СПИСОК ЛИТЕРАТУРЫ

1. Рекомендация МСЭ-Т Y.1541 (02/2006 г.). Требования к сетевым показателям качества для служб, основанных на протоколе IP. 2006.
2. *Алиев Т. И.* Основы моделирования дискретных систем. СПб: СПбГУ ИТМО, 2009. 363 с.
3. *Алиев Т. И.* Характеристики дисциплин обслуживания заявок с несколькими классами приоритетов // Изв. АН СССР. Техническая кибернетика. 1987. № 6. С. 188—191.
4. *Алиев Т. И.* Задачи синтеза систем с потерями // Изв. вузов. Приборостроение. 2012. Т. 55, № 10. С. 57—63.
5. *Богатырев В. А., Богатырев С. В., Богатырев А. В.* Функциональная надежность вычислительных систем с перераспределением запросов // Изв. вузов. Приборостроение. 2012. Т. 55, № 10. С. 53—56.

Сведения об авторе

Тауфик Измайлович Алиев — д-р техн. наук, профессор; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; заведующий кафедрой E-mail: aliev@d1.ifmo.ru

Рекомендована кафедрой вычислительной техники

Поступила в редакцию 23.12.13 г.

УДК 004.89: 002.53

Л. А. Муравьева-Витковская

МЕТОД РАСЧЕТА ХАРАКТЕРИСТИК ЗАМКНУТЫХ ДЕТЕРМИНИРОВАННЫХ МОДЕЛЕЙ МУЛЬТИСЕРВИСНЫХ КОМПЬЮТЕРНЫХ СЕТЕЙ

Рассматривается метод расчета характеристик мультисервисных компьютерных сетей, моделями которых являются замкнутые сети с детерминированным временем обслуживания заявок в узлах. Предположение о детерминированном обслуживании позволяет учитывать реальную статистику распределений фиксированных длин пакетов в мультисервисных компьютерных сетях.

Ключевые слова: *мультисервисные компьютерные сети, детерминированное обслуживание, замкнутые сети массового обслуживания.*

Введение. Для анализа процесса функционирования мультисервисных компьютерных сетей (КС) широко используются сетевые модели массового обслуживания, учитывающие наличие множества ресурсов. Замкнутые сети массового обслуживания (МО), содержащие постоянное число заявок, успешно применяются для моделирования работы мультисервисных КС [1].

Хорошо известны методы расчета характеристик замкнутых сетей МО при распределении по экспоненциальному закону временных интервалов обслуживания заявок в узлах сети. В настоящей статье предлагается метод расчета характеристик мультисервисных КС, моделями которых являются замкнутые сети с детерминированным временем обслуживания заявок в узлах. Предположение о детерминированном обслуживании позволяет учитывать реальную статистику распределений фиксированных длин пакетов [2].

Постановка задачи. В качестве моделей функционирования мультисервисных КС рассматриваются замкнутые сети МО, содержащие n узлов с произвольным количеством обслуживающих приборов, время обслуживания в которых детерминировано. Пакетам в моделях мультисервисных КС будут соответствовать заявки.

Для анализа процесса функционирования мультисервисных КС с кольцевой топологией, узлы которых связаны между собой последовательно, может использоваться циклическая модель с одноканальными узлами. Узлы модели отображают задержки, возникающие в узлах передачи данных, при этом выходная информация предыдущего узла является входной для последующего элемента структуры.

Циклическая модель с многоканальными узлами может применяться для анализа процесса функционирования локальных сетей, имеющих кольцевую топологию, центры коммутации в которых могут содержать один или несколько процессоров.

Для анализа работы мультисервисных КС, имеющих звездообразную структуру, возможно использовать модель центрального обслуживания [3]. В данной модели первый узел — центральный, отображающий работу главной ЭВМ; а узлы 2, ..., n — периферийные, соответствующие терминалам, посылающим сообщения через главную ЭВМ по дуплексным каналам; p_2, …, p_n — вероятность передач из центрального узла в периферийные. В модели выделяется дуга, являющаяся внешней по отношению к сети, на которой отмечается нулевой узел „0“. Относительно нулевого узла определяются временные характеристики функционирования сети [4].

Оценить качество функционирования мультисервисных КС можно, определив характеристики сетевых моделей, основной из которых является производительность λ_0, измеряемая как интенсивность потока заявок, проходящих через нулевой узел.

