

УДК 681.324

Е.Н. Машенко, доцент, канд. техн. наук,**А.А. Лисицкая, студентка,****Н.А. Плюхина, студентка***Севастопольский национальный технический университет**ул. Университетская 33, г. Севастополь, Украина, 99053**E-mail: kvf@sevgtu.sebastopol.ua***ОПТИМИЗАЦИЯ ПРОЦЕССОВ ПЛАНИРОВАНИЯ ЗАДАНИЙ В ГРИД-СИСТЕМАХ С ИСПОЛЬЗОВАНИЕМ КОМПЛЕКСА ИМИТАЦИОННЫХ МОДЕЛЕЙ***Рассмотрена задача оптимизации процессов планирования заданий в Грид-системах.**Описан оптимизационный эксперимент на основе комплекса имитационных моделей.***Ключевые слова:** *Грид-система, планирование заданий, имитационное моделирование.*

Грид (англ. *grid* – решетка, сеть) – согласованная, открытая и стандартизованная компьютерная среда, которая обеспечивает гибкое, безопасное, скоординированное разделение вычислительных ресурсов и ресурсов хранения информации, в рамках одной виртуальной организации.

Грид-система – система, связанная с функциями интеграции, виртуализации и управления службами и ресурсами в распределенной, гетерогенной среде, которая поддерживает совокупность пользователей и ресурсов (виртуальные организации) на совокупности традиционных административных и организационных доменов (фактических организаций).

Грид-технологии – распределение крупных вычислительных задач с помощью специального программного обеспечения на множество удаленных компьютеров.

Основная цель применения Грид-технологий – интеграция большого количества распределенных ресурсов, которые могут взаимодействовать друг с другом для достижения определенной цели.

В настоящее время эти технологии являются перспективными, так как они образуют компьютерную инфраструктуру нового типа, содержащую набор стандартизированных служб, обеспечивающих надежный, удаленный и дешевый доступ к информационным и вычислительным ресурсам [1]. Таким образом, с появлением Грид-концепции стало возможным решать ресурсоемкие задачи, например, моделирование химических процессов, расчеты в области высоких энергий, нанотехнологий и др.

Основная проблема Грид-вычислений – принятие решения о скоординированном разделении динамических ресурсов и планировании множества заданий в условиях большого количества виртуальных организаций. Следовательно, исследования и оптимизация процессов планирования заданий в Грид являются актуальными.

Моделирование работы Грид-систем позволяет на этапе проектирования разработать оптимальную структуру распределенной системы с учетом поставленных задач, достигая значительной экономии финансовых ресурсов. Моделирование нагрузки реализованной Грид-системы выявляет узкие места системы, позволяет оценить масштабируемость решаемых задач и провести планирование дальнейшего усовершенствования системы.

Целью данной статьи является построение математической модели процесса планирования заданий в Грид-системах, постановка задачи оптимизации и проведение оптимизационного эксперимента на основе комплекса имитационных моделей.

С точки зрения архитектуры Грид-системы механизм планирования задач реализован на двух уровнях: локальном и глобальном. Локальные задачи выполняются на уровне отдельных кластеров и не поступают на другие подсистемы, глобальные задачи могут попадать на любой кластер, входящий в состав Грид. Общая схема планирования заданий в Грид-системе изображена на рисунке 1.

Используется следующий механизм взаимодействия: планировщик получает задачи от пользователей, выбирает необходимые ресурсы для запуска задачи, используя информационный сервис, и на основе полученной информации направляет задачу на подходящий ресурс [2].

Существует два основных подхода к планированию заданий. Первый заключается в непосредственном запуске задания на оптимальных вычислительных ресурсах после его постановки на счет. Второй подход подразумевает предварительное составление некоторых планов обработки заданий – расписаний. Для нахождения расписания используются точные методы, гарантирующие нахождение расписания, например методы линейного и целочисленного программирования, ограниченного перебора, метод критического пути. Также существует ряд эвристических методов, не гарантирующих лучшее решение, например иерархический подход, имитация отжига, подходы, основанные на теории искусственного интеллекта (экспертные системы, системы, основанные на знаниях – *knowledge-based*), методы, использующие эвристики планирования и упорядочения, эволюционные алгоритмы.

