

МОДЕЛЬ ДИНАМИЧЕСКОГО УПРАВЛЕНИЯ ОЧЕРЕДЯМИ И ПРОПУСКНОЙ СПОСОБНОСТЬЮ КАНАЛА СВЯЗИ НА МАРШРУТИЗАТОРАХ МУЛЬТИСЕРВИСНОЙ СЕТИ

Введение

Задачи управления очередями в существующих телекоммуникационных технологиях транспортных мультисервисных сетей – IP (Internet Protocol), ATM (Asynchronous Transfer Mode) и MPLS (MultiProtocol Label Switching) являются ключевыми при обеспечении необходимого уровня качества обслуживания (Quality of Service, QoS) запросов пользователей. В условиях гетерогенности и территориальной распределенности современных телекоммуникационных систем (ТКС) ввиду хорошей масштабируемости именно архитектурная модель DiffServ (Differentiated Services), основанная на приоритетной обработке пакетов на сетевых узлах-маршрутизаторах, получила наибольшее распространение. К сожалению, решение многих сетевых задач в рамках модели DiffServ все еще носит статический характер и нередко связано с административным вмешательством человека. Как правило, приоритеты пакетам присваиваются на приграничных узлах и в дальнейшем практически не изменяются, а количество очередей и закрепленная за ними пропускная способность исходящего канала связи может определяться вручную администратором сети в ходе конфигурирования сетевого оборудования [1, 2]. Инерционность и низкая согласованность в решениях задач формирования очередей, их обслуживания и распределения канального ресурса между очередями влекут за собой необходимость пересмотра заложенных в соответствующие механизмы и протоколы схем и математических моделей с целью реализации динамических стратегий управления, способных адаптивно подстраиваться под изменения состояния сети и отдельных сетевых узлов – интенсивности и приоритета поступающего на обслуживание трафика, загруженности канальных и буферных ресурсов ТКС.

Модель динамического управления очередями и пропускной способностью канала связи на маршрутизаторах мультисервисной сети

С целью формализации процесса управления очередями с динамическим распределением пропускной способности (ПС) исходящего канала связи (тракта передачи) предлагается математическая модель, в которой количество приоритетов (классов) обслуживаемых трафиков M определяется особенностью поддерживаемого на приграничных узлах метода маркировки (IP DSCP, IP Precedence, MPLS EXP bits, ATM CLP-bit, Frame-Relay DE-bit, IEEE 802.1Q/p user-priority bits и др.). Общее число приоритетов определяется числом выделенных для этой цели битов в заголовке пакета той или иной технологии. Например, в случае использования трехбитового поля IP Precedence при назначении приоритетов поддерживается маркировка до восьми классов трафика ($M = 8$), а при использовании шести битов кода дифференцированной услуги (Differentiated Services Code Point, DSCP) число классов трафика равно шестидесяти четырем ($M = 64$) (рис. 1).

Кроме того, условимся, что число очередей на сетевом узле также фиксировано (N) или назначается административно и, в общем случае не превышает количества классов поддерживаемых сетью трафиков. Например [1], для алгоритма заказного (обычного) обслуживания (Custom Queuing, CQ) общее число очередей равно семнадцати ($N = 17$), шестнадцать обычных очередей плюс нулевая – системная очередь, в алгоритме приоритетного обслуживания (Priority Queuing, PQ) может быть выделено до четырех очередей для сетевого трафика высокого, среднего, нормального и низкого приоритета ($N = 4$), а в алгоритме взвешенного справедливого обслуживания (Weighted Fair Queuing, WFQ) по умолчанию число очередей равно 256.



Рис.1

Ввиду потокового характера сетевого, преимущественно мультимедийного трафика, при решении задач обеспечения QoS в мультисервисных сетях необходимо оперировать не столько с отдельными пакетами, сколько с интенсивностью трафика, а сами математические модели управления очередями должны носить потоковый характер, а не ориентироваться лишь на описание схем обслуживания отдельных пакетов, как это происходит в существующих алгоритмах PQ, FQ, WFQ и др [2]. 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 динамически выделяется для передачи пакетов из j -й очереди ($j = \overline{1, N}$). При этом необходимо выполнить условие

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

где b – пропускная способность исходящего канала связи.

