأداء المحسن ضمن بيئة Cloud Analyst.

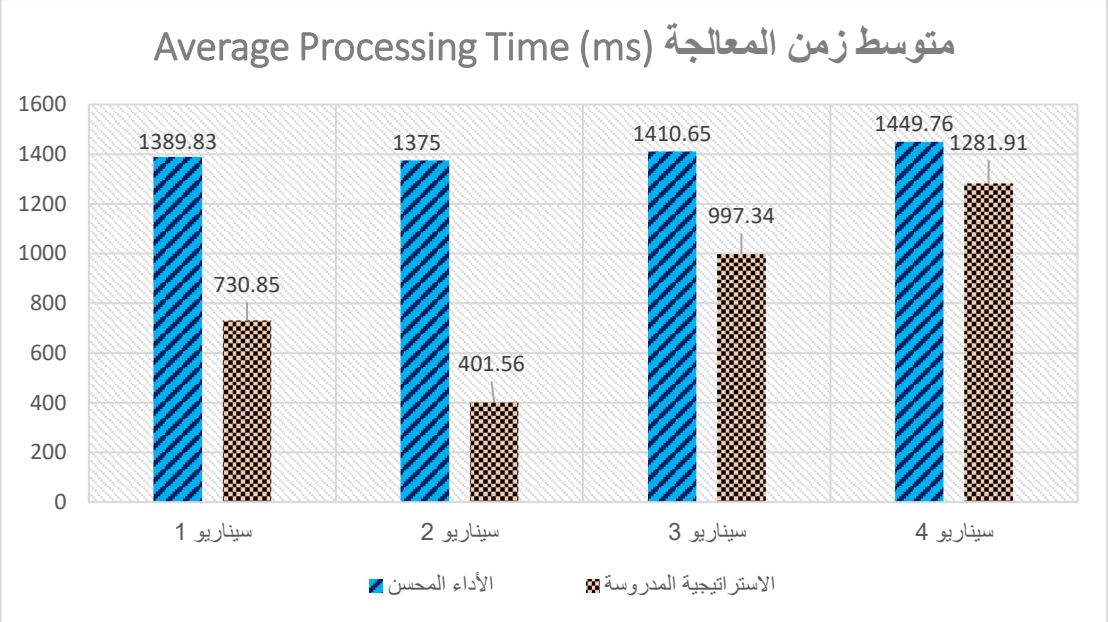
تعمل إستراتيجية الأداء المحسن على اختيار المسار الأفضل باتجاه مركز البيانات اعتماداً على تأخير (استجابة) الشبكة Network Latency، وكذلك الحمل على هذا المركز للحصول على أفضل استجابة ممكنة اعتماداً على زمن الاستجابة لآخر طلب. جرى اختيار هذه الإستراتيجية لمقارنة الإستراتيجية المدروسة معها بنتيجة تقييم سابق لإستراتيجيات التوجيه لمراكز البيانات الموجودة ضمن برنامج CloudAnalyst والتي بيّنت كون إستراتيجية الأداء المحسن أفضلهم أداءً من حيث زمن الاستجابة وزمن المعالجة [19].



ظهرت نتائج الاختبار لمقارنة زمن الاستجابة كما في الشكل (6) الذي بيّن انخفاض قيمة متوسط زمن الاستجابة للإستراتيجية المدروسة مقارنةً مع إستراتيجية الأداء المحسن.

الشكل (6): متوسط زمن الاستجابة

حيث إن زمن الاستجابة Response Time هو عبارة عن تخمين الوقت المطلوب من اللحظة التي يرسل فيها المستخدم طلباً إلى مركز بيانات، إلى اللحظة التي يبدأ فيها المستخدم باستقبال النتائج. [14]



