

УДК 681.324

# Синтез распределенной вычислительной среды с многоуровневой иерархической архитектурой

Р. М. АЛГУЛИЕВ, Р. К. АЛЕКПЕРОВ, И. М. АЛИЕВ

Институт информационных технологий НАН Азербайджана, г. Баку

*Предложена модель синтеза надежной распределенной вычислительной среды с многоуровневой иерархической структурой. В качестве базовой среды служат компьютерные сети. При организации распределенных вычислений в качестве критериев оптимизации используются параметры стоимости и надежности.*

**А**нализ различных подходов к организации распределенных вычислений на базе локальных и глобальных компьютерных сетей показывает, что они наряду с преимуществами имеют ряд недостатков, одним из которых является проблема надежности. Несомненно, объединение большого количества географически распределенных компьютеров в единую распределенную вычислительную среду (РВС) предъявляет жесткие требования к надежности ее функционирования. Поэтому при организации распределенных вычислений необходимо учитывать этот важный фактор.

В данной работе рассмотрены вопросы синтеза РВС по критериям "стоимость—надежность".

Надежность распределенных вычислительных систем относится к числу достаточно продвинутых областей исследования [1, 2]. Для организации распределенных вычислений на базе компьютерных сетей с учетом надежности используется мультиагентная технология, которая в последнее время находит широкое применение [3, 4]. Согласно данной технологии агенты наряду с ведением статистики свободных вычислительных ресурсов (СВР) сети сообщают администратору сети данные об объеме СВР каждого компьютера, осуществляют контроль за ходом выполнения подзадач и передачей результатов вычислений. Они регистрируют сбои в компьютерах, собирают статистические данные о сбоях, и на основе этих статистических данных оценивается надежность их функционирования.

## 1. Математическая постановка задачи

Пусть требуется гарантированное решение задачи  $Z$  за время  $T_0$  с заданной надежностью. Для решения поставленной задачи необходимо организовать РВС с производительностью  $M_Z$ . Решение задачи  $Z$  требует создания многоуровневой архитектуры.

Пусть администратором вычислительной сети выделено  $n + 1$  компьютеров для организации РВС. Один из компьютеров, входящих в состав

РВС, является центральным компьютером, в функции которого входят: распараллеливание задач на отдельные подзадачи; распределение этих подзадач компьютерам; прием результатов решения подзадач; выработка окончательных результатов и т. п. Пусть РВС имеет  $k$ -уровневую архитектуру, где на первом уровне находится центральный компьютер, а на более низких — остальные компьютеры. Общее количество компьютеров, входящих в состав конкретной РВС, и количество компьютеров на каждом уровне определяются сложностью и типом задачи, а также стратегией распараллеливания.

Пусть для каждого уровня определено множество подзадач  $z_k = \{z_{i_k} | i_k = \overline{1, m_k}\}$ , где  $z_{i_k}$  —  $i_k$ -я подзадача  $k$ -го уровня;  $m_k$  — количество подзадач на  $k$ -м уровне.

Обозначим  $T_Z$  фактическое время выполнения задачи  $Z$ , которое имеет следующие основные составляющие: время распараллеливания задачи; среднее время передачи распараллеленных подзадач от главного компьютера к компьютерам РВС; время решения подзадачи и среднее время передачи результатов решения подзадачи к главному компьютеру РВС.

Обозначим  $J_k$  множество индексов компьютеров, входящих в  $k$ -й уровень РВС. В этом случае задача заключается в определении таких подмножеств  $J_k$ , которые обеспечивают

$$P_Z = \prod_{k=1}^L \prod_{j_k \in J_k} p_{j_k} \rightarrow \max \quad (1)$$

при условии, что

$$T_Z \leq T_0; \quad \sum_{k=1}^L \sum_{j_k \in J_k} p_{j_k} \geq \sum_{k=1}^L M_k, \quad (2)$$

где  $P_Z$  — вероятность выполнения задачи  $Z$ ;  $p_{j_k}$  — вероятность того, что сбой (или отказ) не появится при решении произвольно выбранной подзадачи в  $j_k$ -м компьютере  $k$ -го уровня;  $b_{j_k}$  —

