Анализ и оптимизация кластерных систем с использованием аналитических моделей

Михайлова Т.В.
Кафедра ПМиИ ДонНТУ
tanya@ r5.dgtu.donetsk.ua

Abstract

Michailova T. Analysis and optimization of cluster systems with the use of analytical models. *The modified methods of analysis and synthesis of multiprotsessor calculable resources of a different topology by probabilistic models are offered, allowing to analyse and design more wide class of parallel calculable environments.*

Введение

Для анализа и синтеза различных вычислительных структур, в том числе кластерных, используются непрерывные [1,11,12] или дискретные аналитические модели [2]. Дискретные модели Маркова, отражающие работу вычислительной среды более точно, имеют большую размерность и эффективно отображаются на параллельные вычислительные структуры [9]. Менее трудоемкие непрерывные модели применяются для исследования более сложного класса анализируемых структур.

Кластерные вычислительные системы по критерию совместного использования дискового пространства классифицируются следующим образом: с совместным использованием дискового пространства и без предоставления доступа к ресурсам [5].

1. Анализ кластерных систем

На основании методик [3,4] можно построить модели кластеров с совместным использованием дискового пространства и без предоставления доступа к ресурсам. Построим модель кластера топологии NxN (рис.1).

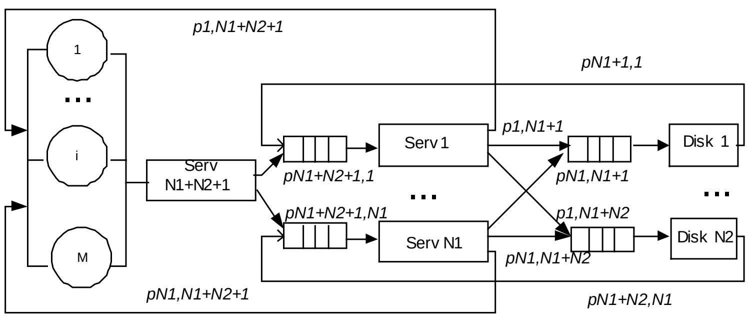

Рисунок 1 – Структура кластера топологии NxN

В ней *N1* кластеров, выполняющих разные приложения, *N2* диска. Дисковое пространство неоднородное. Такие кластеры использует, например, Oracle Parallel Server (для 3- и 4-х узловых кластеров). Эта топология может использоваться с Informix и для организации систем высокой готовности.

Каждая из M задач с вероятностью $p_{N1+N2+1,i}$ $(i=\overline{1,N1})$ посылает запрос к i-му серверу, который, в свою очередь, обрабатывая этот запрос обращаются к j-му диску с вероятностью $p_{i,N1+j}$ $(i=\overline{1,N1},\ j=\overline{1,N2})$.

Функционирование рассматриваемой системы можно представить замкнутой стохастической сетью, содержащей $N1+N2+1$ системы массового обслуживания (СМО), в которой циркулирует M заявок.

По графу передач (рис.2) определяются коэффициенты посещений каждой СМО, решая систему уравнений

$$\begin{cases} \alpha_{N1+N2+1} = 1 + \sum\limits_{i=1}^{N1} p_{i,N1+N2+1} * \alpha_i, \\ \alpha_i = p_{N1+N2+1,i} * \alpha_{N1+N2+1} + \sum\limits_{j=N1+1}^{N1+N2} p_{j,i} * \alpha_j, \qquad i=\overline{1,N1}, \\ \alpha_i = \sum\limits_{j=1}^{N1} p_{j,i} * \alpha_j, \qquad i=\overline{N1+1,N1+N2}. \end{cases}.$$

За состояние модели принимается распределение заявок по всем СМО

$$\overline{m} = (m_{N1+N2+1}, m_1, \ldots, m_{N1}, m_{N1+1}, \ldots, m_{N1+N2}),$$

где $m_{N1+N2+1}+m_1+\ldots+m_{N1}+m_{N1+1}\ldots+m_{N1+N2}=M$.