Циклическая однородная модель с одноканальными узлами. Положим, что в n-узловой замкнутой сетевой модели циркулирует M заявок. Тогда коэффициент загрузки узла j (j = 1, ..., n) определяется как

$$\rho_j = \begin{cases} b_j / b_{\max}, \text{если } M\, b_{\max} \ge B; \\ M\, b_j / B, \text{если } M\, b_{\max} < B, \end{cases} \tag{1}$$

где b_j — время обслуживания заявки в узле j; $b_{\max} = \max\{b_1,...,b_n\}$; $B = \sum_{j=1}^{n} b_j$.

Справедливость выражения (1) вытекает из следующих рассуждений. Параметр B — время обслуживания одной заявки в сети, $M\, b_{\max}$ — время обслуживания всех заявок в максимально загруженном узле сети. Тогда условие $M\, b_{\max} \ge B$ означает, что время обслуживания всех заявок в максимально загруженном узле будет больше времени, затраченного на обслуживание одной заявки во всей сети, т.е. некоторая заявка вернется в максимально загруженный узел раньше, чем в нем завершится обслуживание всех остальных заявок. Таким образом, этот максимально загруженный узел никогда не будет простаивать, т.е. его коэффициент загрузки равен единице: $\rho_{\max} = 1$. Коэффициенты загрузки остальных узлов связаны с рассматриваемым коэффициентом известным соотношением $\rho_j / \rho_{\max} = \lambda_0 b_j / \lambda_0 b_{\max}$, откуда $\rho_j = b_j / b_{\max}$.

При $M\, b_{\max} < B$ коэффициент загрузки максимально загруженного узла меньше единицы. Тогда, рассмотрев интервал времени, равный B, можно заметить, что время, в течение которого узел j занят обслуживанием всех заявок, циркулирующих в сети, равно $M\, b_j$, откуда следует, что коэффициент загрузки определяется как $\rho_j = M\, b_j / B$, что и требовалось доказать.

Циклическая однородная модель с многоканальными узлами. Пусть, как и в предыдущем случае, в сети циркулирует M заявок. Число обслуживающих приборов K_j в узлах сети произвольно. Тогда коэффициент загрузки узла j (j = 1, …, n) определяется следующим образом:

$$\rho_j = \begin{cases} x_j / x_{\max}, \text{если } M\, x_{\max} \geq B; \\ M\, x_j / B, \text{если } M\, x_{\max} < B, \end{cases} \quad (2)$$

где $x_j = b_j/K_j$; $x_{\max} = \max\{x_1, ..., x_n\}$.

Очевидно, что рассуждения, доказывающие справедливость формулы (1), сохраняют силу и в данном случае. Отличие заключается в способе определения значения времени обслуживания всех заявок в узле и коэффициентов загрузок. Для многоканального узла время обслуживания всех заявок в j-м узле определяется как $M\, b_j / K_j$, а коэффициент загрузки

$$\rho_j = \lambda_0\, b_j / K_j. \quad (3)$$

Следовательно, если $\rho_{\max} = 1$, то $\rho_j / \rho_{\max} = x_j / x_{\max}$, откуда $\rho_j = x_j / x_{\max}$. При отсутствии узла с $\rho = 1$ параметр ρ_j можно определить как отношение времени обслуживания всех заявок в узле j ко времени обслуживания одной заявки в сети $\rho_j = M\, x_j / B$.

Однородная модель центрального обслуживания с одноканальными узлами. Заметим, что в этой модели коэффициенты передачи, в отличие от предыдущих моделей, совпадают со значениями вероятности передач: $\alpha_j = p_j$ $(j = 2,...,n)$; $\alpha_1 = 1$. В сети, как и ранее, циркулирует M заявок.

Расчет такой модели осложняет случайный характер переходов из центрального узла в периферийные, описываемый в виде вероятностей передач p_2, ..., p_n. Однако для модели центрального обслуживания (МЦО), в которой соблюдается локальный баланс для периферийных узлов, т.е. $\alpha_j\, b_j = \text{const}$ для j = 2, ..., n, можно предложить достаточно простой метод приближенного расчета характеристик сети. Метод предполагает сведение исходной сети центрального обслуживания к двухузловой циклической, второй узел которой представляет собой многоканальную систему МО, с $n-1$ обслуживающими приборами, в которых время обслуживания определяется из условия эквивалентности экспоненциальных сетей МО с центральным обслуживанием и двухузловой с многоканальным (МК) вторым узлом [4]. Время обслуживания в приборах многоканального узла циклической сети будем полагать равным $\alpha_j\, b_j$ (j = 2, ..., n). Тогда по формуле (2) определим коэффициент загрузки 1-го узла ρ_1, производительность сети $\lambda_0^{\text{МК}} = \rho_1 / b_1$, затем для $\lambda_0^{\text{МЦО}} = \lambda_0^{\text{МК}}$ получим формулу для расчета коэффициентов загрузки узлов j = 2, …, n в модели центрального обслуживания:

$$\rho_j = \alpha_j b_j \rho_1 / b_1. \quad (4)$$

Циклическая неоднородная модель с одноканальными узлами. Рассмотрим замкнутую сетевую модель с n одноканальными узлами, в которой циркулируют заявки H классов. Число заявок в сети $M = \sum_{i=1}^{H} m_i$, где m_i — число заявок класса i. Время обслуживания заявок класса i в j-м узле b_{ij} детерминировано, тогда $B_j = \sum_{i=1}^{H} m_i\, b_{ij}$ — время обслуживания всех заявок в j-м узле (j = 1,..., n), а $B_{\max} = \max\{B_1,...,B_n\}$ — время обслуживания всех заявок в максимально загруженном узле. Время обслуживания заявки класса i в сети, т.е. время обслуживания заявки класса i в сети $T_i = \sum_{j=1}^{n} b_{ij}$ $(i = 1,...,H)$, а $T_{\max} = \max\{T_1,...,T_H\}$ — максимальное время обслуживания в сети. Аналогично формуле (1) получим выражение для коэффициента загрузки узла j заявками класса i (i = 1, ..., H; j = 1, ... n):

$$\rho_{ij} = \begin{cases} b_{ij}\, m_i \,/\, B_{\max}, \text{ если } B_{\max} \ge T_{\max}; \\ b_{ij}\, m_i \,/\, T_{\max}, \text{ если } B_{\max} < T_{\max}, \end{cases} \quad (5)$$

тогда коэффициенты загрузок узлов $\rho_j \ (j = 1, \ ..., \ n)$ можно определить как $\rho_j = \sum_{i=1}^{H} \rho_{ij}$.

Расчет характеристик замкнутых сетей МО. Характеристики замкнутых сетей МО, рассмотренных в работе, рассчитываются следующим образом. Определяются коэффициенты передач α_{ij} для всех узлов сети. Для циклических моделей $\alpha_{ij} = 1$, для модели центрального обслуживания $\alpha_{ij} = p_{ij}$ (i = 1,..., H; j = 1,.... n). Коэффициенты загрузок приборов для каждого узла сети определяются по формулам (1), (2), (4), (5) в зависимости от типа модели. Тогда производительность однородной замкнутой сети определяется как $\lambda_0 = \rho_j K_j / \alpha_j b_j$ $(j = 1, \ ..., \ n)$. Из формулы Литтла [5] находится среднее время пребывания заявки в замкнутой сети $U = M / \lambda_0$. Для неоднородной замкнутой циклической модели $\lambda_0 = \sum_{i=1}^{H} \lambda_{0i}$, где $\lambda_{0i} = \rho_{ij} / b_{ij} \ (i = 1, \ ..., \ H; \ j = 1, \ ..., \ n)$. Тогда среднее время пребывания заявки класса i в сети $U_i = m_i / \lambda_{0i}$, а среднее время пребывания заявок объединенного потока $U = M / \lambda_0$.

Примеры. Рассмотрим замкнутую циклическую модель с двумя узлами, число приборов и время обслуживания в которых соответственно: K_1 = 3, b_1 = 2 с; K_2 = 5, b_2 = 4 с. В сети циркулирует M = 7 заявок, тогда, согласно (2), коэффициенты загрузок $\rho_1 = 7/9$ и $\rho_2 = 14/15$. Производительность сети $\lambda_0 = 7/6$ с$^{-1}$, тогда среднее время пребывания заявки в ней U = 6 с.

Рассчитаем характеристики циклической неоднородной модели с тремя узлами, в которой циркулируют заявки двух классов. Число заявок каждого класса и время обслуживания заявок в узлах соответственно: m_1=2; m_2=1; b_{11}=2 с; b_{21}=3 с; b_{12}=l с; b_{22}=2,5 с; b_{13}=l,5 с; b_{23}=4,5 с. Тогда, согласно (5), коэффициенты загрузок узлов заявками разных классов $\rho_{11} = 0,4$; $\rho_{21} = 0,3$; $\rho_{12} = 0,2$; $\rho_{22} = 0,25$; $\rho_{13} = 0,3$; $\rho_{23} = 0,45$, а коэффициенты загрузок узлов $\rho_1 = 0,7$; $\rho_2 = 0,45$; $\rho_3 = 0,75$ производительность сети λ_0=3 с$^{-1}$; среднее время пребывания заявок объединенного потока U = 10 с.