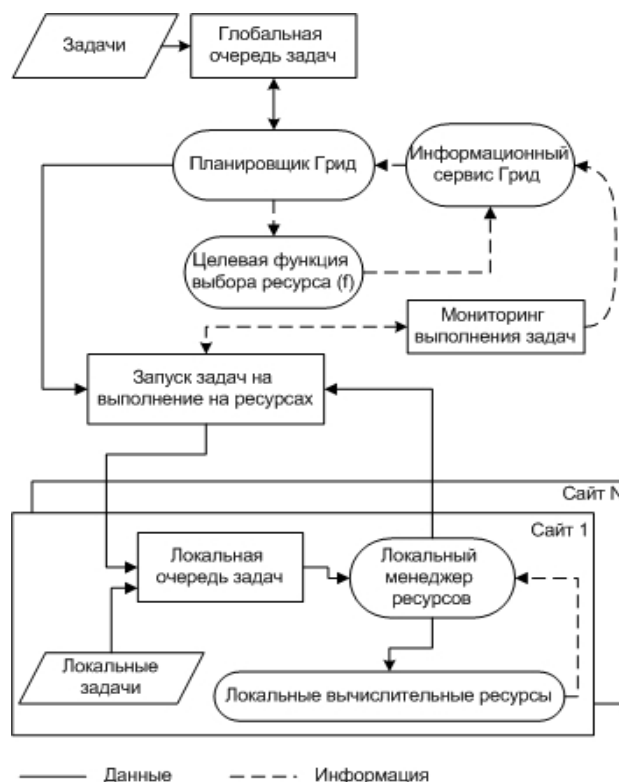


Рисунок 1 – Общая схема планирования заданий в Грид-системе

В данной работе исследуются процессы планирования, реализующие первый подход. В таблице 1 приведена краткая классификация основных методов планирования заданий в Грид-системах. Более подробно особенности методов описаны в [1].

Далее рассматриваются два класса алгоритмов, основанные на очередях:

- 1) алгоритмы планирования с разделением ресурсов (FCFS (First-Come, First-Served), Round Robin, SPN (shortest process next) и SRT (shortest remaining time)) [1];
- 2) алгоритмы распределенного планирования (JobRandom, JobLeastLoaded [2]).

Таблица 1 – Основные методы планирования заданий в Грид

Группы методов	Основные идеи	Особенности		
Методы разделения времени	Разделение времени процессоров между несколькими заданиями	Трудно применимы в условиях Грид, направлены на улучшение среднего времени обработки заданий		
Методы разделения ресурсов	Разделение пространства ресурсов между заданиями	Основанные на интегральных критериях	Оптимизируют загрузку ресурсов или общее время счёта пакета заданий	
		Приоритетного планирования	Обеспечивают возможность управления порядком выделения ресурсов для отдельных заданий	Основанные на очередях (Queuing)
				Основанные на расписаниях (Planning)

В модели, реализующей алгоритмы с разделением ресурсов, задания поступают в общую очередь, разделяемую всеми ресурсами системы. Распределение заданий по очередям к локальному планировщику осуществляется с помощью алгоритма взвешенных очередей (Weighted Queuing) согласно установленным приоритетам. По данным мониторинга текущего состояния ресурсов системы и критериям выбранного алгоритма планирования производится выборка заданий и предоставление им ресурсов в требуемых или установленных алгоритмом объемах.

При реализации алгоритмов распределенного планирования задания, поступающие в Грид-систему, регистрируются в одной из очередей распределенной системы очередей. Каждая очередь обслуживается

собственным локальным планировщиком, принимающим для очередного задания одно из трех решений: задание запланировано для выполнения на ресурсах одного или нескольких узлов Грид; оставлено в очереди для его последующего планирования; передано в другую очередь. При этом необходимо обеспечить возможность независимого одновременного планирования заданий, находящихся в разных очередях, локальными планировщиками и согласованно использовать ресурсы Грид-системы.

