

УДК 621.3

**АНАЛИЗ ХАРАКТЕРИСТИК СЕТЕЙ С ДИНАМИЧЕСКИМИ
ПРИОРИТЕТАМИ И НЕОГРАНИЧЕННЫМ БУФЕРОМ****Б.Г.ИСМАЙЛОВ, Ф.А.ДАДГАР***Бакинский Государственный Университет*
balemi@rambler.ru, Dadgar_52@yahoo.com

Анализируются распределенные компьютерные сети с динамическими приоритетами и неограниченным буфером, время пребывания требований в сети ограничено. В сети нарушение ограничения приводит к потере требований. Решены задачи минимизации потери требований за счет превентивного удаления части из них, не дожидаясь окончания обслуживания. Разработаны алгоритмы анализа характеристик компьютерных сетей с динамическими приоритетами и неограниченным буфером. Проведены эксперименты и получены численные результаты.

Ключевые слова: системы обслуживания, динамические приоритеты, оптимизация

Рассматривается процесс обслуживания распределенных компьютерных сетей с динамическими приоритетами требований и ограниченным буфером. Подобные обслуживания применяются в системах обслуживания (в станциях мобильных телефонных технологиях) при, которых абоненты находятся в движущимся транспортных средствах. Под обслуживанием понимается предварительная обработка информации о координатах и скорости транспортного средства. Время между моментами прихода требований в систему и окончания обслуживания ограничено некоторой величиной, причем, если это время превысит данную величину, требований будет потерянным. При этом информация, поступившая потребителю слишком поздно, оказывается бесполезной.

В [3] рассмотрены четыре модели организации обслуживания в многопроцессорных системах с ограниченным и неограниченным числом мест ожидания в очереди. Показано, что нижняя граница суммарных потерь достигается при конечном значении длины очереди. Используя эти результаты можно решить задачи определения оптимальных характери-

стик процесса обслуживания в распределенных компьютерных сетях и зачастую это в определенной мере усложняет операционную систему.

Однако разработки более развитых алгоритмов организации процесса обслуживания могут позволять уменьшить суммарные потери требований, сопровождающиеся ростом производительности сети обслуживания. Особенностью данного алгоритмического подхода является минимизация потери требований за счет превентивного удаления части из них, не дожидаясь окончания обслуживания.

Особенностью рассматриваемых сетей является минимизация потери требований за счет превентивного удаления части из них, не дожидаясь окончания обслуживания. С целью исследования функционирования такой сети её можно рассмотреть как модель системы массового обслуживания (СМО)[1,2].

При этом необходимо решить задачу оптимизации по параметру, определяющему характер превентивного удаления требований, что приводит к минимизации их потерь. Следует отметить, что в [3] доказано существование оптимального способа обслуживания требований в таких системах при различных вариантах взаимодействия алгоритмов очереди и обработки.

Постановка задачи

На вход очереди системы поступает поток требований. Из очереди требования, определённые по некоторому правилу, поступают равновероятно (полнодоступная система массового обслуживания) на любой исправный и свободный от обслуживания компьютер. После обслуживания требования покидают систему. Обслуживанию предполагается однородный входной поток требований с равными априорными приоритетами. В процессе обслуживания динамические приоритеты требований могут меняться в зависимости от ситуации. Время обслуживания всех требований распределено по одинаковому закону. В сети при функционировании возможно изменения ситуации в двух местах: в очереди и в обслуживании. Пребывание требования в системе состоит из двух фаз: ожидание и обслуживание. Общее время пребывания требований в системе не должно превышать τ_s^* .

На фазе ожидания с помощью алгоритма организации очереди требование может быть удалено из системы по некоторому правилу. Такие требования составляют потери первого рода:

$$P_1 = \frac{(\rho^N / N!)}{\left(\sum_{k=0}^{N-1} \rho^k / k! + (\rho^N / N!) \sum_{k=0}^m (\rho / N)^k\right)} \left(\frac{\rho}{N}\right)^m,$$

где $\rho = \lambda / \mu$ — приведенная интенсивность.

