

МОДЕЛИ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ С ДИНАМИЧЕСКИМ РАСПРЕДЕЛЕНИЕМ ПРОПУСКНОЙ СПОСОБНОСТИ ИСХОДЯЩЕГО КАНАЛА СВЯЗИ

Лемешко А.В., Симоненко А.В., Ватти Махмуд
Харьковский национальный университет радиозлектроники
61166, Харьков, пр. Ленина, каф. ТКС, тел. (057) 702-13-20
E-mail: avlem@ukr.net

In the given work is proposed a mathematical model of queue management on the basis of which the task of packages allocation on the queue is represented in the form of optimization task, which belongs to the class of triplane transporting tasks of linear programming

Введение

Превалирующей тенденцией в развитии средств обеспечения качества обслуживания является (Quality of Service, QoS) переход к динамическим стратегиям управления сетевыми ресурсами – буферным пространством сетевых узлов, пропускной способностью каналов связи и сетевым трафиком. Это, прежде всего, касается протоколов динамической маршрутизации, пришедших на смену статическим планам распределения пакетов. С целью повышения гибкости управления сетевыми ресурсами динамический характер становится характерным и для других средств, в частности и для механизмов управления очередями, от эффективности которых во многом зависят значения таких показателей качества обслуживания, как средняя задержка, джиттер и потери пакетов.

В этой связи, на смену статическим и квазидинамическим схемам, положенных в основу существующих механизмов управления очередями [1], должны прийти математические модели динамического управления очередями, ориентированных на получение адаптивных решений по отношению к загруженности узла и телекоммуникационной сети (ТКС) в целом, а также к требуемым значениям показателей QoS.

Модель динамического управления очередями на узлах «тупой» сети

Под «тупой» сетью (Stupid network) понимается сеть, в которой большинство «интеллектуальных» функций управления вынесены на ее периметр, т.е. они сосредоточены на приграничных узлах. К числу таких функций, например, относится назначение приоритетов пакетам трафиков различных приложений, маршрутизация «от источника» и др. В подобного рода сетях транзитные узлы сети реализуют лишь самые простейшие функции коммутации, т.е. функции пересылки с входного порта на выходной поступающих на узел пакетов в соответствии с содержанием маршрутной таблицы (в IP-сетях) или таблицы коммутации меток (в MPLS-сети).

Для формализации процесса управления очередями с динамическим распределением пропускной способности исходящего канала связи (тракта передачи) в рамках «тупой» сети предлагается математическая модель, в которой считается фиксированным количество классов (приоритетов) обслуживаемых трафиков M . Например, в случае использования IP Precedence при назначении приоритетов поддерживается маркировка до восьми классов трафика ($M = 8$). При использовании поля кода дифференцированной услуги (Differentiated Services Code Point, DSCP), которое является расширением 3-битового поля IP Precedence, задействуется уже 6 бит байта типа обслуживания (Type of Service, TOS) из заголовка пакета IPv4 или байта класса трафика (Traffic Class) пакета в IPv6. В последнем случае число классов трафика равно шестидесяти четырем ($M = 64$).

Кроме того, условимся, что число очередей на сетевом узле также фиксировано (N) или назначается административно, которое, в общем случае не превышает количества классов поддерживаемых сетью трафиков. Например, для алгоритма заказного (обычного) обслуживания (Custom Queuing, CQ) число очередей равно шестнадцати (плюс нулевая – системная очередь), а в алгоритме приоритетного обслуживания (Priority Queuing, PQ) сетевой администратор может выделить до четырех очередей для сетевого трафика высокого, среднего, нормального и низкого приоритета ($N = 4$).

Ввиду потокового характера современного, преимущественно мультимедийного трафика при решении задач обеспечения QoS необходимо оперировать с интенсивностью трафика, поэтому обозначим через a_i ($i = \overline{1, M}$) – интенсивность трафика i -го класса, поступающего на обслуживание сетевым узлом. Кроме того, пусть b_j ($j = \overline{1, N}$) – часть пропускной способности исходящего тракта передачи, которая закреплена за j -й очередью ($j = \overline{1, N}$), что характерно, например, для алгоритма CBWFQ. При этом необходимо выполнить условие

$$\sum_{j=1}^N b_j = b, \quad (1)$$

где b – пропускная способность исходящего тракта передачи.