свободный вычислительный ресурс  $j_k$ -го компьютера;  $M_k$  — требуемая вычислительная мощность для решения задачи на  $k$ -м уровне. Здесь 
$$\sum_{k=1}^L M_k = M_Z.$$

Для определения СВР отдельных компьютеров мультиагенты по запросу центрального компьютера передают данные о фактической загрузке компьютеров. Обозначим коэффициент загрузки  $j_k$ -го компьютера  $v_{j_k}$ . По этим данным главный компьютер вычисляет СВР каждого компьютера и составляет матрицу  $B_k = \{b_{j_k}\}$ ,  $j_k = \overline{1, n_k}$ , где  $b_{j_k} = (1 - v_{j_k})\pi_{j_k}$ . Здесь  $\pi_{j_k}$  — реальная производительность  $j_k$ -го компьютера;  $n_k$  — число компьютеров на  $k$ -м уровне.

В случае, если РВС имеет одноуровневую архитектуру, т. е.  $L = 1$ , то модель (1) преобразуется к следующему виду:

$$P_Z = \prod_{j \in J^*} p_j \rightarrow \max, \quad (3)$$

где  $p_j$  — вероятность того, что сбой (или отказ) не появится при решении подзадачи в  $j$ -м компьютере. Естественно, при этом необходимо соблюдать условие  $\sum b_j \geq M$ , где  $b_j$  — свободные вычислительные ресурсы  $j$ -го компьютера.

Для решения моделей (1) и (3) все компьютеры классифицируют по параметру надежности и для РВС выбирают компьютеры, имеющие максимальную надежность. При определении надежности отдельных компьютеров используются статистические данные, собранные с помощью агентов, установленных непосредственно в компьютерах.

Рассмотрим стоимостную модель. Пусть требуется решить задачу  $Z$  за время  $T_0$  с заданной надежностью при минимальной стоимости. Для этого создается  $k$ -уровневая архитектура РВС.

Пусть известна стоимость  $c_{j_k}$  ( $j_k = \overline{1, n_k}$ ) использования  $j_k$ -го компьютера с реальной производительностью за единицу времени. Тогда с учетом загрузки компьютеров составляем матрицу

$$C = \{C_{i_k j_k} | i_k = \overline{1, m_k}, j_k = \overline{1, n_k}\},$$

где  $C_{i_k j_k} = (1 - v_{j_k})c_{j_k}t_{i_k}$  — стоимость выполнения подзадачи  $z_{i_k}$  на  $j_k$ -м компьютере с загрузкой  $v_{j_k}$ ;  $t_{i_k j_k}$  — время решения подзадачи  $z_{i_k}$  в  $j_k$ -м компьютере.

Введем следующую булеву переменную:

$$x_{i_k j_k} = \begin{cases} 1, & \text{если подзадача } z_{i_k} \text{ решается в} \\ & j_k\text{-м компьютере;} \\ 0 & \text{— в противном случае.} \end{cases}$$

Таким образом, задача состоит в определении таких значений переменной  $x_{i_k j_k}$ ,  $i_k = \overline{1, m_k}$ ,  $j_k = \overline{1, n_k}$ , для которых целевая функция  $C_Z$  имеет минимальное значение:

$$C_Z = \sum_{k=1}^L \sum_{i_k=1}^{m_k} \sum_{j_k=1}^{n_k} c_{i_k j_k} x_{i_k j_k} \rightarrow \min \quad (4)$$

при условии

$$T_Z \leq T_0; \quad P_Z \geq P_0,$$

где  $P_0$  — заданный уровень надежности решения задачи  $Z$ .

Поскольку для каждой подзадачи  $z_{i_k}$ ,  $i_k = \overline{1, m_k}$ , выбирается только один компьютер, дополнительно должно быть учтено:

$$\sum_{i_k=1}^{m_k} \sum_{j_k=1}^{n_k} x_{i_k j_k} = m_k, \quad k = \overline{1, L};$$

$$\sum_{k=1}^L M_k = M.$$

В случае, если РВС имеет одноуровневую архитектуру, модель (4) будет иметь вид

$$C_Z = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min$$

при тех же условиях, что и модель (4). Здесь  $x_{ij}$  определяется следующим образом:

$$x_{ij} = \begin{cases} 1, & \text{если подзадача } z_i \text{ решается в} \\ & j\text{-м компьютере;} \\ 0 & \text{— в противном случае.} \end{cases}$$

