تقنية ذات مستويين لتحمل الأخطاء في تطبيقات الحوسبة عالية الأداء

عائشة محمد عسيري

المشرفة : الدكتورة .مي فاضل

الملخص

تعتبر الموثوقية أكبر مصدر للقلق الذي تواجهه أنظمة الحوسبة العالية الأداء المستقبلية. ضمن نطاق الجيل الحالي لأنظمة الحوسبة العالية الاداء تشير التوقعات إلى حدوث أخطاء بمعدلات عالية جدًا في الأنظمة المستقبلية. مثل هذه الأخطاء الناتجة عن الأعطال يمكن أن تنتج العديد من أنواع الفشل، والتي قد تؤدي إلى نتائج تتراوح من فساد النتائج إلى تعرض التطبيق إلى انهيار كارثي.

في الوقت الذي نقترب فيه من الاكساسكيل ، سيصبح تحدي الصمود حرجًا بسبب الزيادة في حجم النظام.

وبالتالي ، فمن الأساسي أن نتصدى لهذه الأخطاء التي يمكن أن تكون مصدر قلق على أداء التطبيقات العلمية التي تعمل على هذه الأنظمة المستقبلية ، من أجل تقليل تأثيرها السلبي على النظام ككل.

وهذا يتطلب إدارة العديد من تقنيات الأجهزة والبرامج التي تكون قادرة على تحمل الخطأ ومواجهة انخفاض متوسط وقت الفشل .لذلك يجب معالجة هذه المخاوف من خلال تصميم آلية فعالة لتحسين القدرة على الصمود من خلال تحمل الخطأ ، والذي يمكن تحقيقه من خلال كشف الأخطاء.

تكتسب وحدات معالجة الرسومات (GPU) استخدامًا واسعًا في الحوسبة عالية الأداء نظرًا لمزايا أدائها بالنسبة إلى وحدات المعالجة المركزية (CPU). من المتوقع أن تكون وحدات معالجة الرسوميات جزءً من المسار المتوقع نحو الأنظمة المستقبلية ، نظرًا لمزيجها الفريد من حيث كفاءة الطاقة والأداء والكثافة والتكلفة ، والتي أثبتت أنها مسرعات فعالة في انظمة الحوسبة عالية الاداء.

في هذه الرسالة ، نقترح تقنية ذات مستويين للكشف عن الأخطاء وتصنيفها لوحدات معالجة الرسومات المعتمدة على CUDA. تعتمد هذا التقنية على مفهوم تقنيات تحمّل الأخطاء في البرمجيات الذي يعتمد على التنوع في التصميم الذي يوفر اكتشاف الأخطاء للتطبيقات المتوازية ، من خلال إنشاء نسخ متعددة لتحمل الأخطاء والتحقق من صحة نتائج هذه النسخ بواسطة آلية القرار، والكشف عن سبب الخطأ سواء كان بسبب البرمجيات أو الأجهزة .

A two-level fault-tolerance technique for high performance computing applications

Aishah Mohammad Aseeri

Supervised By

Dr. Mai Fadel

ABSTRACT

Reliability is the biggest concern facing future extreme-scale, high-performance computing (HPC) systems. Within the current generation of HPC systems, projections suggest that errors will occur with very high rates in future systems. Such errors resulting from faults can generate numerous kinds of failures, which may lead to the result in outcomes ranging from result corruptions to expose the application to catastrophic crash.

As we approach exascale, resilience challenges will become critical due to increases in system-scale. It is thus fundamental that we address the errors that can be a source concern on the performance of scientific applications running on these future systems ,in order to reduce their negative impact on the overall system. This requires management of various hardware and software techniques that are capable of fault tolerance in the face of a decreasing Mean Time To Failure (MTTF).

Graphics Processing Units (GPUs) are increasingly common in high-performance computing because of their performance advantages relative to CPUs. GPUs are anticipated to be a part of the projected path toward future systems, due to their unique combination of energy-efficiency, performance, density, cost, also proved to be effective accelerators in HPC systems.

In this thesis, we propose a two-level fault-tolerance method for the detection and classification of errors for CUDA-based GPUs. In the first level, it detects the existence of errors by using software redundancy that applies design diversity. In the second level, it investigates the problematic software version and re-executes it on a different hardware component to classify whether the error is a permanent hardware error or a software error. We implemented our method on General-Purpose Graphics Processing Units (GPGPUs) and conducted proof of concept experiments by running three versions of matrix multiplications with different error scenarios and results show the feasibility of the proposed method.