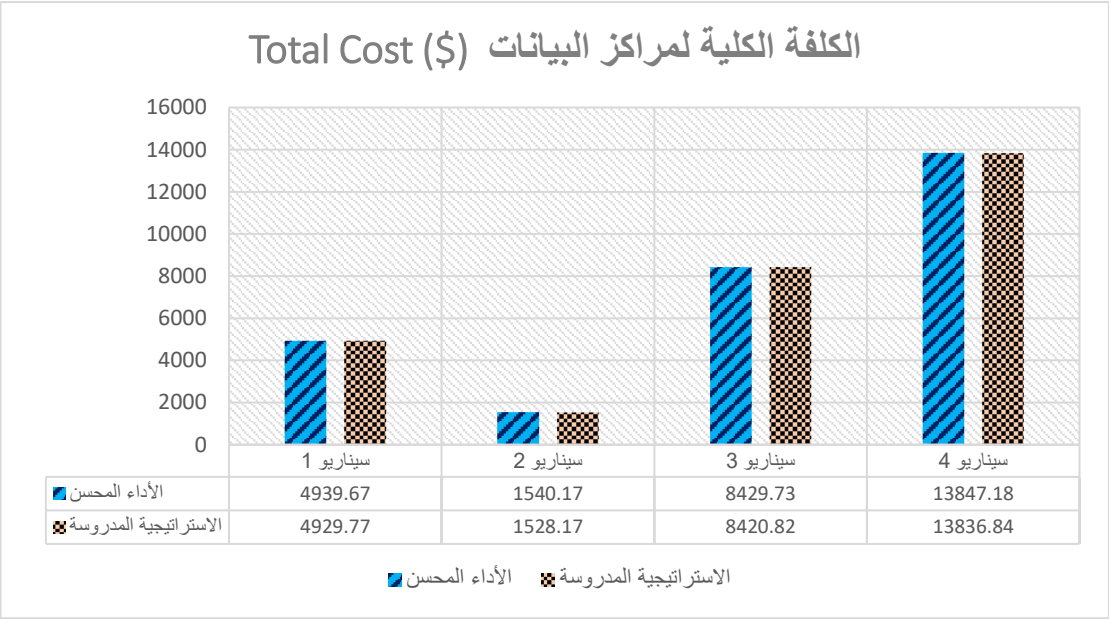
فيما يخص زمن المعالجة ضمن مراكز البيانات فقد بيّنت المقارنة كما في الشكل (7) انخفاض قيمة زمن المعالجة للإستراتيجية المدروسة وفي جميع السيناريوهات مقارنةً مع إستراتيجية الأداء المحسن.

الشكل (7): مقارنة زمن المعالجة

أما الكلفة الكلية فتتكون من مجموع كلفتين هما كلفة استخدام الآلات الافتراضية وكلفة نقل البيانات إلى مراكز البيانات [20]، وتظهر قيم هذه الكلف للسيناريوهات الأربعة كما في الجدولين (4) و (5).

	الأداء المحسّن	الإستراتيجية المدروسة
السيناريو الأول	72.02	62.12
السيناريو الثاني	72.01	60.01
السيناريو الثالث	48.01	39.10
السيناريو الرابع	72.01	61.67

الجدول (4): كلفة استخدام الآلات الافتراضية



الجدول (5): كلفة نقل البيانات

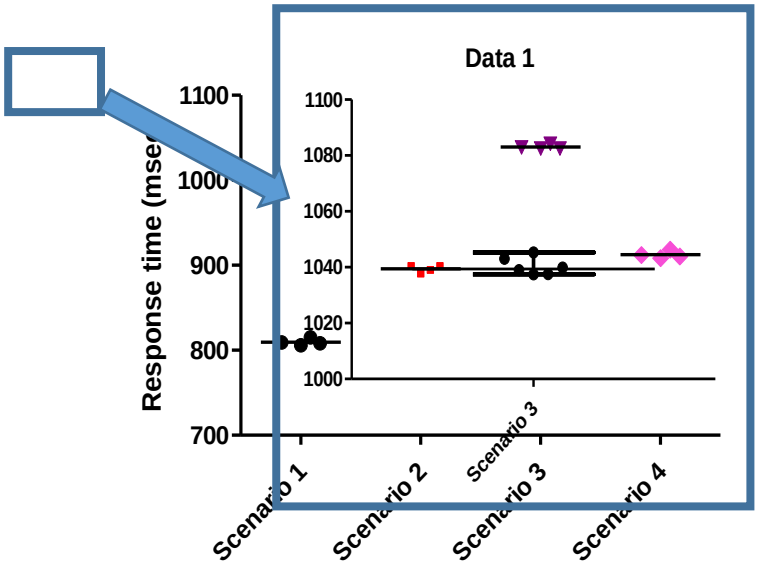
يظهر من الجدولين السابقين بأن الإستراتيجية المدروسة أعطت قيم أقل للكلفة في أغلب السيناريوهات مقارنةً مع الإستراتيجية الأخرى، وبأخذ المجموع للكلفتين الجزئيتين تنتج لدينا الكلفة الكلية كما هو مبين في الشكل (8).