На фазе обслуживания с помощью алгоритма организации обслуживания (или сразу после её окончания) требование может быть удалено из

системы, если её пребывание в системе превысит τ_s^* . Эти требования составляют потери второго рода:

$$p_2 = \frac{(\rho^N / N!)}{\left(\sum_{k=0}^{N-1} \rho^k / k! + (\rho^N / N!) \sum_{k=0}^{m+1} (\rho / N)^k \right)} \left(\frac{\rho}{N} \right)^m ,$$

Особенностью рассматриваемых моделей СМО от традиционных заключается в наличии ограничения на время пребывания (ожидание и обслуживание) требований в системе, что и определяет целесообразность превентивного удаления из системы части требований не дожидаясь окончания обслуживания.

В общем случае, целью является выбор такого правила удаления ожидающих и обслуживаемых требований, при котором минимизация математическое ожидание значений суммарных потерь обоих родов, т.е.

$$M[P(p_1, p_2)] \rightarrow \min \quad (1)$$

$$\tau_s \leq \tau_s^* \quad (2)$$

$$L_q \leq m \quad (3)$$

где τ_s, L_q, τ_s^*, m – значения времени пребывания требования в системе, значения длина очереди, максимально допустимое значение времени пребывания требования в системе, максимально допустимое значение количество мест ожидания в очереди.

Оптимальная организация очереди и обслуживание позволяет получить максимальную эффективность системы за счёт удаления из системы до начала или во время обслуживания тех требований, которые не выполняются условиям (2).

Следует отметить, что учитывая сложность реальных систем, нахождение оптимальных параметров для удаления требований из системы аналитическими методами обычно не представляется возможным. Поэтому для этого используются методы имитационного моделирования.

Однако необходимость оптимизации влечёт за собой проблему прямого перебора значений оптимизируемых параметров, которая даже в однопараметрической задаче делает метод имитационного моделирования практически непригодным.

В работе [3] данная проблема решается путём ограничения перебора, т.е. показывается либо существование оптимального значения параметра и вычисляются границы интервала, на котором его следует искать, либо определяется граничное значение параметра, выход за которое не приводит к практически значимому улучшению качества работы системы.

Для организации обслуживания в системе в данной работе предлагаются различные варианты алгоритма взаимодействия организации оче-

реди и обслуживании, при котором могут быть использованы разработки программного обеспечения подобных систем различного назначения.

При этом в соответствии [3] предлагается четыре вариантов алгоритма взаимодействия организации очереди и обслуживании модели СМО.

В первом варианте алгоритма очередь не ограничена $L_q = \infty$, и возможно измерения времени пребывания требования в системе. Система имеет возможность измерения времени пребывания требования в системе, а при обслуживании не имеет. Поэтому в алгоритме устанавливается ограничение на время ожидания в очереди $\tau_q \leq \tau_q^*$, а факт превышения допустимого времени пребывания требования в системе $\tau_s \leq \tau_s^*$ в алгоритме, так же как и в первом варианте, может быть выявлен после окончания обслуживания.

Во втором варианте алгоритма очередь не ограничена $L_q = \infty$ и в обоих фазах имеется возможность измерения времени пребывания требования в системе. Поэтому факт превышения допустимого времени пребывания требования в системе $\tau_s \leq \tau_s^*$ выявляется на любой из двух фаз, т.е. в момент его возникновения.

Следует отметить, что рассматриваемые варианты алгоритма взаимодействия организации очереди и обслуживании соответствуют различным техническим системам и они описывают различные с точки зрения эффективности возможности и взаимодействия организации очереди и обслуживании. Отметим, что если обслуживания требований и проверка на выполнение условия (2) производятся в разных компьютерах (первый вариант алгоритма), то не имеется возможность прекратить обслуживание требования, даже если нарушено условие (2). Такая ситуация имеет место в первом варианте алгоритма (в обслуживании не имеется возможность измерения времени τ_s).

