

АНОТАЦІЯ

Дослідження процесів завантаження обчислювального кластеру системи в пакеті GridSim // Дипломна робота ОКР “Магістр” // Фельдман (Мандзій) Світлана Михайлівна // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп’ютерно-інформаційних систем і програмної інженерії, кафедра комп’ютерних наук, група СНм-51 // Тернопіль, 2014 // С. , рис. – , табл. – , кресл. – , додат. – , бібліогр. – .

Ключові слова: КЛАСТЕР, МІНІМАЛЬНЕ ПОКРИТТЯ, МЕТОДИ ПЛАНУВАННЯ, МЕТРИКА, МОДУЛЮВАННЯ, РЕСУРСИ, РОЗПОДІЛЕНІ ОБЧИСЛЕННЯ, GRID, GRIDSIM

В першому розділі проаналізовано основні аспекти організації GRID-систем. В першу чергу розглянуто принципи функціонування та стандарти побудови архітектури таких систем. Основним завданням GRID є забезпечення доступу до ресурсів, а оскільки ресурси розподілені, то функціонування забезпечується спеціальним програмним забезпеченням – службами. На відміну від моделі «клієнт-сервер», набір служб встановлюється на кожному ресурсі. Проведений аналіз завдань планування розподілу ресурсів в GRID-системах.

В другому розділі розглянуто архітектуру та принципи моделювання в пакеті GridSim. За допомогою GridSim можна проводити відтворені експерименти, які були б неможливі в цьому оточенні для динамічних GRID-систем. Основні можливості: моделювання різних характеристик ресурсів GRID; моделювання різних політик планування завдань на обчислювальному кластері, як вже реалізованих (FCFS, EasyBackfill, Earliest Suitable Gap), так і використання розроблених користувачем алгоритмів планування; використання даних про завантаження реальних суперкомп’ютерів для проведення експериментів; підтримка механізму аукціону для планування завдань; моделювання різних конфігурацій обчислювальної мережі GRID системи з різних топологій; моделювання регіональних компонентів GRID інформаційних сервісів моделювання DataGrid.

Розроблено метод вирішення задачі про мінімальне покриття для планування завантаження кластера.

Розроблено модифікацію існуючих алгоритмів планування завантаження кластера, а також алгоритм планування, який заснований на знаходженні мінімального покриття.

В третьому розділі проведено експериментальні дослідження в розрізі завантаження ресурсів (кластерів).

Мета: розробка методів планування завантаження обчислювальних ресурсів GRID-системи для їх моделювання в пакеті GridSim.

Об'єкт дослідження: процеси планування завантаження ресурсів GRID-системи.

Предмет дослідження – методи планування завантаження ресурсів GRID-системи для їх моделювання в пакеті GridSim.

Методи дослідження: аналіз, дискретно-подієве моделювання, методи статистичного аналізу.

Для розробки моделі планувальника використана мова високого рівня Java та пакет GridSim.

Основні результати: проаналізовано існуючі методи планування завантаження кластерів GRID. Проведений аналіз архітектури та принципів моделювання в пакеті GridSim. Розроблено та інтегровано до пакету GridSim програмну реалізацію модифікацій методу FCFS та методу на основі вирішення задачі про мінімальне покриття. Здійснена оцінка ефективності отриманих результатів за різні періоди часу на реальних даних роботи GRID на основі запропонованих метрик. Програмний комплекс реалізує методи планування завантаження кластерів, реалізована розширенням пакету GridSim на мові Java.

ABSTRACT

Investigation of downloads computational cluster system package GridSim // Bachelor Thesis "Master" // Feldman (Mandzij) Svitlana // Ternopil Ivan Pul'uj National Technical University, Department of Computer Information Systems and Software Engineering, Department of Computer Science, a group of SNM-51 // Ternopil, 2014 // C., Fig. -, Table. -, Draw. - Add. -, Ref. -.

The first section analyzes the main aspects of GRID-systems. First of all Principles of operation and construction standards architecture of such systems. GRID is the main objective of providing access to resources, and as resources are distributed, then the operation is provided by special software - services. Unlike the model of "client-server", a set of services installed on each resource. The analysis task planning resource allocation in GRID-systems.