الشكل (8): مقارنة الكلفة الكلية

	سياسة الأداء المحسّن	السياسة المقترحة
السيناريو الأول	4867.65	4867.65
السيناريو الثاني	1468.16	1468.16
السيناريو الثالث	8381.72	8381.72
السيناريو الرابع	13775.17	13775.17

نجد من القيم الناتجة للكلفة تخفيضاً في كلفة استخدام الآلات الافتراضية للإستراتيجية المدروسة وهذا بسبب استخدام الآلة الافتراضية (المخدم) الأسرع غير المشغولة، وبالتالي إنجاز الطلب بوقت استخدام أقل للمخدمات الأمر الذي يعني كلفة أقل. أما كلفة نقل البيانات فهي متماثلة فيما بين الإستراتيجيتين لكونها تعتمد على إعدادات الشبكة وعرض الحزمة على مساراتها.

تمت عملية الاختبار لهذه الإستراتيجية بتكرار المحاكاة عدداً من المرات؛ وذلك للتحقق من ثبات النتائج التي تم الحصول عليها. يبين الشكل (9) قيم زمن الاستجابة لعدة تكرارات للمحاكاة.



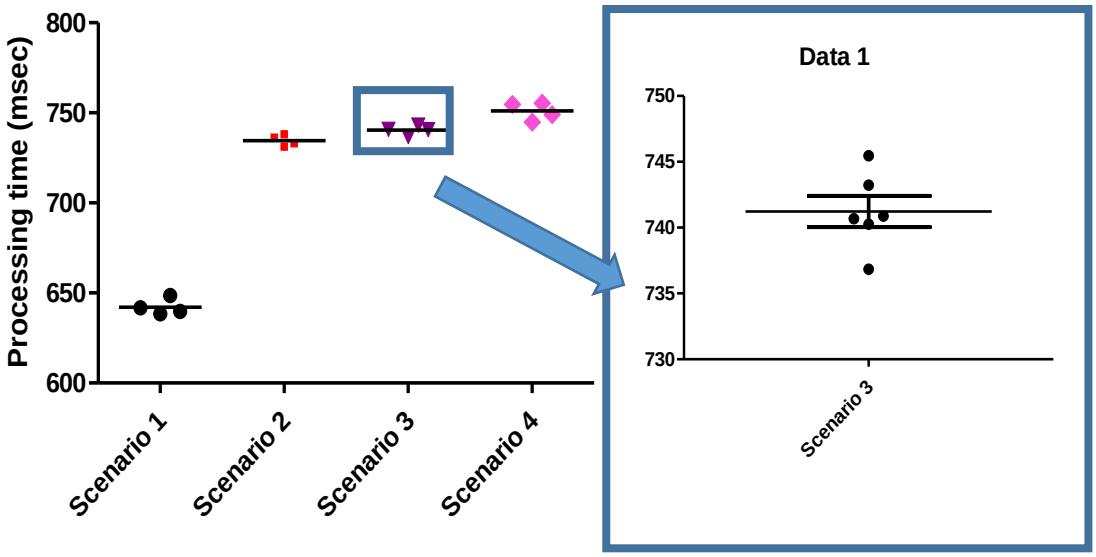
الشكل (9): زمن الاستجابة لعدة تكرارات للمحاكاة

نجد أن قيم زمن الاستجابة لكل سيناريو ولعدة تكرارات تتراوح ضمن مجال صغير نسبياً (تم تقريب نتائج تكرار السيناريو الثالث)؛ وهذا يدل على ثبات القيم الناتجة من المحاكاة لهذه الإستراتيجية. يظهر الجدول (7) بعض القيم الإحصائية لنتائج المحاكاة ولجميع السيناريوهات من أجل زمن الاستجابة.