С целью предотвращения перегрузки узла и сети в целом необходимо обеспечить выполнение условий

$$\sum_{i=1}^M a_i \leq b. \quad (2)$$

Выполнение условия (2) обеспечивается, прежде всего, за счет эффективной маршрутизации трафика в сети, в противном случае – путем использования механизмов предотвращения ограничения длины очереди (RED, WRED).

Если в известных алгоритмах обслуживания очередей порядок распределения пакетов того или иного трафика в очереди сетевого узла, как правило, строго фиксирован, то придать динамический характер этому процессу удалось в предлагаемой модели за счет введения переменной x_{ij} , под которой подразумевалась часть i -го трафика, который будет направлен для обслуживания в j -ю очередь. Согласно физическому смыслу переменной x_{ij} имеют место следующие дополнительные условия:

$$x_{ij} \geq 0 \quad (i = \overline{1, M}, j = \overline{1, N}), \quad (3)$$

$$\sum_{j=1}^N x_{ij} = a_i \quad (i = \overline{1, M}), \quad (4)$$

$$\sum_{i=1}^M x_{ij} \leq b_j \quad (j = \overline{1, N}). \quad (5)$$

Выполнение условия (4) гарантирует обслуживание всех пакетов трафика, поступающих на рассматриваемый сетевой узел. Условия (5) вводятся для предотвращения перегрузки очередей сетевого узла в процессе управления. В отличие от ранее известных решений [2] в качестве искомого вектора выберем вектор

$$X = \begin{bmatrix} x_{ij} \\ - \\ b_j \end{bmatrix} \quad (i = \overline{1, M}, j = \overline{1, N}), \quad (6)$$

в ходе расчета которого удастся обеспечить согласованность в решении задач обслуживания очередей и динамического распределения за ними пропускной способности исходящего канала связи (тракта передачи). В связи с тем, что, в общем случае, выбор x_{ij} и b_j в рамках ограничений (1), (3)-(5) можно произвести множеством случаев, то целесообразно задачу, связанную с расчетом искомых переменных, сформулировать в виде оптимизационной со следующей (в дальнейшем минимизируемой) целевой функцией:

$$T(x) = c^t X, \quad (7)$$

в которой координаты вектора весовых коэффициентов

$$c = \begin{bmatrix} c_{ij} \\ \dots \\ c_j \end{bmatrix} \quad (8)$$

характеризуют условную стоимость (c_{ij}) использования пакетами i -го трафика ресурсов j -й очереди, а также стоимость (c_j) выделения j -й очереди того или иного объема пропускной способности исходящего канала связи. С помощью подбора численных значений координат c_{ij} и c_j можно добиться реализации в рамках предложенной модели различных типов очередей – заказных, приоритетных, справедливых и др.

Предложенная математическая модель управления очередями (1)-(8) представлена в виде задачи линейного программирования, для решения которой известны точные методы решения, например симплекс метод, метод ветвей и границ и др.

Модель динамического управления очередями на узлах активной сети

Для учета особенностей обслуживания пакетов на узлах активных сетей предложенная выше модель потребует своего некоторого развития и совершенствования. Наличие K параллельно функционирующих активных процессоров (АП) в структуре сетевого узла [3, 4], отвечающих за выполнение предварительных функций обработки пакетов на транзитном сетевом узле, предполагает введение новых ограничений (условий) в содержание модели. Тогда под c_k ($k = \overline{1, K}$) обозначим номинальную производительность k -го АП. С целью предотвращения перегрузки активного узла условия (2) дополним выражениями вида

$$b \leq \sum_{k=1}^K c_k \quad \text{и} \quad \sum_{i=1}^M a_i \leq \sum_{k=1}^K c_k. \quad (9)$$

Выполнение условий (9) необходимо удовлетворить на этапе проектирования сети, при выборе производительности сетевых узлов и АП в частности.