Заключение. Сравнительный анализ характеристик замкнутых сетей МО, рассчитанных в предположении об экспоненциальном и детерминированном характере обслуживания в узлах сети, показывает их существенное различие. В частности, результаты расчета характеристик обслуживания пакетов, полученные на сетевых моделях с детерминированным и экспоненциальным временем обслуживания в узлах, могут различаться на 30 % и более. Следует отметить, что различие в результатах существенно зависит от параметров модели. Наиболее сильно оно проявляется при больших загрузках узлов и их сбалансированности.

СПИСОК ЛИТЕРАТУРЫ

1. *Клейнрок Л.* Вычислительные системы с очередями. М.: Мир, 1979. 600 с.
2. *Шварц М.* Сети ЭВМ. Анализ и проектирование. М.: Радио и связь, 1981. 336 с.
3. *Алиев Т. И.* Стохастические модели информационно-вычислительных систем // Современные технологии. Сб. науч. статей. СПб: СПбГУ ИТМО, 2003. С. 6—17.
4. *Алиев Т. И., Никульский И. Е., Пяттаев В. О.* Моделирование ядра мультисервисной сети с относительной приоритезацией неоднородного трафика // Научно-технический вестник СПбГУ ИТМО. 2009. Вып. 4 (62). С. 88—96.
5. *Little J. D. C.* A proof for the queuing formula $L = l * w$ // Operations Research. 1961. Vol. 9, N 3. P. 383—387.

Сведения об авторе

Людмила Александровна Муравьева-Витковская — канд. техн. наук; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники;
E-mail: muravyeva-vitkovskaya@yandex.ru

Рекомендована кафедрой вычислительной техники

Поступила в редакцию 23.12.13 г.

УДК 004.89: 002.53

В. В. Соснин

ВРЕМЯ ОЖИДАНИЯ В НЕОДНОРОДНЫХ СИСТЕМАХ С ОЧЕРЕДЯМИ ПРИ ОБСЛУЖИВАНИИ ЗАЯВОК В ПОРЯДКЕ ПОСТУПЛЕНИЯ

Для системы массового обслуживания с заявками двух классов с помощью имитационного моделирования рассчитано среднее время ожидания в очереди при использовании бесприоритетной дисциплины обслуживания. Показаны условия, при которых различается среднее время ожидания заявок разных классов.

Ключевые слова: *неоднородная система массового обслуживания, среднее время ожидания в очереди, бесприоритетная дисциплина обслуживания.*

Введение. В теории массового обслуживания важное место занимает бесприоритетная дисциплина обслуживания (ДОБП). Традиционно считается [1], что при ДОБП качество обслуживания заявок разных классов одинаково (показателем качества считается среднее время ожидания заявки в очереди). В работе [2] показано, что в системе M/G/1 ДОБП ни один из классов не имеет преимуществ в качестве обслуживания, т.е. если значения времени ожидания в очереди заявок k классов суть случайные $W_1, W_2, ..., W_k$, то их математические ожидания равны: $M[W_1] = M[W_2] = ... = M[W_k]$. Однако цель настоящей работы — проверить, обладает ли этим свойством весь класс систем GI/GI/1 с ДОБП [3]. Для подтверждения корректности полученных автором результатов были проведены дополнительные исследования.

Поставленная задача решалась с помощью имитационного моделирования. Рассмотрим пример исследования системы массового обслуживания (СМО) GI/GI/1 с заявками двух классов (НК — низконагружающий, ВК — высоконагружающий класс), которые создают загрузки $\rho_{\text{ВК}} = 0,3$ и $\rho_{\text{НК}} = 0,03$. Время обслуживания заявок ВК и НК — случайная величина $B_{\text{ВК}}$ и $B_{\text{НК}}$ такая, что $M[B_{\text{ВК}}] = M[B_{\text{НК}}] = 10$ у.е. Время между приходом заявок ВК и НК — случайная величина $A_{\text{ВК}}$ и $A_{\text{НК}}$. Для моделирования $A_{\text{ВК}}$, $A_{\text{НК}}$, $B_{\text{ВК}}$ и $B_{\text{НК}}$ используется гамма-распределение, каждая из этих величин имеет фиксированное значение математического ожидания, а коэффициент вариации ν изменяется от 0 до 3 с шагом 0,1 так, что $\nu = \nu[A_{\text{ВК}}] = \nu[A_{\text{НК}}] = \nu[B_{\text{ВК}}] = \nu[B_{\text{НК}}]$. При проведении имитационных экспериментов измеряются время ожидания в очереди заявок каждого из классов ($W_{\text{ВК}}$ и $W_{\text{НК}}$), а также относительное различие их средних значений:

$$\varepsilon = \frac{M[W_{\text{ВК}}] - M[W_{\text{НК}}]}{M[W_{\text{НК}}]} \cdot 100\,\%.$$