الجدول (6): قيم زمن الاستجابة

	السيناريو 1	السيناريو 2	السيناريو 3	السيناريو 4
أصغر قيمة	805.9	890.2	1037	908.7
أكبر قيمة	815.0	899.2	1043	918.6
المتوسط	809.5	895.7	1039	912.5

وكذلك الأمر بالنسبة إلى زمن المعالجة بحيث بيّنت عملية المحاكاة تقارباً كبيراً للنتائج عند كل تكرار، كما في الشكل (10).



الشكل (10): زمن المعالجة لعدة تكرارات للمحاكاة

يظهر الجدول (7) بعض القيم الإحصائية لنتائج المحاكاة ولجميع السيناريوهات من أجل زمن المعالجة.

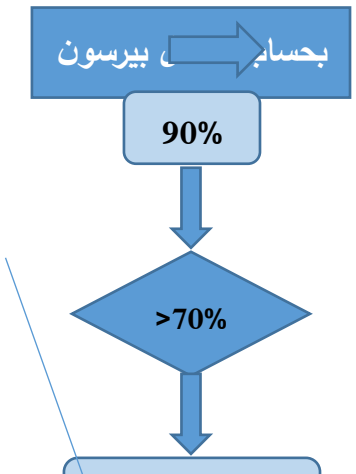
الجدول (7): قيم زمن المعالجة

	السيناريو 1	السيناريو 2	السيناريو 3	السيناريو 4
أصغر قيمة	638.5	731.2	736.9	745.0
أكبر قيمة	648.5	738.1	743.2	755.4
المتوسط	642.1	734.7	740.4	751.1

وبدراسة الترابط بين قيم زمن الاستجابة وزمن المعالجة السابقة لمعرفة طبيعة العلاقة بين هذين المتغيرين، ومدى قوتها أو انعدامها، وجدنا وبحساب معامل بيرسون الذي يعتبر من أكثر معاملات الارتباط استخداماً [21]، وجدنا قيمة ارتباط مساوية إلى 90% والتي تعبر عن ارتباط قوي (كونها أكبر من 70%) كما يبدو في الشكل (11) بأخذ مجموعة من القيم المتقابلة الناتجة من التكرارات لعملية المحاكاة.

زمن الاستجابة
1037.55
1038.96
1037.44
1043.05
1039.84
1045.27

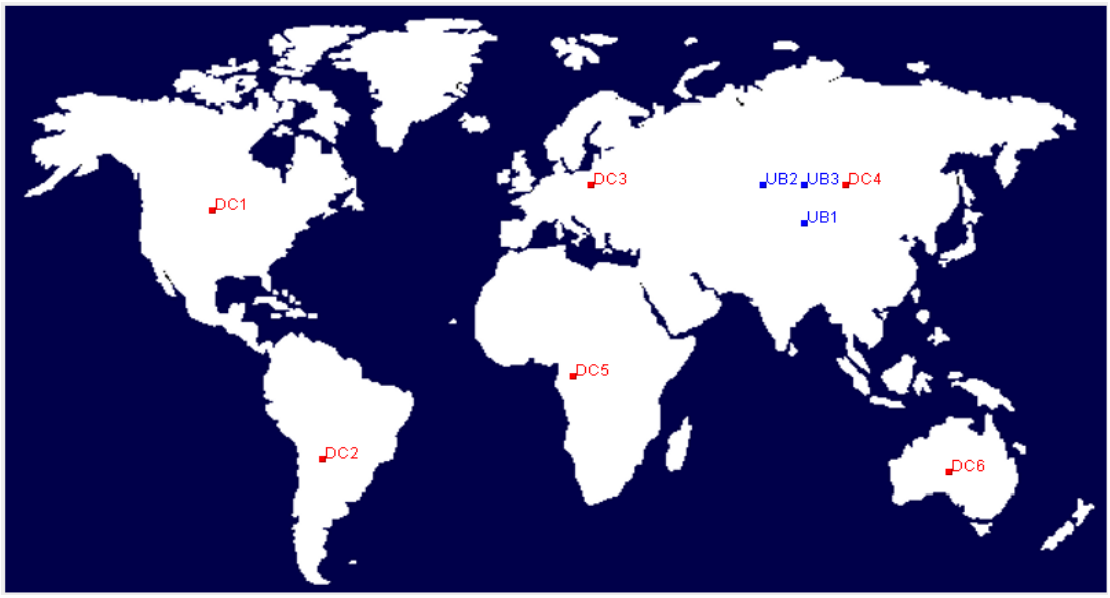
زمن المعالجة
740.68
740.88
736.[illegible]
743.22
740.25
745.45