В случае, когда обслуживания требований и проверка условия $\tau_s \leq \tau_s^*$ производится в одном компьютере сети, требование, задержавшееся и системе дольше времени τ_s^* , удаляется из нее, не дожидаясь окончания обслуживания. Такая ситуация имеет место во втором варианте (в обслуживании имеется возможность измерения времени τ_s .)

Эти варианты алгоритма имеют неограниченную по числу мест очередь $L_q = \infty$, а удаление из очереди в них производится по критерию времени пребывания в ней $\tau_s \leq \tau_s^*$ (в обслуживании имеется возможность

измерении времени τ_s .)

Замечание 1. В [3] показывается, что нахождение границ множества перебора для определения оптимального критерия удаления из очереди для этих двух вариантов алгоритма сложности не представляет.

Замечание 2. Вариант, который позволяет исключить из системы «бесперспективные» требования раньше, чем другой, естественно является более эффективным и «интеллектуальным». Кроме того, следует отметить, что для каждого варианта взаимодействия организации очереди и обслуживания в зависимости от уровня деградации системы правила удаления должны корректироваться.

Анализ алгоритмов модели обслуживания сети

В первом варианте алгоритма модели обслуживания очередь не ограничено, но время ожидания в очереди не должно превышать t^* . При $t = t^*$ заявка удаляется из очереди и из системы вообще (потери первого рода). Очевидно, что для любых входного потока и распределения времени обслуживания, а также для произвольной дисциплины выбора требований на обслуживание существует по меньшей мере одно значение t^* при котором суммарные потери минимальны причём $0 < t^* \leq T^*$

В втором варианте алгоритма модели обслуживания очередь организуется аналогично второму варианту модели, а обслуживание на компьютере аналогично третьему варианту модели.

Для экспоненциального времени обслуживания значения L_q могут быть определены следующим образом [4]:

$$L_q = P\rho / (N - \rho)$$

В этом формуле в зависимости от характера объекта применения системы можно допускать следующие аппроксимации:

$$\begin{aligned} \text{при } \rho \ll 1 \quad L_q &\rightarrow \rho^{N+1} / N^2; \\ \text{при } \lambda / \mu N \rightarrow 1 \quad L_q &\rightarrow \rho / (N - \rho); \end{aligned}$$

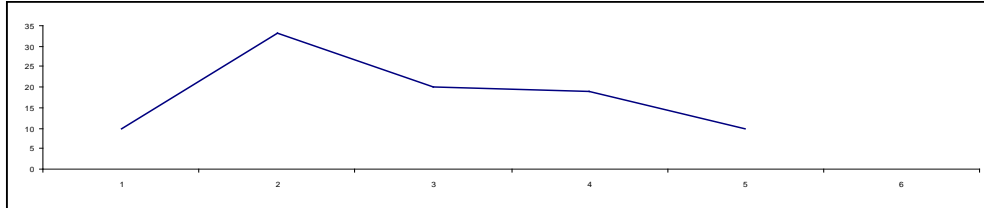
При известном L_q можно определить также время ожидания требования в очереди τ_q , время пребывания требования в системе, ожидаемое количество z требования в системе L_s :

$$\tau_q = L_q / \lambda; \quad \tau_s = L_s / \lambda; \quad \lambda / \mu N < 1; \quad L_s = L_q + \rho$$

Для $m = 10$ $\tau_s^* = 100$ и ($\lambda = 0.001 - 0.005$, $\mu = 0,010 - 0.022$), $\rho = 0.1 - 0,9$ проведены объемные вычислительные эксперименты и получены численные результаты. На основе полученных результатов построены следующие

