

УДК 681.511

В.В. Кирюхин, доцент, канд. физ.-мат. наук*Севастопольский национальный технический университет**ул. Университетская, 33, г. Севастополь, Украина, 99053**E-mail: kvk@sevtu.sebastopol.ua***АЛГОРИТМЫ ДИНАМИЧЕСКОЙ РЕСТРУКТУРИЗАЦИИ ОДНОРОДНОГО КЛАСТЕРА ВЫСОКОЙ ДОСТУПНОСТИ**

Рассматривается задача оптимального распределения серверов по узлам высокодоступного однородного кластера с целью минимизации средних потерь из-за отказов серверов в условиях изменения нагрузки и параметров кластера. Предлагаются алгоритмы решения соответствующей задачи целочисленного программирования.

Ключевые слова: *облачное вычисление, кластер высокой доступности, динамическая реструктуризация, целочисленное программирование.*

Постановка проблемы. Одним из основных средств реализации современных компьютерных технологий типа грид-компьютинга, «облачных» вычислений и т.п. являются кластеры — группы компьютеров, объединенных высокоскоростными коммуникациями и функционирующие под управлением единой сетевой операционной системы. За тридцатилетний период своего существования кластеры превратились в мощные вычислительные ассоциации, успешно конкурирующие с суперкомпьютерами в первую очередь благодаря относительной дешевизне и высокой надежности вычислений при сравнимой производительности. Наметилась также довольно четкая дифференциация кластеров по применению, определяющему их внутреннюю структуру. Наиболее востребованными в настоящее время, в том числе по линии «облачных» вычислений, являются *HA (High Availability)* — кластеры, или кластеры высокой доступности. Кластеры этого типа создаются для обеспечения высокой доступности сервиса, предоставляемого клиенту. Ведущим критерием эффективности при этом является, как правило, коэффициент готовности, определяемый как вероятность того, что кластер готов предоставить необходимый сервис любому клиенту в любой момент времени. Этот критерий хорошо работает лишь при условии, что все клиенты с точки зрения кластера равноценны. Однако в ситуации, где клиенты дифференцированы по степени важности своих работ и/или веса рекламаций в случае отказов при их выполнении, коэффициент готовности — вряд ли адекватная мера эффективности.

Целью данной работы является: разработка достаточно адекватного критерия для оценки эффективности функционирования *HA*-кластера звездообразной структуры, обслуживающего неравноценные группы клиентов в условиях отказов отдельных его серверов; формулировка оптимизационной задачи в терминах целочисленного программирования; разработка алгоритма оптимизации стартовой структуры кластера и алгоритма реструктуризации при изменениях параметров кластера и характеристик потоков заданий клиентов.

Математическая формулировка задачи. Рассмотрим функционирование кластера звездообразной структуры типа *BeoWulf*, в котором центральный узел 0 распределяет потоки заданий клиентов по n рабочим узлам, связанным с ним высокоскоростными прямыми каналами. Структуру такого кластера удобно описывать целочисленным вектором $x = (x_1, \dots, x_n)$, где x_i — количество серверов в i -м рабочем узле, включая резервные. Эффективность функционирования кластера будем оценивать функцией $\pi(x)$ потерь, возникающих вследствие отказов серверов в рабочих узлах,

$$\pi(x) = \sum_{i=1}^n a_i \lambda_i Q_i(x_i, q_i), \quad (1)$$

где a_i — коэффициент важности задания, направляемого на обработку в i -й узел; λ_i — интенсивность соответствующего потока; $Q_i(x_i, q_i)$ — вероятность отказа (коэффициент «неготовности») i -го узла; q_i — вероятность отказа отдельного сервера в этом узле. Конкретный вид функций Q_i определяется режимом включения резервных серверов и может быть установлен в результате аналитических исследований либо на основе данных имитационного моделирования. В любом случае известно, что при каждом фиксированном значении q_i функция $Q_i(x_i, q_i)$ является убывающей функцией целочисленной переменной x_i :

$$Q_i(x_i + 1, q_i) < Q_i(x_i, q_i), \quad i = 1, \dots, n. \quad (2)$$

Пусть на построение кластера выделено известное количество m серверов, одинаковых по своим техническим характеристикам (в этом смысле рассматриваемый кластер можно назвать однородным). Проблема состоит в выборе такого размещения $(x_1, \dots, x_n)$ серверов по узлам кластера, которое минимизировало бы функцию (1) при соблюдении условий:

$$x_1 + \dots + x_n \leq m, \quad (3)$$

$$x_i = 0, 1, 2, \dots; i = 1, \dots, n. \quad (4)$$

В процессе функционирования кластера могут изменяться как ресурс m , так и параметры a_i , λ_i потоков заданий в узлах кластера. В этих условиях дополнительно к вышеуказанной проблеме необходимо корректировать размещение $(x_1, \dots, x_n)$ либо периодически, либо всякий раз при изменении перечисленных величин.