Расширение области применения предлагаемой модели влечет за собой расширение и функциональной нагрузки искомого вектора (6)

$$X = \begin{bmatrix} x_{ijk} \\ \dots \\ b_j \end{bmatrix}, \quad (10)$$

координаты которого x_{ijk} характеризуют теперь интенсивность i -го трафика, который будет направлен для обслуживания в j -ю очередь через k -й активный процессор. Согласно физическому смыслу переменной x_{ijk} по аналогии с выражениями (3)-(5) имеют место следующие дополнительные условия:

$$x_{ijk} \geq 0 \quad (i = \overline{1, M}, j = \overline{1, N}, k = \overline{1, K}), \quad (11)$$

$$\sum_{j=1}^N \sum_{k=1}^K x_{ijk} = a_i \quad (i = \overline{1, M}), \quad (12)$$

$$\sum_{i=1}^M \sum_{k=1}^K x_{ijk} \leq b_j \quad (j = \overline{1, N}) \quad \text{и} \quad \sum_{i=1}^M \sum_{j=1}^N x_{ijk} \leq c_k \quad (k = \overline{1, K}). \quad (13)$$

Целевая функция, подлежащая минимизации в ходе решения задачи управления очередями на активном узле, будет иметь прежний вид (7), но координаты вектора

$$c = \begin{bmatrix} c_{ijk} \\ \dots \\ c_j \end{bmatrix} \quad (14)$$

теперь будут численно характеризовать условную стоимость использования пакетами i -го трафика ресурсов j -й очереди и k -го АП (c_{ijk}).

Предложенная математическая модель управления очередями на активном узле (9)-(14) представлена в виде многоиндексной задачи линейного программирования, а именно в виде трипланарной транспортной задачи [5], для решения которой известны точные и приближенные методы решения, например, метод потенциалов и метод минимального элемента в сечении.

Выводы

Таким образом, рассмотрен подход к динамическому управлению очередями (буферным ресурсом) в ТКС, поддерживающей или не поддерживающей технологию активных сетей. Подход основан на использовании предложенных математических моделей управления очередями, в рамках которых в отличие от ранее известных решений обеспечивается согласованное решение задач распределения пакетов и пропускной способности исходящих трактов передачи по очередям сетевого узла. Динамический характер решений достигается за счет периодического или аperiodического перерасчета порядка использования буферного ресурса сетевых узлов адаптивно, т.е. в зависимости от изменения требований к качеству обслуживания, интенсивности поступающего на узел информационного трафика, количества поддерживаемых очередей, пропускной способности активных процессоров или исходящих трактов передачи.

Развитие данного подхода видится в расширении функциональности предложенной модели в сторону учета разнородности используемых активных процессоров, ориентированных на выполнение определенных (специализированных) функций обработки пакетов на сетевых узлах. Кроме того, необходимо предусмотреть учет в явном виде требований к качеству обслуживания в рамках вновь разрабатываемых математических моделей управления очередями. Важное направление совершенствования предложенных моделей состоит в адаптивном определении числа поддерживаемых очередей в ходе решения задач управления в зависимости от числа и интенсивности поступающего на обслуживание пользовательского трафика.

Литература:

1. Вегенша Ш. Качество обслуживания в сетях IP: Пер. с англ. – М.: Издательский дом «Вильямс», 2003. – 386 с.
2. Лемешко А.В., Ватти М., Симоненко А.В. Управление очередями на узлах активной сети // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2007. – Вып. 151. – С. 92-97.
3. Tennenhouse D. L., Smith J. M., Sincoskie W. D., Wetherall D. J., Minden G. J. A Survey of Active Network Research // IEEE Communications Magazine. 1997. – Vol. 35, №1. – P. 80-86.
4. Иванов П. Активные сети // Сети. 1999. – №10. – С. 14-24.
5. Раскин Л.Г., Кириченко И.О. Многоиндексные задачи линейного программирования. – М.: Радио и связь, 1982. – 240 с.