The second section examines the architecture and principles simulation package GridSim. With the help of GridSim can be made reproducible experiments that would be impossible in this environment for dynamic GRID-systems. Key features: simulation of different characteristics of resources GRID; modeling of various scheduling policies on a computing cluster, as already implemented (FCFS, EasyBackfill, Earliest Suitable Gap), and the use of algorithms developed by planning; Data about downloading real supercomputers to conduct experiments; support mechanism for scheduling the auction; simulation of different configurations GRID computing network system with different topologies; regional modeling components GRID information services modeling DataGrid.

The method of solving the problem of the minimum coverage for planning boot cluster.

A modification of the existing algorithms for scheduling loading cluster, as well as scheduling algorithm, which is based on finding the minimum coverage.

In the third section, experimental studies conducted in the context of loading resources (clusters).

Objective: To develop planning methods download computing resources GRID-systems for modeling package GridSim.

Implementation of the objectives of the study necessitated the definition and the following tasks:

1) to analyze the principles of operation and construction standards architecture GRID-systems;

2) analyze the existing planning methods download GRID-computing resources system;

3) justify the choice of development environment for modeling work scheduler GRID-system;

4) develop a modification of the existing planning methods download GRID-computing resources systems used in the package GridSim;

5) develop metrics to assess the effectiveness of the planning methods download GRID-computing resources system;

6) develop a software implementation planners downloading GRID-computing resources system and integrate it into the package GridSim;

7) to study and analyze the results on the effectiveness of the proposed methods of computational resources planning download GRID-system on the data of the French GRID-system AuverGrid package GridSim.

Object of study: the planning download resources GRID-system.

Purpose of the study - methods of planning GRIDsystem download resources for their modeling package GridSim.

Methods: analysis, discrete Event simulation and statistical analysis.

Main results: The existing planning methods download cluster GRID. The analysis of the architecture and principles simulation package GridSim. Developed and integrated into GridSim package software implementation of modifications FCFS method and the method based on the solution of the problem of minimum coverage. The estimation of the effectiveness of the results obtained for different periods of time on real data of GRID based on the proposed metrics. The software implements a set of methods for planning boot cluster expansion pack GridSim is implemented in Java.

Thus, the results of the simulation are as follows:

1 Experimental studies in terms of average download resources (clusters) for the entire simulation period for each mode queuing (scheduling algorithm) showed that increasing the number of computing resources GRID-system methods for FCFS and Easy Backfill increases uneven distribution of loading between computing resources - these methods load more resources that are closer to the top of the list of all resources and less - those who are near the end of the list; on the other hand - modifying «FCFS + Ring», «FCFS + RoundRobin» and planning method based on finding the minimum coverage, increasing the number of computing resources GRID-systems remain uniform distribution of loading resources.

2 Experimental studies in the context cluster factor loading for each mode of service for the entire period of simulation (scheduling algorithm) showed that increasing the number of computing resources GRID-system methods for FCFS and Easy Backfill reduced rate cluster download, indicating that the low performance of GRID-systems using techniques FCFS and Easy Backfill; on the other hand - modifying «FCFS + Ring», «FCFS + RoundRobin», method Earliest Suitable Gap and planning method based on finding the minimum coverage, increasing the number of computing resources GRID-systems remain coefficient cluster download about 0.9, indicating that high level loading balancing between resources.

3 Experimental research in terms of average download resources (clusters) and cut cluster factor loading for each mode of service per week simulation has shown that the methods of FCFS and Easy Backfill not adapt to changes in the input stream tasks, on the other hand - modified methods «FCFS + Ring »,« FCFS + RoundRobin »and planning method based on finding the minimum coverage adapt to dynamic changes in the input stream and do not lose their effectiveness.

Keywords: CLUSTER, DISTRIBUTED COMPUTING, GRID, GRIDSIM, METRIC, MINIMAL COVER, RESOURCES, SHEDULLING METHODS.