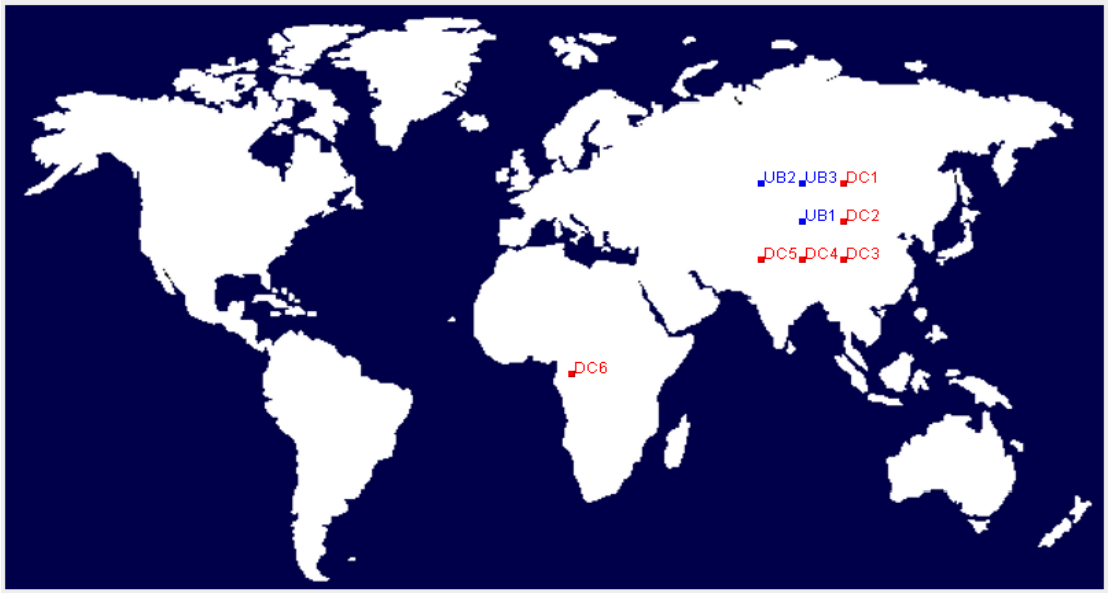
الشكل (11): دراسة الارتباط والتنبؤ

يمكننا بالتالي الاستفادة من هذا الارتباط القوي وباستخدام تابع FORECAST يمكننا التنبؤ بقيمة زمن الاستجابة انطلاقاً من قيمة زمن معالجة معطى، الأمر الذي يمكننا من توقع التحسين الممكن الحصول عليه في زمن الاستجابة لطلب المستخدم في حال إجراء تحسين معين في زمن المعالجة.

في حال قمنا بإنشاء سيناريو خامس مشابه للسيناريو الثاني ولكن وحدات المستخدم الثلاث تكون جميعها في منطقة واحدة كما في الشكل (12)، وسيناريو سادس مشابه للسيناريو الخامس ولكن تتوضع فيه خمس مراكز بيانات ضمن نفس المنطقة مع وحدات المستخدم الثلاث كما في الشكل (13).

