

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ БАЛАНСИРОВКИ НАГРУЗКИ НА ВЫЧИСЛИТЕЛЬНУЮ СЕТЬ

А.Ю. Чичканов, В.Е. Подольский

ФГБОУ ВПО «Тамбовский государственный технический университет», г. Тамбов

Рецензент д-р техн. наук, профессор В.А. Немтинов

Ключевые слова и фразы: балансировка нагрузки; вычислительные сети; облачные технологии; обработка данных.

Аннотация: Описана программная система для повышения эффективности использования имеющихся технических средств. Предложен алгоритм балансировки нагрузки, обеспечивающий повышение производительности в обработке данных, что существенно повысит мобильность оконечных терминалов и позволит снизить затраты на обработку информации.

В современном мире широкое распространение получили портативные устройства, позволяющие получать, модифицировать и хранить данные. Они призваны повысить мобильность пользователя, предоставив функциональность персонального компьютера. Однако зачастую они не могут обеспечить производительность, необходимую для комфортной работы с данными. В то же время, развитие широкополосных линий связи дает возможность осуществлять передачу данных на больших скоростях. Для компенсации недостаточной производительности мобильных устройств были предложены так называемые «облачные технологии». Данный термин можно отнести к большому количеству сервисов, предоставляемых в сети Internet. Облачные технологии позволяют обрабатывать данные удаленно, а широкополосный Интернет обеспечивает удобное их использование [2].

Однако с распространением облачных технологий возникла проблема рационального их использования. Обработывающие сервера, объединенные в вычислительные сети, являются дорогостоящим оборудованием, эффективное использование которого снижает затраты на содержание.

В качестве критерия эффективности использования вычислительных сетей можно принять время, необходимое для вычисления. В вычисли-

Чичканов Алексей Юрьевич – студент группы ССП51, e-mail: kolob_ok@bk.ru; Подольский Владимир Ефимович – доктор технических наук, профессор кафедры «Системы автоматизированного проектирования», проректор по информатизации, ТамбГТУ, г. Тамбов.

тельных сетях использование распределения нагрузки для оптимизации использования ресурсов называют балансировкой нагрузки [3].

В настоящее время существуют проектные решения, разработанные для повышения эффективности работы вычислительных сетей. Среди них есть как программные продукты (например, Perlbal), так и программно-аппаратные комплексы (F5 Networks, Cisco, Citrix).

Для балансировки нагрузки могут использоваться различные алгоритмы, наиболее очевидным из них является циклический, который состоит в следующем: пусть имеется N объектов, способных выполнить заданное действие, и M задач, которые должны быть выполнены этими объектами. Подразумевается, что объекты n равны по своим свойствам между собой, задачи m имеют равный приоритет. Тогда первая задача $m = 1$ назначается для выполнения первому объекту $n = 1$, вторая – второму и т.д., до достижения последнего объекта $m = N$. Тогда следующая задача $m = N + 1$ будет назначена снова первому объекту и т.п. В результате, происходит перебор выполняющих задания объектов по циклу, и по достижении последнего объекта следующая задача будет также назначена первому объекту.

Однако данный алгоритм имеет ряд существенных недостатков. Наиболее критичным из них стоит считать неэффективное распределение ресурсов: алгоритм рассчитан на сеть из нескольких одинаковых серверов и не позволяет распределять различные типы задач по серверам различных конфигураций.

При разработке автоматизированной системы балансировки нагрузки основной задачей ставилось решение проблемы различной производительности обрабатывающих серверов.

Перед началом работы системы необходимо выполнить настройку. Для запуска системы в режиме настройки необходимо перейти на страницу настроек и указать необходимые параметры: количество серверов и их IP-адреса.

После этого начнет работу подсистема определения производительности обрабатывающих серверов: подсистема поочередно будет производить серии тестов с каждым сервером, вычислять коэффициенты обработки для каждого кодека на сервере, составлять таблицы коэффициентов и записывать их в базу данных. По окончании работы подсистемы в базе данных будет создана таблица коэффициентов. В случае успешного выполнения настройки пользователю будет выведено сообщение и предложена возможность просмотреть таблицу коэффициентов.

Для начала работы системы в стандартном режиме пользователь осуществляет подключение к системе. После установления соединения, пользователь попадает на web-страницу, на которой он может выбрать файлы для загрузки со своего жесткого диска и указать каким образом они должны быть обработаны. После того, как он указал требования к обработке, он отправляет форму, содержащую информацию по обработке на сервер.

Информация о файлах передается в подсистему оптимального распределения, под которым понимают задачу нахождения экстремального размещения, сформулированную следующим образом: используя известное количество файлов p и число серверов m , получить все

возможные размещения и среди размещений r , образующих множество размещений R , найти размещение r^* , при котором

$$T_0(r^*) = \min_{r \in R} T_0(r), \quad (1)$$

где T_0 – общее время обработки.

