

Modeling of Two-Level Checkpointing With Silent and Fail-Stop Errors in Grid Computing Systems

Rahaf Maher Ghazal, Damascus University, Syria

Samir Jafar, Damascus University, Syria

Mohammad Mounaf Alhamad, Damascus University, Syria

ABSTRACT

With the increase in high-performance computing platform size, it makes the system reliability more challenging, and system mean time between failures (MTBF) may be too short to supply a total fault-free run. Thereby, to achieve greater benefit from these systems, the applications must include fault tolerance mechanisms to satisfy the required reliability. This manuscript focuses on grid computing platform that exposes to two types of threats: crash and silent data corruption faults, which cause the application's failure. This manuscript also addresses the problem of modeling resource availability and aims to minimize the overhead of checkpoint/recovery-fault tolerance techniques. Modeling resources faults has commonly been addressed with exponential distribution, but that isn't fully realistic for the transient errors, which appear randomly. In the manuscript, the authors use Weibull distribution to express these random faults to create the optimal time to save checkpoints.

KEYWORDS

Checkpoint/Recovery, Fault Tolerance, Grid Computing, Modeling of Availability, Weibull Distribution

INTRODUCTION

A Grid is a collection of distributed computing resources which work together and connect by internet to resolve the large parallel computations (Nandhini, 2013; Foster & Kesselman & Nick & Tuecke 2008) however, all that widening and the huge number of the nodes involved has created great challenges to achieve reliability in this large environment (Dabrowski, 2009; Avizienis & Jean-Claude & Brian, 2001; Jafar & J-L., 2004). The probability of the failure is much greater than in traditional distribution platforms (such as which uses local communication network), and the failure of resources leads to stop or deviation of service (Avizienis & Jean-Claude & Brian, 2001). It is, therefore, necessary to investigate the parallel application of fault tolerant techniques.

Fault tolerance is studied to maintain the delivery of correct service in the existence of active faults (Kaur & Singh, 2014). It is generally implemented by error detection and later application recovery. Grid system is exposed to different types of faults, that make the Grid unable to provide

DOI: 10.4018/IJGHP.2021010104

its services properly, including permanent faults and transient faults (Gramacho, 2009; Haider & Babar, 2016). A permanent fault will remain in the system until replacement or repair of the disabled component and then recovery the app. On the other hand, a transient fault appears, disappears and will maybe never occurs again in the overall of system lifetime (Gramacho, 2009; Kaur & Singh, 2014; Menychtas & Kleopatra, 2012). The major difference between the two types of faults is that the permanent faults are determined immediately when it occurs, and system will rollback to a correct previous state, while the transient faults can occur and strike the data inside the application without detection. So transient fault could propagates from one system's component to another, and it remains so until it is activated (once the resource with the error is used by the system to do some computations) (Menychtas & Kleopatra, 2012). Transient fault affect internal processor components, particularly in modern processors that have auxiliary memories and internal buses.

In our research we will addresses to two special types of faults that attack the Grid computing: crash fault (as an example for permanent faults), that shows a constant behavior of system during the work, either failure or work at for a moment (Neocleous & Marios & Paraskevi & Evangelos, 2006).

Silent Data Corruption(SDC) (as an example for transient faults) the most dangerous outcome of a transient faults (Gramacho, 2009), won't stop any application execution and might only be noticed by the users if the final result generated by the application differs from the expected result.

In our manuscript, we propose that the communications layer is to be protected from those faults, which means that these faults have an effect on the app layer.