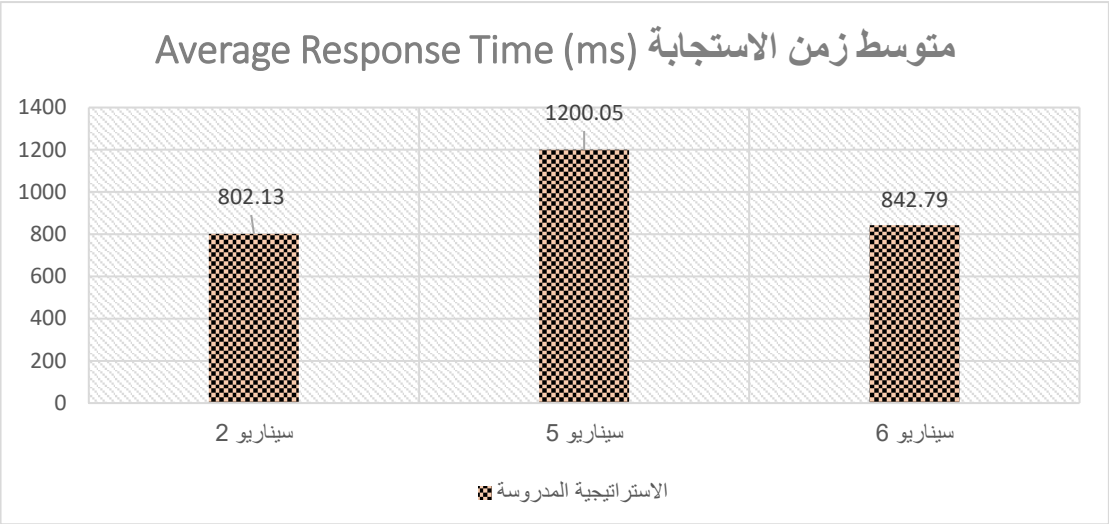


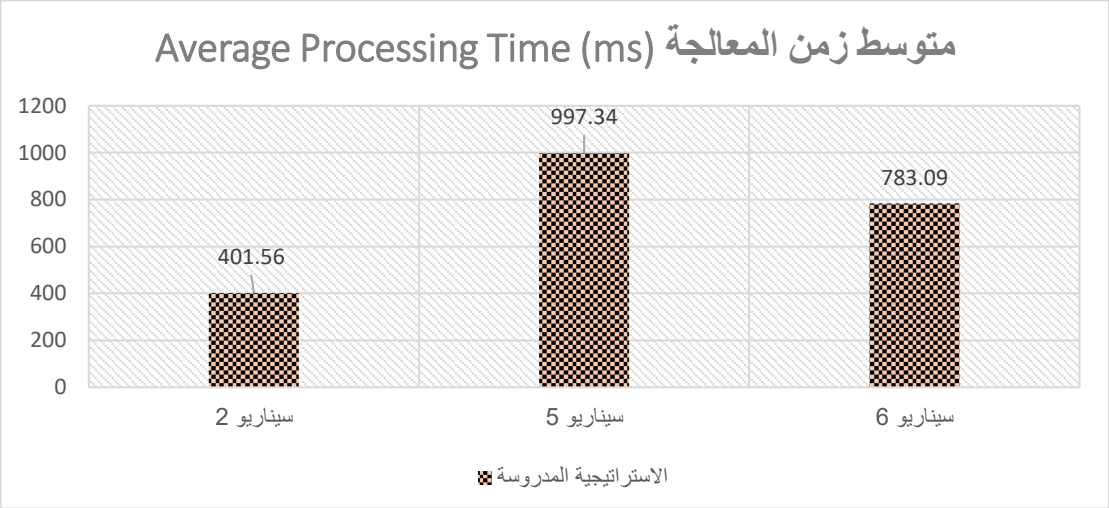
الشكل (12): السيناريو الخامس.