Построим модель процесса планирования и выполнения заданий с использованием теории массового обслуживания. Представим исследуемую Грид-систему в виде сети массового обслуживания (СтМО), структура которой показана на рисунке 2.

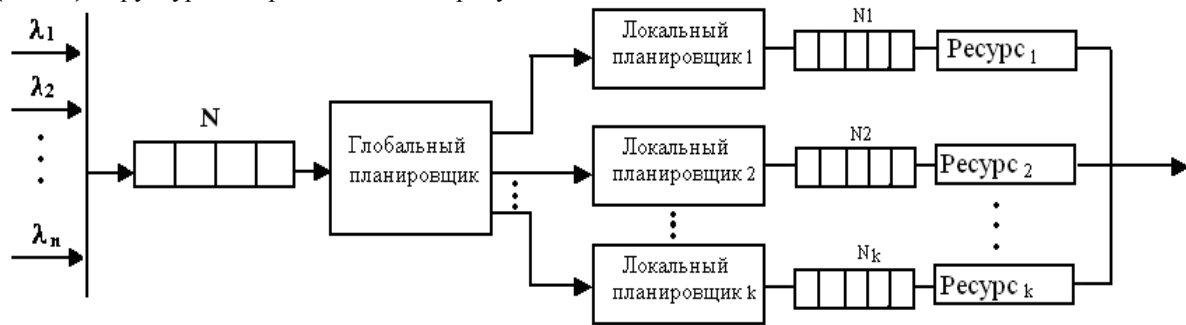


Рисунок 2 – Представление Грид-системы в виде СтМО

Введем следующие параметры модели, отражающие базовые показатели системы и ресурсов: $\lambda_1, \lambda_2, \dots, \lambda_n$ – интенсивности поступления заданий в Грид; T – временной интервал, на протяжении которого велось наблюдение системы; A – количество запросов, поступивших в систему за время T ; C – количество обработанных запросов за время T ; B – временной интервал, на протяжении которого происходила обработка запросов (в пределах интервала T); N – среднее количество запросов в системе; S – структура Грид-системы.

Задача оптимизационных исследований – выяснить, какие из алгоритмов планирования заданий в Грид обеспечивают максимальное число обработанных заданий (1):

$$C_{onm} = \text{Max}\{C_1, C_2, \dots, C_q\}, \quad (1)$$

где q – количество рассматриваемых алгоритмов;

и минимальное время нахождения задания в системе (2):

$$B_{onm} = \text{Min}\{B_1, B_2, \dots, B_q\}. \quad (2)$$

Таким образом, выбранный алгоритм планирования должен обеспечивать оптимум некоторого векторного критерия:

$$K_{onm} = F(C_{onm}(\lambda, A, T, B, N, S), B_{onm}(\lambda, A, T, B, N, S)). \quad (3)$$

Для проведения исследований разработан программный комплекс имитационного моделирования с использованием принципа многоверсионности алгоритмов планирования заданий [3]. Параметры имитационных моделей основаны на статистике функционирования Грид-систем. Разработанные имитационные модели построены в среде имитационного моделирования AnyLogic. Результаты исследований Грид GWA-T-11 приведены в [3]. В данной статье в качестве опорного источника параметров модели Грид-системы был использован отчет о работе GWA-T-4 Trace analysis report (anon_jobs.gwf / GWA-T-4) [4].

Параметры моделей, построенных на основе алгоритмов распределенного планирования и алгоритмов с разделением ресурсов, представлены в таблице 2. Результаты моделирования приведены в таблице 3.

Анализ полученных результатов позволяет сделать следующие выводы:

1) при использовании алгоритмов с разделением ресурсов – при осуществлении невытесняющего планирования суммарное число обработанных заданий максимально при использовании алгоритма FCFS;

2) при вытесняющем планировании лучший результат по количеству обработанных заданий получен при использовании алгоритма Round Robin;

3) при использовании алгоритмов распределенного планирования оптимальные критерии качества обеспечивает алгоритм JobRandom.