To the best of our knowledge, Checkpointing/Recovery(C/R) is the most popular fault tolerance technique (Haider & Babar,2016; Jafar & J-L., 2004). However, in the recent years, researchers have developed this technique to cover all system levels, and used multi-levels of checkpointing (Benoit & Cavelan & Robert & Sun,2018; Di & Bouguerra & Bautista-Gomez & Cappello, 2014; Mohror & Moody & Bronevetsky & Supinski,2013). Each level of checkpointing deal with special fault, where fault rate decreases when we go to higher levels. C/R create a save of a previous correct state in the stable memory at the free-fault period, it recovers from save in case of a system failure (Jafar & J-L., 2004; Kaur & Singh, 2014). This definition is usually used for deal the crash fault. To deal with SDC fault we will study in-memory checkpoints coupled with verification mechanism to detect that error (Benoit & Cavelan & Robert & Sun,2018; Benoit & Cavelan & Robert & Sun,2016), which means not needed to additional cost for stable storage process. We would explain the difference between two types of storage in environment model paragraph.

Using the currently available resilience procedures, new challenges created, is the huge overhead for checkpoint and verification techniques on free-fault execution, thereby it took a long time to finish execution of an application. To avoid such scenarios and to get an optimal execution time for parallel applications in Grid computing, we must introduce an accurate analytical studying for resilience technique behavior. We deal with data flow applications which consist of tasks executed in parallel and other tasks may implement sequentially and forming the critical path (Jafar & Krings & Gautier, 2008). In this paper, we are going to introduce a mathematical model for the parallel applications execution time in presences of two studied faults and valid checkpoints/recovery techniques for fault tolerance coupled with verification, we have been able determine an optimal time to perform save the two types of checkpoints, in order to minimize the total execution time of parallel application.

RELATED WORKS

In the past decades, there is a lot of researches focusing on the fault-tolerant problem, which play an important role to maintain high reliability of the systems. But in the earlier studies, there was a tendency to analyze the performance of these techniques, to get the most benefit from them and to reduce their associated overhead (Ahmed & Hasan & Tahar, 2016; Benoit & Cavelan & Robert & Sun, 2016; Benoit & Cavelan & Ciorba & Le Fèvre, & Robert,2019; Benoit & Raina & Robert, 2014; Liu & Nassar & Leangsuksun & Naksinehaboon & Paun & Scott,2008; Munir & Antoon & Gordon-

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/modeling-of-two-level-checkpointing-with-silent-and-fail-stop-errors-in-grid-computing-systems/266219

Related Content

Assembling of Parallel Programs for Large Scale Numerical Modeling

V.E. Malyskin (2010). *Handbook of Research on Scalable Computing Technologies* (pp. 295-311).

www.irma-international.org/chapter/assembling-parallel-programs-large-scale/36413

Potential of General Purpose Graphic Processing Unit for Energy Management System

Jean-Charles Tournier, Vaibhav Donde, Zhao Liand Martin Naef (2012). *International Journal of Distributed Systems and Technologies* (pp. 72-82).

www.irma-international.org/article/potential-general-purpose-graphic-processing/66058

IoT and Big Data Technologies for Monitoring and Processing Real-Time Healthcare Data

Abdelhak Kharbouch, Youssef Naitmalek, Hamza Elkhouchi, Mohamed Bakhouya, Vincenzo De Florio, Moulay Driss El Ouadghiri, Steven Latreand Chris Blondia (2019). *International Journal of Distributed Systems and Technologies* (pp. 17-30).

www.irma-international.org/article/iot-and-big-data-technologies-for-monitoring-and-processing-real-time-healthcare-data/240251

Multi-Route Plan for Reliable Services in Fog-Based Healthcare Monitoring Systems

Nour El Imane Zeghib, Ali A. Alwan, Abedallah Zaid Abualkishikand Yonis Gulzar (2022). *International Journal of Grid and High Performance Computing* (pp. 1-20).

www.irma-international.org/article/multi-route-plan-for-reliable-services-in-fog-based-healthcare-monitoring-systems/304908

Advanced Topics GPU Programming and CUDA Architecture

Mainak Adhikariand Sukhendu Kar (2016). *Emerging Research Surrounding Power Consumption and Performance Issues in Utility Computing* (pp. 175-203).

www.irma-international.org/chapter/advanced-topics-gpu-programming-and-cuda-architecture/139843