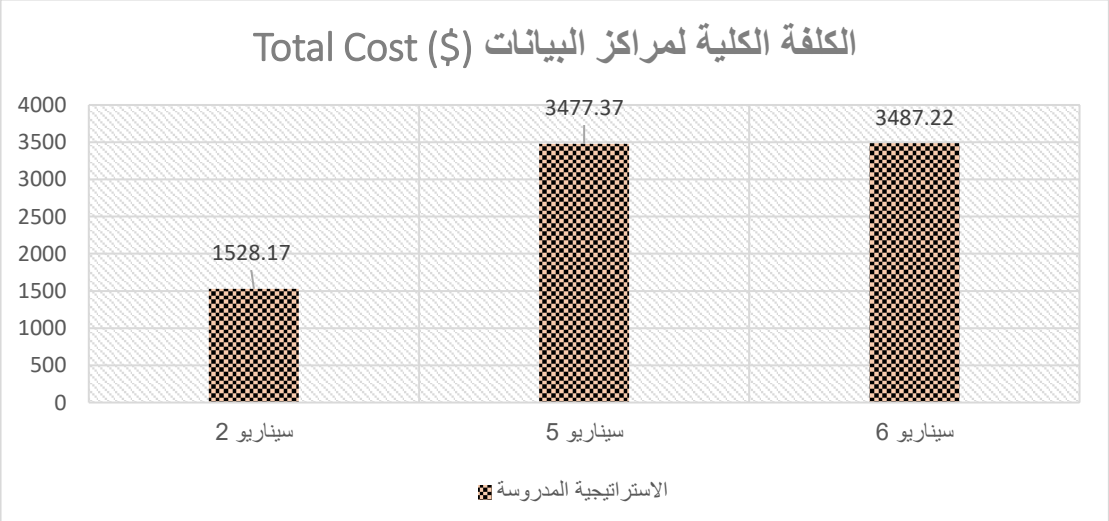


الشكل (13): السيناريو السادس.

نجد بمقارنة متوسط زمن الاستجابة ومتوسط زمن المعالجة والكلفة لكل من السيناريو الثاني والخامس والسادس كما في الأشكال (14) و (15) و (16)، نجد أن السيناريو الثاني (الذي يتضمن ست مراكز بيانات موزعة في جميع المناطق وثلاث وحدات مستخدم موزعة كذلك) أعطى متوسط زمن استجابة وزمن معالجة وكلفة أفضل من السيناريو الخامس (الذي يتضمن ست مراكز بيانات موزعة في جميع المناطق وثلاث وحدات مستخدم موجودة جميعها في منطقة واحدة)، والسيناريو السادس (الذي يتضمن مراكز مجمعة في نفس المنطقة مع وحدات المستخدم)، وكذلك أعطى السيناريو السادس نتائج أفضل مقارنةً مع السيناريو الخامس. الأمر الذي يبين أهمية توضع مراكز البيانات وتوزعها وعددها بالنسبة إلى أماكن وجود المستخدمين وتأثير ذلك على الاستجابة لطلباتهم.







الشكل (14): مقارنة متوسط زمن الاستجابة.

الشكل (15): مقارنة متوسط زمن المعالجة.

الشكل (16): مقارنة الكلفة الكلية لمراكز البيانات.

4-الاستنتاجات والتوصيات

يمكن تلخيص مجموعة من الاستنتاجات بعد دراسة الإستراتيجية السابقة بالنقاط الآتية:

- تقوم هذه الإستراتيجية باختيار مركز البيانات الأفضل لاستقبال طلبات المستخدم ومعالجتها بالاعتماد على مجموعة من البارامترات (تأخير الشبكة، عرض الحزمة المتاحة، حجم طلب المستخدم، سرعة الآلات الافتراضية (المخدمات) ضمن المراكز).
- نتج تحسين في زمن الاستجابة وزمن المعالجة والكلفة الكلية وذلك باختبار الإستراتيجية ومقارنتها مع إستراتيجية الأداء المحسن الموجودة بشكل أساسي ضمن محاكي CloudAnalyst.
- تحقق هذه الإستراتيجية إمكانية التعامل مع أحجام طلبات مستخدم متغيرة، من دون التأثير على التحسين الحاصل في زمن الاستجابة وزمن المعالجة.
- يساعد اختيار الآلة الافتراضية الأسرع غير المشغولة في معالجة طلب المستخدم بوقت أقصر وبالتالي تحسين زمن المعالجة والاستجابة وكذلك تقليل استهلاك الطاقة ضمن مركز البيانات.
- نتج ارتباط قوي في نتائج المحاكاة بين قيم زمن الاستجابة وزمن المعالجة، الأمر الذي يعطي إمكانية التنبؤ بقيمة زمن الاستجابة انطلاقاً من زمن معالجة معطى؛ وبالتالي التنبؤ بأي تحسين ممكن إجراؤه لاحقاً في حال العمل ضمن بنية مراكز البيانات لتقليل الكلفة واستهلاك الطاقة وزمن المعالجة.
- يؤثر توزع مراكز البيانات وعددها في كل منطقة على الاستجابة لطلبات المستخدمين؛ وبالتالي يمكن العمل على دراسة آلية لتشبيك هذه المراكز مع بعضها بطريقة تحقق أفضل استجابة ممكنة.
- هناك حاجة لتقليل الكلفة الكلية الناتجة عن اتباع هذه الإستراتيجية بشكل أكبر مما ظهر بنتيجة مقارنتها مع إستراتيجية الأداء المحسن، بحيث لم يكن التحسين كبيراً كما هو بالنسبة إلى زمن الاستجابة وزمن المعالجة.
- تبدو الحاجة واضحة لتطوير حلول فعالة من ناحية الطاقة لتقنيات الاتصال والبنية التحتية لمراكز البيانات، بحيث يبين تحليل توزع حركة البيانات ضمن السحابة أن أغلبية الحركة تنتقل بين المراكز، وعلى اعتبار أن هناك حاجة للسرعة والسعة العالية فإن تقنية الاتصال الأكثر ملاءمةً لمراكز البيانات السحابية هي الضوئية [18].