Разработанный комплекс имитационных моделей можно использовать при анализе и оптимизации Грид-систем различной структуры.

Таблица 2 – Основные параметры имитационных моделей

Параметр	Значение
Интенсивность поступления заданий в Грид-систему (нормальное распределение)	(17,8–1,12) с ⁻¹
Приоритет (стоимость) заданий	1–3
Номинальный объем ресурса	100%
Требуемый заданием объем ресурса	(15–44)%
Время, необходимое для выполнения задания	(5–29) с
Для алгоритмов Round Robin и SPN квант времени на выполнение задания	15 с
Типы ресурсов CR, SE, RAM (SE – элементы физической памяти, CR – вычислительные ресурсы, RAM – элементы оперативной памяти)	

Таблица 3 – Сводная таблица результатов моделирования

Параметр Алгоритм	Среднее время пребывания задания в системе	Среднеквад- ратичное отклонение	Довери- и- тельная граница	Дисперсия	Суммарное число обработанных заданий	Время модели- рования, с
Алгоритмы с разделением ресурсов						28800
FCFS	24,74	8,57	0,24	73,55	4855	
SPN	25,2	7,78	0,25	77,23	4848	
Round Robin	24,17	15,46	0,44	239,2	4851	
SRT	24,15	15,33	0,43	235,11	4844	
Алгоритмы распределенного планирования						
Job Random	19,28	7,44	0,25	55,4	4851	
JobLeast Loaded	37,34	33,62	0,53	1130,35	4849	

Таким образом, построена математическая модель процесса планирования заданий в Грид-системах с использованием теории массового обслуживания, приведена постановка задачи оптимизации и проведен оптимизационный эксперимент на основе комплекса имитационных моделей. Получены результаты, позволяющие повысить эффективность процесса планирования заданий в Грид-системах.

Перспективы дальнейших исследований заключаются в том, чтобы выработать рекомендации для систем принятия решений. Также предполагается расширить комплекс моделей с использованием других типов алгоритмов планирования, структур Грид-систем и параметров потока задач.

Бibliографический список использованной литературы

1. Корнеев В.В. Распределенный метапланировщик ГРИД / В.В. Корнеев, Д.В. Семенов // Вычислительные методы и программирование. — 2010. — Т. 11. — С. 69–76.
2. Куссуль Н.Н. Grid-системы для задач исследования Земли / Н.Н. Куссуль, А.Ю. Шелестов. — К.: Наукова думка, 2008. — 452 с.
3. Мащенко Е.Н. Система моделирования процессов планирования заданий в Грид с использованием принципа многоверсионности алгоритмов / Е.Н. Мащенко, А.А. Лисицкая, Н.А. Плюхина // Управление, автоматизация и окружающая среда: матер. междунар. науч.-техн. конф. молодых ученых, аспирантов и студентов; Севастополь, 25–28 апреля 2011 г. — Севастополь, 2011. — 254 с. — С. 118–121.
4. The Grid Workloads Archive [Электронный ресурс]. — Электрон. текстовые данные (84205 bytes). — Режим доступа: <http://gwa.ewi.tudelft.nl/pmwiki/pmwiki.php?n=Main.Home>

Поступила в редакцию 15.06.2011 г.

Мащенко О.М., Лисицка А.О., Плюхина Н.А. Оптимізація процесів планування завдань у грід-системах з використанням комплексу імітаційних моделей

Розглянуто задачу оптимізації процесів планування завдань в грід-системах. Описано оптимізаційний експеримент на основі комплексу імітаційних моделей.

Ключові слова: грід-система, планування завдань, імітаційне моделювання.

Maschenko E.N., Lisitskaya A.A., Pluhina N.A. Optimization of the task planning processes in the Grid systems using the complex of simulation models

The problem of optimizing the task planning in the Grid systems is considered. The optimization experiment based on complex simulation models is described.

Keywords: grid system, planning jobs, simulation modeling.