щие зависимости, которые показаны на рис.1.-5. $L_q = f(\rho)$

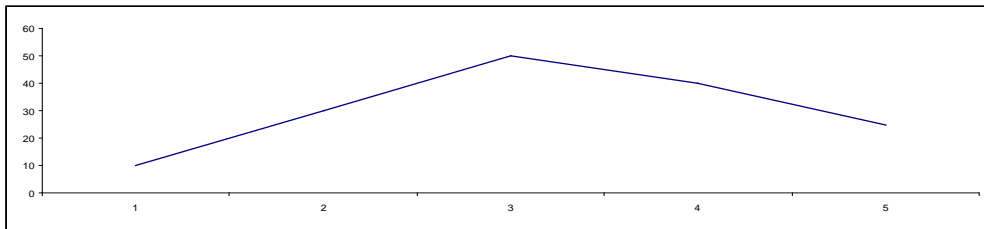
L_q



ρ

Рис.1.Зависимость $L_q = f(\rho)$

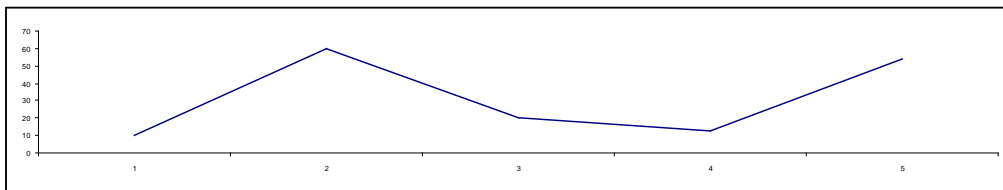
L_s



ρ

Рис.2.Зависимость $L_s = f(\rho)$

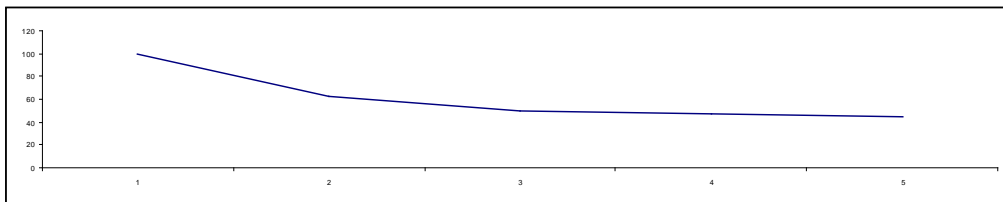
τ_q



λ

Рис.3.Зависимость $\tau_q = f(\lambda)$

τ_s



λ

Рис.4.Зависимость $\tau_s = f(\lambda)$

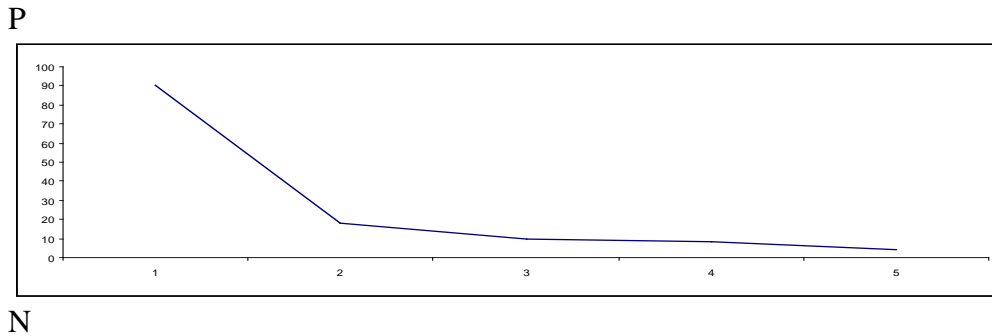


Рис. 5. Зависимость $P = f(N)$

Анализ этих зависимостей показывает, что значение длина очереди и количество требования в системе уменьшаются при $\rho = 0.3$.

Значение время ожидания требования в очереди τ_a , время пребывания требования в системе τ_s достигают свою минимальную значений при $\lambda = 0.004$ и следовательно минимальное значение суммарные потери информации достигается при $N=2$.

Следует отметить, что эти результаты могут быть использованы при построении распределенных сетей обслуживания различного назначения.