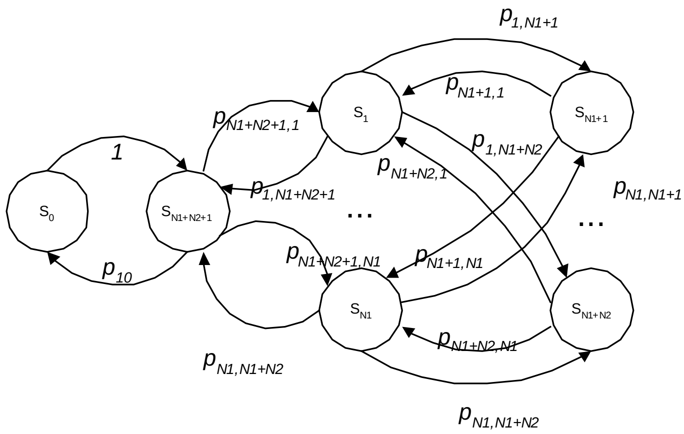

Рисунок 2 – Граф передач кластера топологии NxN

Эта модель может быть использована при анализе топологии кластерных кластерных пар при условии: $N1=N2$ – четное количество (для образования пар); вероятности обращения пар серверов к своим дискам: $p_{i,N1+i}+p_{i,N1+i+1}=1$, $p_{i+1,N1+i}+p_{i+1,N1+i+1}=1$ $(i=1,3,5,\ldots,N1-1)$.

Для исследования топологии кольцо вводятся следующие корректировки: $N1=N2$, $p_{i,N1+i}=1$ $(i=1,2,3,\ldots,N1-1)$, $p_{N1,i}$ $(i=1,2,\ldots,N1)$.

Исследуем эту вычислительную среду при решении задач различных классов.

Вероятности обращения к серверам приведены в табл.1. Суммарные вероятности обращения к дискам - в табл.2. Деление потоков к дискам между собой (в долях) приведено в табл.3.

Таблица 1 Вероятности обращения к серверам

	Номер варианта			
	1	2	3	4
p(5,1)	0.83125	0.59375	0.35625	0.11875
p(5,2)	0.11875	0.35625	0.59375	0.83125

Таблица 2 Суммарные вероятности обращения к дискам

	Номер варианта			
	5	6	7	8
(p(1,3)+p(1,4))	0.125	0.375	0.625	0.875
(p(2,4)+p(2,3))	0.875	0.625	0.375	0.125

Таблица 3 Деление потоков к дискам между собой (в долях)

	Номер варианта			
	9	10	11	12
p(1,3)	0.875	0.125	0.875	0.125
p(1,4)	0.125	0.875	0.125	0.875
p(2,4)	0.875	0.875	0.125	0.125
p(2,3)	0.125	0.125	0.875	0.875

Зафиксируем вероятность обращения к серверу (вариант 1) и для каждого варианта обращения к диску (варианты 5-8) будем изменять потоки между дисками (варианты 9 – 12). Минимальное время решения задачи (рис.3) равно 0.34с для варианта (1,6,10).

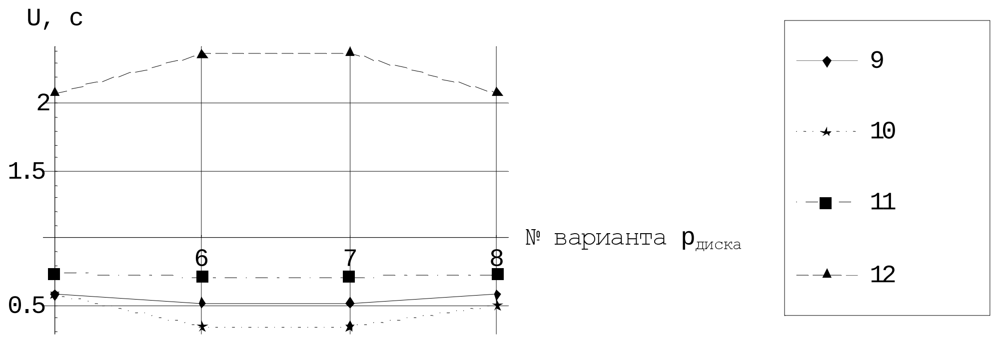

Рисунок 3 – Зависимость времени отклика при различных вероятностях обращения к дискам

В целом, независимо от распределения потока между дисками класс задач, соответствующий варианту (1,10) выполняется быстрее всего (на рис.3- вариант 10). Несколько хуже классы задач, соответствующие вариантам (1,11), (1,9), что можно увидеть на рис.3 – варианты 9, 11. Хуже всего для класса задач(1,12)- на рис.3 – вариант 12. Проанализируем причины такого разброса, построив графики загрузок серверов для лучшего (вариант 10) и худшего вариантов (вариант 12).