Пусть компьютеры, организующие распределенную вычислительную среду, характеризуются пуассоновским потоком сбоев с интенсивностью  $\lambda$ . Из-за изменения сводных ресурсов компьютеров длительность решения подзадачи является случайной величиной. В общем виде вероятность безотказной работы компьютеров при случайной длительности выполнения задачи определяется как [5]

$$P = \int_0^{\infty} P(t) dG(t) = \int_0^{\infty} P(t) g(t) dt,$$

где  $P(t)$  — вероятность безотказной работы компьютеров за время  $t$ ;  $G(t)$  — вероятность того, что длительность подзадачи  $z_j$  не превысит величину  $t$ ;  $g(t)$  — плотность распределения  $G(t)$ .

Если длительность решения подзадач является случайной величиной с произвольным распределением  $G(t)$ , то вероятность того, что сбой не появится при решении произвольно выбранной подзадачи, может определяться как

$$P = 1 - \sum_{a=1}^{\infty} (-1)^a \frac{\lambda^a x_a}{a!},$$

где  $x_a$  — начальный момент распределения  $G(t)$ ,  $x_1 = t_0$ ,  $x_2 = t_0^2 + \sigma^2$ . Здесь  $\sigma$  — среднеквадратичное отклонение. На практике можно использовать приближенное выражение:

$$P = 1 - \lambda t_0 + \frac{\lambda[t_0^2 + \sigma^2]}{2!}.$$

Если длительность решения подзадач распределена по нормальному закону со средним  $t_{cp}$  и среднеквадратическим  $\sigma$ , то в этом случае вероятность  $p_j$  того, что сбой не появится при решении произвольно выбранной подзадачи в  $j$ -м компьютере, определяется как

$$p_j = e^{-\lambda_j t_{cp} + \frac{\sigma^2 \lambda_j^2}{2}},$$

где  $\lambda_j$  — интенсивность отказа  $j$ -го компьютера.

● Сформулированная задача принадлежит к классу комбинаторных задач, традиционно трудно решаемых. Во многих случаях единственным возможным способом решения таких задач является полный перебор всех комбинаций значений варьируемых переменных. Объемы вычислений при полном переборе вариантов катастрофически растут с увеличением размерности задачи. Для решения данной модели можно использовать нейронные сети, генетические и другие алгоритмы.

УДК 681.324(07)

## Управление трафиком в среде теледоступа к ресурсам СКЦ

М. Л. ГОЛЬДШТЕЙН, Н. В. ЗАКУРДАЕВ, Д. Д. КАЙКЫ, А. В. МЕРКУЛЬЕВ

Институт математики и механики УрО РАН, г. Екатеринбург

*Предложена технология управления информационными потоками в среде теледоступа к ресурсам суперкомпьютерного центра Института математики и механики Уральского отделения РАН в целях обеспечения "бесконфликтного" трафика в рамках иерархической структуры сети с консультационно-диагностическими и интеллектуальными элементами.*

**К**орректное построение системы коллективного удаленного доступа регионального суперкомпьютерного центра (СКЦ) [1], оснащенного разнообразными вычислительными, информационными, отладочными, учебными и т. п. ресурсами и предоставляющего эти ресурсы научной общественности Уральского отделения РАН, — залог комфортной работы и эффективной загрузки дорогостоящего оборудования.

По мере увеличения объема вычислительных и информационных ресурсов, числа пользователей и появления новых мультимедийных технологий требования к пропускной способности системы доступа возрастают. Все это приводит к

возникновению "заторов" при доступе к единому ресурсу. В качестве ресурса может выступать физический канал, сетевое оборудование, сервер, вычислитель. Применение технологий QoS — одно из эффективных и относительно недорогих средств достижения "бесконфликтного" трафика в среде теледоступа к ресурсам СКЦ.

На сегодняшний день локальная сеть СКЦ имеет открытую, масштабируемую двухступенчатую структуру на основе гигабитной распределенной магистрали (рис. 1) с применением сетевых средств (СОС) фирмы 3Com. В качестве магистральных (1-я ступень) выбраны коммутаторы серии 4005, в качестве коммутаторов 2-й ступени — коммутатор SuperStack