

Модели управления системами серверного обслуживания

Ревотюк М.П.; Зобов В.В.; Шешко Е.В.

Кафедра информационных технологий автоматизированных систем
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
e-mail: rmp@bsuir.by

Аннотация— Рассмотрены модели управления в реальном времени системами серверного обслуживания потока запросов с регламентированными сроками обслуживания. Предложена процедура оптимизации управления методом динамического программирования.

Ключевые слова: системы серверного обслуживания; задачи управления; сетевые модели

Технологии построения распределенных систем (например, CORBA, DCOM, Java-RMI, SOAP и RPC) порождают потребность управления конфигурацией среды серверного обслуживания [1]. Возникает задача оптимизации управления многоэтапными процессами с учетом временных ограничений на сроки выполнения отдельных этапов. Ее решение будем проводить методом динамического программирования на основе стратегии ЛТ (Just-In-Time). Формализация задачи для простоты изложения будет отражать отношение взаимозависимости этапов процесса на основе линейной временной логики [2].

Пусть в некоторый момент времени t определено множество агентов $S = \{S_i, i = \overline{1, m}\}$. Каждый из агентов характеризуется интервалом доступности во времени $[a^i, b^i]$ причем $(a^i > t) \wedge (a^i < b^i)$, $i = \overline{1, m}$. Агенты S должны решить задачи из множества $\Pi = \{\Pi_j, j = \overline{1, n}\}$. Процесс решения каждой задачи может требовать использования обеспечивающей системы S_0 , играющей роль элемента синхронизации всех процессов обслуживания. Интервал $[a_0, b_0]$ существования синхронизирующей системы S_0 должен перекрывать интервалы функционирования агентов, что соответствует условию $(a^0 < \min\{a^i, i = \overline{1, m}\}) \wedge (b^0 > \max\{a^i, i = \overline{1, m}\})$.

Каждую из задач будем характеризовать тройкой $\Pi_j = \langle s_j, f_j, c_j \rangle$, где $[s_j, f_j]$ – заданный интервал решения задачи Π_j , а c_j – продолжительность взаимодействия агента, решающего эту задачу, с синхронизирующей системой, $j = \overline{1, n}$.

Порядок их решения естественно определить относительно номеров этапов взаимодействия с синхронизирующей системой S_0 :

$$\Pi_i \prec \Pi_{i+1} : (t_i + c_i < t_{i+1}) \wedge ((s_i < s_{i+1}) \vee ((s_i = s_{i+1}) \wedge (f_i < f_{i+1}))), i = \overline{1, n}.$$

Оптимальное управление пусть соответствует минимизации смещения интервалов решения задач от

идеальных значений. Используя метод математической индукции, можно показать, что решение здесь может быть найдено следующим алгоритмом.

Шаг 1. Фиксация начальных условий

$$\begin{cases} T_0 = t - 1; c_0 = 0; \\ R_j^1 = a^j, j = \overline{1, m}. \end{cases}$$

Шаг 2. Прямое движение

$$\begin{cases} k^i = \arg \min_j \{R_j^i, j | f_i + \max(0, R_j^i - s_i) < b^j\}; \\ T_i = \max(T_{i-1} + c_{i-1}, R_{k^i}^i); \\ R_{k^i}^{i+1} = f_i + \max(0, T_i - s_i), k^i > 0; \\ R_j^{i+1} = \max(R_j^i, T_i + c_i), j \neq k^i, i = \overline{1, n}. \end{cases}$$

Шаг 3. Обратное движение

$$\begin{cases} t_{n+1} = b^0 + 1, \\ t_i = \max(T_i, \min(t_{i+1} - c_i, s_i)), i = \overline{n, 1}. \end{cases}$$

Здесь значения R_j^i – момент готовности агента S_j к решению задачи Π_i , $i = \overline{1, n}$, $j = \overline{1, m}$.

Последовательность пар $\{(k^i, t_i), i = \overline{1, n}\}$ определяет назначение агента S_{k^i} для решения Π_i в момент времени t_i , $i = \overline{1, n}$. Если $k^i = 0$, то решение задачи Π_i в сложившихся условиях невозможно.

Фактически обсужденное отношение вида “один ко многим” в задаче управления синхронизацией может рассматриваться как элемент шаблона представления произвольных структур взаимозависимости между процессами. Таким образом, задача управления обслуживанием потока запросов серверной средой оказывается представленной в виде сетевой модели.

[1] Tianying Chang, Mustaque Ahamad, "Improving Service Performance through Object Replication in Middleware: A Peer-to-Peer Approach," Fifth IEEE International Conference on Peer-to-Peer Computing (P2P'05), 2005, pp.245-252.

[2] Ревотюк М.П., Чан З.А., Шешко Е.В. Модели оптимизации управления на основе сетей свободного выбора// Труды V Межд. конф. “Идентификация систем и задачи управления” SICPRO'06, Москва, 30 января – 2 февраля 2006 г./М: ИПУ им. В.А. Трапезникова РАН, 2006. – С. 2159-2170.