Процесс нахождения всех размещений называют конструированием вариантов и осуществляют в виде построения дерева, называемого порфириан.

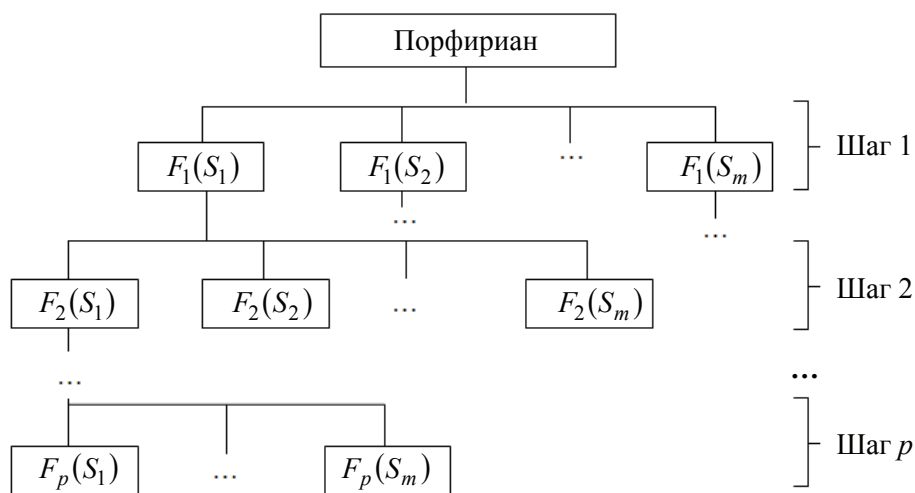
Пусть имеется множество файлов $F = \{F_1, F_2, \dots, F_p\}$ и множество серверов $S = \{S_1, S_2, \dots, S_m\}$, тогда дерево размещений может быть построено следующим образом: F_1 будет обрабатываться на сервере S_1 . Тогда файл F_2 может обрабатываться S_1, S_2 , или на любом из m серверов. Для каждого из этих вариантов верно и то, что файл F_3 может быть обработан на одном из m серверов. Аналогично для всех p файлов. Процесс такого последовательного разбиения представлен на рисунке. В результате работы алгоритма будет получено дерево, содержащее все варианты размещений [1].

Для нахождения времени на обработку k -го файла используем следующую формулу

$$T_k = \frac{S_k}{P_n} Z_k, \quad (2)$$

где T_k – время на обработку файла; S_k – размер k -го файла (находится подсистемой ввода-вывода на этапе загрузки); P_n – скорость копирования файла (находится подсистемой определения производительности для каждого сервера); Z_k – коэффициент, показывающий во сколько раз время, потраченное на обработку сервером файла, больше времени затраченного на копирование этого файла.

Тогда общее время на обработку находим расчетом времени обработки последнего файла. Для этого необходимо вычислить время обработки каждого списка файлов и определить наибольшее. Рассчитываем время



Дерево размещений

обработки списка файлов сложением времен обработки всех файлов

$$T_S = \sum_{l=0}^l S_k / P_n Z_k, \quad (3)$$

где T_m – время обработки одного списка, $S < m$; l – количество файлов в списке.

Общее время на обработку

$$T_o = \max T, \quad (4)$$

где $T = \{T_0, T_1, \dots, T_m\}$ – множество времен обработки каждого списка.

После того, как время для обработки каждого размещения найдено, находится наименьшее время и размещение, соответствующее наименьшему времени, которое считается оптимальным и записывается в базу данных.

Разработанная система позволяет осуществлять выполнение задач за минимальное время и тем самым существенно снизить затраты, связанные с обработкой данных. Это, в свою очередь, позволяет наращивать мощность вычислительных сетей, чтобы расширить возможности конечных терминалов, компенсировав недостаточную производительность.

Список литературы

1. Альфс, Б. Деревья. Структуры данных / Б. Альфс. – М. : Статистика, 1974. – 131 с.

2. Подольский, В.Е. Повышение эффективности региональных образовательных компьютерных сетей с использованием элементов структурного анализа и теории сложности : монография / В.Е. Подольский, С.С. Толстых. – М. : Машиностроение, 2006. – 175 с.

3. Информационные технологии в САПР. Вычислительные сети и компьютерная графика / С.А. Васильев [и др.]. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2008. – 80 с.

Development of Automated System for Network Load Balancing

A.Yu. Chichkanov, V.E. Podolsky

Tambov State Technical University, Tambov

Key words and phrases: “cloud computing”; computer networks; data processing; load balancing.

Abstract: This article describes the software system to improve effective use of available resources. The authors proposed an algorithm of load balancing, which provides increased productivity in data processing. This greatly enhances the mobility of end terminals and reduces the cost of data processing.

© А.Ю. Чичканов, В.Е. Подольский, 2012