早稲田大学大学院国際情報通信研究科

博　士　論　文　概　要

論　文　題　目

Checkpoint Systems in P2P-Grid Environment
P2P グリッド環境におけるチェックポイントシステム

申　請　者

王	歓
Huan	Wang

国際情報通信学専攻
分散システムII

2016 年 02 月

Abstract

P2P-Grid computing explores the convergence of Grid and P2P technologies. Combining Grid and P2P technologies has the potential to lead to the development of large-scale computational Grids based on edge resources. Deploying P2P-Grid architecture on multiple computers enables an organization to automatically exchange computing time with other Internet-connected organizations. Such P2P exchanges of computing time enable individual Peers, i.e. organizations, to transparently aggregate large amounts of computational power with minimal infrastructure requirements or administrative cost.

Reliability of tasks executing becomes a challenge with collective churn effect and increasing heterogeneity and complexity of P2P-Grids environment. A specificity of P2P-Grids is that each node can depart at any time from the computational force of worker nodes supplied to other peers, will cause bursts of execution preemption. The computation results of leaving nodes are no longer accessible, and their computation processes are cut off when the nodes leave their systems. This highly dynamic nature significantly affects the robustness of P2P-Grid systems. It holds the potential to increase the distributed application execution time so that system performance is retarded.

In this dissertation, redundancy to reconfigure the P2P-Grid system and to recover from error states as soon as possible is our target. First, mechanisms to detect and localize failures must be provided. Failure detection design in P2P-Grid systems involves several issues, including consideration and complexity of system dynamics, as well as reasonable storage and time requirements. An interesting issue is rapid response to any node failure or departure. Ideally, node departure or node crash should be immediately detected.

Therefore, perfect failure detection reqiures detecting a node failure immediately during task execution. The details of each chapter will be introduced as follow.

In Chapter 1, I shortly introduced a overview of combination of grid computing and P2P system to provide context for our work. P2P networks and applications may be conceptually as old as the Internet itself but they have become well-known with the widespread use of popular sharing applications. First, I discussed the background of grid computing system and analyzed the advantages of integral on P2P and grid system. Then, failure problem of node arbitrary departure in P2P-Grid systems is discussed. This highly dynamic nature significantly affects the robustness of P2P-Grid systems. Finally, we present our approaches to integrate fault-tolerance techniques into P2P-Grid systems, including a checkpointing system that enables a P2P-Grid system to continue the task execution operation in the presence of failures, and failure probe and recovery are two basic consecutive operation.

Chapter 2 discusses the effect of checkpointing interval on the total exectution time of tasks. For P2P-Grid tasks that perform periodic checkpointing, the choice of the checkpointing interval, the period between checkpoints, can have a significant impact on the total execution time. Seeking the optimal checkpointing interval that minimizes the total task execution time has been a hot topic of research over the last decade. We present an accurate analytical cost function based on our checkpointing architecture. An optimal checkpointing interval calculation method is studied for minimizing the total task execution time.

In Chapter 3, we propose a new fast restartable P2P-Grid checkpointing system. each role including task exectuion, checkpointing etc. in P2P-Grid system are studied. When one task is processed in P2P-Grid over one machine, one is interested how to build a checkpointing system model so as to reduce the execution loss as much as possible and restart quickly during checkpointing interval. Within this architecture a suite of criteria to select each role are presented. Finally a simple simulator based on PeerSim is implemented to validate the improvement in failure recovery performance by our model.

In Chapter 4, to show the real effect of our checkpointing system, we are exploring two methods on the design and implementation of our set of application checkpointing technique whose goal is to checkpoint the application data through hypervisor on the virtualized environments without modification of guest operating system based on Xen. We present evaluation results using purely application-level checkpoint and restart implementation based on Xen to demonstrate the effectiveness of our approach compared to VM system-level checkpointing.

In Chapter 5, the design and implementation of failure probe algorithms in our model is studied. The average failure probe time (AFPT) is used as a metric to evaluate failure probe performance. To capture real failure condition, we implemented our failure probe algorithms on a real P2P-Grid system based on Xen virtual environment. Then we compared three failure probe algorithms on a real working checkpointing system by average failure probe time (AFPT).

In Chapter 6, I concluded the dissertation and stated the future work.