Заключение

В данной работе разработанные алгоритмы организации взаимодействия очереди и обслуживания некоторых моделей СМО, для произвольного числа компьютеров N (любого уровня деградации), позволяет разработчику системы определять оптимальное значение длины очереди и правила удаления «бесперспективных» с точки зрения масштаба реального времени требований. Рассмотренная модель охватывает достаточно широкий класс современных сетей различного назначения. Показано, что для этого класса всегда целесообразно принудительно удалять часть ожидающих требований из очереди, таким образом уменьшая нагрузку на обслуживающие приборы, а в конечном счете минимизируя суммарные потери обоих родов. Применение более развитых механизмов программно-технических обеспечений позволяет уменьшить суммарные потери требований.

Таким образом, для всех вариантов взаимодействия программно-технических обеспечений всегда существуют параметры, минимизирующие суммарные потери в системе. Целесообразность такой оптимизации зависит от конкретной ситуации - некоторое усложнение операционной системы компенсируется ростом производительности.

ЛИТЕРАТУРА

1. Клейнрок Л. Теория массового обслуживания. М.: Машиностроение, 1979.
2. Клейнрок Л. Коммуникационные сети. М.: Наука, 1970, с.256.
3. Мамедли Э. М., Слипченко А.И., Хусидман В.Б. Модели организации диспетчеризации в много процессорных вычислительных системах реального времени. Автоматика и Телемеханика, М.:1991с.167-175.
4. Kim C.S., Melikov A.Z., Ponomarenko L.A. Approximation Method for Performance Analysis of Queuing System with Multimedia Traffics, // Applied and Comput. Math. 2007, v.6, No.2, pp.218-226.
5. Lee Y., Choi B. D. Queueing System with Multiple Delay and Loss Priorities for ATM Networks // Information Systems. 2001, v 138, pp.7-29.
6. Pitts J.M., Schormans J.A. Introduction to ATM Design and Performance. N.Y.: John Wiley & Sons., 1997.
7. Chao H.J., Peckan I.H. Queue Management with Multiple Delay and Loss Priorities for ATM Switches // Proc. ICC'94, 1994, pp.1184-1189.

QEYRİ-MƏHDUD BUFERLİ VƏ DİNAMİK PRİORİTETLİ ŞƏBƏKƏLƏRİN XARAKTERİSTİKALARININ TƏHLİLİ

B.Q.İSMAYILOV, F.A.DADİGƏR

XÜLASƏ

Qeyri-məhdud buferli və dinamik prioritetli paylanmış kompüter şəbəkələri təhlil edilir, şəbəkədə tələblərin gəlmə vaxtı məhduddur. Şəbəkədə məhdudiyətin pozulması tələblərin itməsinə səbəb olur. Tələblərin itkisinin minimallaşdırılması məsələsi onların qabaqlayıcı tədbir kimi xidməti gözləmədən sistemi tərk etməsi hesabına həll edilmişdir. Qeyri-məhdud buferli və dinamik prioritetli kompüter şəbəkələrinin xarakteristikalarının təhlil alqoritmi işlənməmişdir. Eksperiment aparılmış və ədədi nəticələr alınmışdır.

Açar sözlər: *xidmət sistemi, dinamik prioritetlər, optimallaşdırma*

THE ANALYSIS OF THE CHARACTERISTICS OF NETWORKS WITH DYNAMIC PRIORITIES AND UNLIMITED BUFFERS

B.G.ISMAYILOV, F.A.DADGAR

SUMMARY

The paper deals with distributed computer networks with dynamic priorities and unlimited buffers. Online constraint violation results in loss of orders. The paper solves the problem of minimizing the loss of orders due to the preventive removal of some of them, without waiting for the end of service. The analysis algorithm of the characteristics of the computer networks with dynamic priorities and unlimited buffers is worked out. As a result of the carried out experiments, numerical results have been obtained .

Keywords: *service systems , dynamic priorities , optimization*

Поступила в редакцию: 20.06.2014 г.

Подписано к печати: 26.11.2014 г.