Следуя принятой терминологии, назовем структуру $x = (x_1, \dots, x_n)$, удовлетворяющую условиям (3) и (4), допустимой структурой кластера. Допустимую структуру, обеспечивающую минимум функции $\pi(x)$, будем именовать оптимальной структурой.

Из (2), в частности, следует, что оптимальная структура $(x_1, \dots, x_n)$ находится на границе n -мерной целочисленной решетки (3), и поэтому ограничение (3) можно заменить на более жесткое:

$$x_1 + \dots + x_n = m. \quad (5)$$

Следовательно, поиск оптимальной структуры можно вести в дискретном пространстве, определяемом ограничениями (4) и (5).

Алгоритмы оптимизации. Сама процедура поиска заключается в последовательном применении двух алгоритмов: алгоритма инициализации, с помощью которого определяется начальная, или стартовая, оптимальная структура кластера, и алгоритма реструктуризации, изменяющего, в случае необходимости, эту структуру при изменении числа m серверов и/или параметров потоков заданий, направляемых в рабочие узлы кластера. Алгоритм инициализации, по существу, решает задачу минимизации функции (1) при условиях (4) и (5). Он запускается в начале функционирования кластера. Алгоритм реструктуризации запускается либо периодически, либо по специальному сигналу от центрального узла всякий раз при изменении значений вышеуказанных параметров. Вопрос о правильном выборе интервала сканирования при реструктуризации — это отдельная оптимизационная проблема и здесь не рассматривается.

На основе результатов анализа возможных алгоритмов инициализации и реструктуризации [1] нами предлагаются следующие алгоритмы.

Алгоритм инициализации

Начало

1. Ввод параметров $(a_1, \dots, a_n)$, $(\lambda_1, \dots, \lambda_n)$, $(q_1, \dots, q_n)$, m и функциональных (либо табличных) зависимостей $Q_1(x_1, q_1), \dots, Q_n(x_n, q_n)$.

2. Ввод начальных значений компонентов структуры $x = (x_1, \dots, x_n)$ и рабочей переменной y : $x_i = 0$ для $i = 1, \dots, n$; $y = 0$.

3. Проверка условия $y = m$. При его выполнении — переход к п. 6.

4. Отыскание индекса k , $k = \arg \max_{1 \leq j \leq n} R_j(x_j)$,

где $R_j(x_j) = a_j \lambda_j (Q_j(x_j, q_j) - Q_j(x_j + 1, q_j))$. (6)

5. Выбор следующего приближения к оптимальной структуре:

$$x_k := x_k + 1; \quad y := y + 1.$$

Переход к п. 3.

6. Вывод стартовой оптимальной структуры $x = (x_1, \dots, x_n)$ и соответствующего ей значения минимальных потерь $\pi(x)$, вычисляемых по формуле (1).

Конец.

Алгоритм реструктуризации кластера

Начало.

1. Ввод $(a_1, \dots, a_n)$, $(\lambda_1, \dots, \lambda_n)$, $(q_1, \dots, q_n)$, m , функций $Q_1(x_1, q_1), \dots, Q_n(x_n, q_n)$ и стартовой структуры $(x_1, \dots, x_n)$.

2. Проверка условия $y = m$, где $y = x_1 + \dots + x_n$. При невыполнении условия — переход к п.3, иначе — к п.4.

3. Выбор новой допустимой структуры $x = (x_1, \dots, x_n)$, где $x_i = b$, $i = 1, \dots, n-1$, $x_n = b + m \bmod n$, и $b = [m/n]$. Здесь $[z]$ — целая часть z .