На рис.4 приведены загрузки серверов и дисков для варианта 10 (вероятности обращения к серверу). Устройства оптимально загружены в середине, то есть для вариантов 6 и 7, приближенных к равномерному распределению вероятностей обращений к дискам.

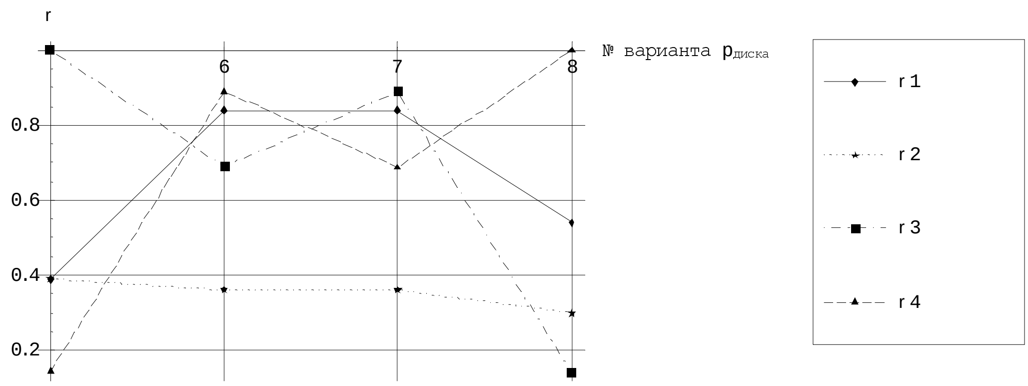

Рисунок 4 – Зависимость загрузки серверов (ρ_1, ρ_2) и загрузки дисков (ρ_3, ρ_4) от вероятности обращения к диску для варианта 10

Используя полученные при моделировании стационарные вероятности, можно вычислить основные характеристики кластерных систем [2,4] для определения эффективности вычислительной среды по раличным критериям.

При использовании критерия минимального времени отклика стоимость кластерной системы несоизмеримо велика. Поэтому в ряде случаев для подбора оптимального коэффициента мультипрограммирования и снижения стоимости можно использовать методику [6], в которой предлагается критерий сбалансированности, составляющие которого цена простоя оборудования и штраф за задержку выполнения запроса.

Аналитические модели используются для выбора вида кластера с определенным классом решаемых на них задач, сравнивая получаемые при моделировании характеристики [7].

Полученные вероятностные модели можно использовать для задач синтеза кластерных структур.

2. Синтез кластерных систем

Для оптимизации состава и структуры вычислительных систем можно использовать методы [7], позволяющие определить структуру вычислительной среды минимальной стоимости при заданном времени отклика или, наоборот, с минимальным временем отклика заданной стоимости.

Предлагается модифицированный алгоритм оптимизации состава и структуры кластерных систем с использованием метода средних.

1.Выбирается исходный вектор производительности устройств V_k ($k=1,\ldots,N$) из условия ограничения. Задаются начальные значения величины шага h, времени отклика $Uopt=\infty$, признак определения лучшей точки $flag$=0 ($flag$=1, если определена лучшая точка).

2.Методом средних считаются основные характеристики: средние времена пребывания в i-ой СМО, время отклика задачи $U(M)$, среднее число задач, находящихся в i-ой СМО.

3.Если $U(M) < U\text{opt}$, запоминаем новый вектор V_k ($k=1,\ldots,N$), новый рекорд по времени отклика $U(M) = U\text{opt}$, $flag$=1 (определена лучшая точка на данном шаге).

4.Если нет лучшей точки с шагом h ($flag$=0), уменьшаем его $h=h/2$.

5.Если $h>\varepsilon$, переходим к п.5, иначе, конец алгоритма.