5- المراجع

[1] Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., et al, 2010. "***A view of cloud computing***". Commun. ACM 53 (4), 50–58.

[2] Srinivas Jagirdar, et al. 2013, "***Cloud Computing Basics***", https://www.researchgate.net/publication/255994786.

[3] Kremer, J., 2013. "***Virtualization and Cloud Computing, Steps in the Evolution from Virtualization to Private Cloud Infrastructure as a Service***", White paper on virtualization, USA.

[4] Shaista Habib, Fawaz S. Bokhari and Samee U. Khan. 2015, "***Routing Techniques in Data Center Networks***", Springer Science+Business Media New York.

[5] Sharma, V., Rathi, R., Bola, S.K., 2013. "***Round-robin data center selection in single region for service proximity service broker in CloudAnalyst***". Int. J. Comput. Technol. 4 (2a1), 254–260.

[6] Zhang, Q., Cheng, L., Boutaba, R., 2010. "***Cloud computing: state-of the-art and research challenges***". J. Internet Serv. Appl. 1 (1), 7–18.

[7] Rekha, P.M., Dakshayini, M., 2014. "***Cost based data center selection policy for large scale networks***". In: Paper presented at the Computation of Power, Energy, Information and Communication (ICCPEIC), 2014 International Conference on 16–17 April 2014.

[8] Calheiros, R.N., Ranjan, R., Beloglazov, A., De Rose, C.A.F., Buyya, R., 2011. "***CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms***". Software: Pract. Exp. 41 (1), 23–50.

[9] Hetal V. P., Ritesh P., January 2015, "***Cloud Analyst: An Insight of Service Broker Policy***", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 1.

[10] Seyed R., Seyed K., Abolfazl G., 2014, "***Comparison Of CloudSim, CloudAnalyst And CloudReports Simulator in Cloud Computing***", International journal of Computer Science & Network Solutions, Volume 2.No5.

[11] Jogesh K. M., Malathi V., ***"Data Center Networking",*** IEEE ANTS 2012 Tutorial.

[12] Ahmad M. et al, 2015, "***A Variable Service Broker Routing Policy for data Center selection in cloud analyst***".

[13] Zhang, Q., Cheng, L., Boutaba, R., 2010. "***Cloud computing: state-of the- art and research challenges***". J. Internet Serv. Appl. 1 (1), 7–18.

[14] Suakanto, S., Supangkat, S.H., Saragih, R., 2012. "***Performance Measurement of Cloud Computing Services***". ArXiv preprint arXiv: 1205.1622.

[15] Rekha P M, Dakshayini M. 2018, "***Dynamic Cost-Load Aware Service Broker Load Balancing in Virtualization Environment",*** ICCIDS.

[16] Iosup, A. et al., 2011. ***"Performance analysis of cloud computing services for many-tasks scientific computing"***. IEEE Trans. Parallel Distrib. Syst. 22 (6), 931–945.

[17] Mahdee J., 2018, ***"A Predictive Workload Balancing Algorithm in Cloud Services"***, ATHABASCA UNIVERSITY.

[18] Dzmitry K., Pascal B., et al., 2000, "***Energy Consumption Optimization in Cloud Data Centers***".

[19] Sunny Nandwani, 2Mohit Achhra, et al., 2015, ***"Analysis of Service Broker and Load Balancing in Cloud Computing"***, Veermata Jijabai Technological Institute, Mumbai.

[20] Rekha, P.M., Dakshayini, M., 2014. *"Cost based data center selection policy for large scale networks"*. In: Paper presented at ICCPEIC, 2014 International Conference on 16–17 April 2014.

[21] "*Correlation & Simple Linear Regression*", Chapter 4, 2015.