4. Проверка неравенства

$$R_s(x_s - 1) \geq R_t(x_t), \quad (7)$$

где индексы s и t находятся соответственно как $s = \arg \min_{\{j | x_j > 0\}} \{R_j(x_j - 1)\}$, $t = \arg \max_{\{j | x_j < m\}} \{R_j(x_j)\}$, и $R_j(x_j)$

определено как (6). При выполнении (7) — переход к п.6. Иначе — к п.5.

5. Выбор очередного приближения $x = (x_1, \dots, x_n)$ к оптимальной структуре: $x_s := x_s - 1$, $x_t := x_t + 1$, $x_i := x_i$ для $i \neq s, t$; переход к п. 4.

6. Вывод откорректированной структуры x и соответствующего ей значения потерь $\pi(x)$, вычисляемых по формуле (1).

Конец.

При каждом последующем запуске алгоритма реструктуризации кластера в качестве стартовой структуры для него выбирается скорректированная оптимальная структура, полученная при предыдущем запуске.

Примеры. Одной из главных задач при кодировании алгоритмов и инсталляции соответствующих программ является задача выбора функций $Q(x, q)$ при конкретных режимах использования резервных серверов. Для многих простых режимов можно воспользоваться известными оценками [2]. Так, например, для кластера типа связки *DRBD* и *HeartBeat* [3] (холодное резервирование)

$$Q(x, q) = \begin{cases} 1 & \text{при } x = 0, \\ 1 - \sum_{j=0}^{x-1} \frac{q^j}{j!} e^{-q} & \text{при } x > 0; \end{cases}$$

для кластера типа Microsoft Cluster Server (горячее резервирование) — $Q(x, q) = q^x$.

Существуют также аналитические оценки [2] и для различных типов скользящего резервирования, которые, например, можно использовать в кластерах с топологией « $n + 1$ » [3]. Если режимы сочетаются в любой комбинации (например, в *Linux-HA*), то можно рекомендовать табличный способ задания функций $Q(x, q)$; элементы таблицы получаются в результате статистической оценки ситуаций в соответствующих натурных или модельных экспериментах.

Выводы. В работе предложена модель оценки эффективности однородного кластера высокой доступности. На основе этой модели в терминах целочисленного программирования сформулирована задача оптимального распределения серверов по узлам кластера определенной (звездообразной) топологии по критерию минимума ожидаемых потерь, возникающих вследствие ненадежности (недоступности) серверов. Предложены стартовый алгоритм оптимального распределения, а также алгоритм оптимальной реструктуризации кластера в условиях изменения его ресурсов и флуктуаций параметров потоков заданий от клиентов. Разработаны компьютерные программы, реализующие алгоритмы. Программы могут быть использованы в составе инструментария центрального узла при оптимизации функционирования кластера в целом.

В перспективе дальнейших исследований предполагаются постановка и решение подобных задач для неоднородных (гетерогенных) кластеров, а также решение задачи выбора оптимального периода включения процедуры реструктуризации.

Бibliографический список использованной литературы

1. Кирюхин В.В. Оптимальное распределение средств при аддитивных функциях цели / В.В. Кирюхин // Известия АН СССР. Сер. Техническая кибернетика. — 1967. — № 6. — С. 49–53.
2. Козлов Б.А. Справочник по расчету надежности аппаратуры радиоэлектроники и автоматики / Б.А. Козлов, И.А. Ушаков. — М.: Советское радио, 1975. — 472 с.
3. Кластеры, практическое руководство по параллельным вычислениям. [Электронный ресурс]. — Режим доступа <http://cluster.linux-ekb.info/map.gif/netware2.php>

Поступила в редакцию 04.07.2011 г.

Кірюхін В.В. Алгоритми динамічної реструктуризації однорідного кластера високої доступності

Розглядається задача оптимального розподілу серверів у вузлах високо доступного однорідного кластера з метою мінімізації середніх втрат через відмови серверів в умовах змінності навантаження і параметрів кластера. Пропонуються алгоритми вирішення відповідної задачі цілочислового програмування.

Ключові слова: хмарно обчислення, кластер високої доступності, динамічна реструктуризація, цілочисельне програмування.

Kiryukhin V.V. Algorithms for dynamic restructuring of a homogeneous HA-cluster

Paper deals with the problem of optimal allocation of servers to nodes of homogeneous HA-cluster to minimize the average loss due to failures of servers, in the presence of flows and resources variations. Algorithms for solving the corresponding integer programming problem is proposed.

Keywords: cloud computing, HA-cluster, dynamic restructuring, integer programming.