5.Строим новый вектор V_k ($k=1,\ldots,N$)= V_k ($k=1,\ldots,N$)+h из условия (3). Переходим к п.2.

Начальным приближением для этого метода используется точка, полученная градиентным методом [5]. Затем на границе с единичным шагом выбирается направление координаты, по которой функция улучшается, и вычисляется новая точка. Если нет лучшей точки, шаг уменьшается. Процесс продолжается, пока не будет достигнута заданная точность.

Одна итерация с использованием теоремы о среднем требует $L*N*M$ операций сложения и умножения (где L - константа) и еще столько же операций для уточнения на границе. Сложность одной итерации этого алгоритма - полиномиальная.

Алгоритм с использованием теоремы Джексона имеет комбинаторный порядок, а алгоритмы с использованием теоремы о среднем [7] – полиномиальный, что позволяет решать задачи, которые вообще не решаются аналитическим методом на современных ЭВМ в течение реального времени.

Выводы

Предложенные методики анализа эффективности вычислительных систем позволяют получить характеристики функционирования этих сред при решении на них различных классов задач, позволяющие оценить работу кластера по выбранному критерию. Рассмотренные способы оптимизации состава и структур кластерных систем можно использовать для проектирования и комплектации кластеров.

Использование вероятностных моделей при проектировании, эксплуатации и оптимизации кластерных систем позволяет вырабатывать рекомендации по рациональному использованию ресурсов этой вычислительной среды.

Литература

1. Авен О. И. и др. Оценка качества и оптимизация вычислительных систем. – М.: Наука, 1982, 464с.
2. Клейнрок Л. Вычислительные системы с очередями. – М.:Мир, 1979, 600с. Последовательно - параллельные вычисления: Пер. с англ. - М.: Мир, 1985. - 456 с.
3. Михайлова Т.В. Анализ оценки эффективности кластерных систем с использованием вероятностных моделей //Системний аналіз та інформаційні технології. Тези доповідей учасників Міжнародної науково-практичної конференції студентів, аспирантів та молодих вчених. 1-3 липня 2003р., м.Київ, 83-85с.
4. Основы теории вычислительных систем/С.А.Майоров, Г.И.Новиков, Т.И.Алиев и др.М.: Высшая школа, 1978, 408с.
5. Спортак М., Франк Ч., Паппас Ч. и др. Высокопроизводительные сети. Энциклопедия пользователя.-К.:"ДиаСофт", 1998.-432с.
6. Фельдман Л.П., Михайлова Т.В. Оценка эффективности кластерных систем с использованием моделей Маркова. //Известия ТРТУ. Тематический выпуск: Материалы Всероссийской научно-технической конференции с международным участием «Компьютерные технологии в инженерной и управленческой деятельности». – Таганрог: ТРТУ, 2002. – №2 (25). – С. 50–53.
7. Фельдман Л.П., Михайлова Т.В. Способы оптимизации состава и структуры высокопроизводительных вычислительных систем //Научные труды Донецкого государственного технического университета. Серия «Информатика, кибернетика и вычислительная техника»(ИКВТ-2001).- Донецк: ДонГТУ.- 2000. С. 80-85.
8. Фельдман Л.П., Дедищев В.А. Математическое обеспечение САПР. Моделирование вычислительных и управляющих систем. – Киев: УМК ВО, 1992, 256с.
9. Фельдман Л.П., Михайлова Т.В. Параллельный алгоритм построения дискретной марковской модели /Высокопроизводительные параллельные вычисления на кластерных системах. Материалы четвертого Международного научно-практического семинара и Всероссийской молодежной школы. /Под редакцией член-корреспондента РАН В.А. Сойфера. – Самара, 2004. – С. 249–255.
10. Шнитман В. Современные высокопроизводительные компьютеры. Информа-ционно- аналитические материалы центра информационных технологий, 1996: http://hardware/app_kis
11. Cremonesi P., Gennaro C. Integrated Performance Models for SPMD Applications and MIMD Architectures //IEEE Transactions on Parallel and Distributed Systems, vol. 13, №7, jul.2002, PP.745-757
12 Varki E. Response Time Analysis of Parallel Computer and Storage Systems //IEEE Transactions on Parallel and Distributed Systems, vol. 12, №11, nov.2001, PP.1146-1161