С целью предотвращения перегрузки узла и сети в целом необходимо обеспечить выполнение следующих 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 алгоритмов произвольного раннего обнаружения перегрузки и ограничения длины очереди RED (Random Early Detection), WRED (Weighted RED).

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

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

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

согласно выражению (2)

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

Выполнение условия (4) гарантирует отсутствие потерь пакетов на сетевых узлах, наносящая на обслуживание всех пакетов трафика, поступающих на рассматриваемый сетевой узел. Условия (5) вводятся для предотвращения перегрузки очередей сетевого узла в процессе управления.

В отличие от известных решений [3, 4] в качестве искомого вектора выберем вектор

$$\vec{x} = \begin{bmatrix} x_{1,1} \\ x_{1,2} \\ \vdots \\ x_{M,N} \end{bmatrix} \quad (i = \overline{1, M}; j = \overline{1, N}), \quad (6)$$

расчет которого целесообразно производить в ходе решения оптимизационной задачи по минимизации целевой функции

$$T(x) = \vec{c}' \vec{x}, \quad (7)$$

характеризующей суммарную стоимость обслуживания пакетов на сетевом узле при том, что

$$\vec{c} = \begin{bmatrix} c_{1,1} \\ c_{1,2} \\ \vdots \\ c_{M,N} \end{bmatrix} \quad - \text{ это вектор удельных стоимостей обслуживания пакетов, а } c_{ij} \text{ - удельная стоимость}$$

обслуживания пакетов i -го класса (приоритета) j -й очередью. Варьируя численными значениями координат вектора \vec{c} , можно смоделировать ту или иную дисциплину приоритетного обслуживания пакетов [3, 4].

Оптимизационная задача, связанная с минимизацией линейной целевой функции (7) при наличии линейных ограничений (2)-(5), решается известными методами линейного программирования. Причем, по результатам решения сформулированной задачи можно рассчитать величину пропускной способности исходящего от рассматриваемого сетевого узла канала связи, которая динамически выделяется для передачи пакетов из j -й очереди:

$$b_j = \sum_{i=1}^M x_{ij} \quad (j = \overline{1, N}). \quad (8)$$

Анализ предложенной модели управления очередями

В работе [3] величина b_j считается заранее заданным параметром, а в работе [4] величины b_j ($j = \overline{1, N}$), наряду с x_{ij} , также являлись координатами искомого вектора \vec{x} . В предлагаемой модели сохраняется согласованность в решении задач обслуживания очередей (x_{ij}) и динамического распределения пропускной способности исходящего канала связи (b_j), но снижается размерность оптимизационной задачи. При практически идентичном получаемом результате размерность вектора искомых параметров в предлагаемой модели (модель 1) составляет $(N \cdot M)$, а та же размерность в модели (модель 2) [4] оставляет $(N \cdot M + N)$. Выигрыш по размерности решаемой оптимизационной задачи для различных моделей управления очередями составляет $(1/M) \cdot 100\%$. Таким образом, при использовании IP Precedence выигрыш составит минимум 12.5%, а с ростом числа классов обслуживания выигрыш уменьшается и для IP DSCP составит около 2%.

Это граничные значения, когда допускается, что пакеты различных классов трафика могут обслуживаться всеми очередями без исключения. На практике каждой очереди соответствует, как правило, один класс трафика. Допускается, в случае «простоя» высокоприоритет-

ной очереди, обслуживание ею менее приоритетного трафика. На рис 2 представлены соотношение размерности модели 2 к размерности модели 1 для случаев обслуживания определенного класса трафика только одной очередью (рис 2, а), одновременно двумя (рис. 2, б) и тремя (рис. 2, в) очередями.

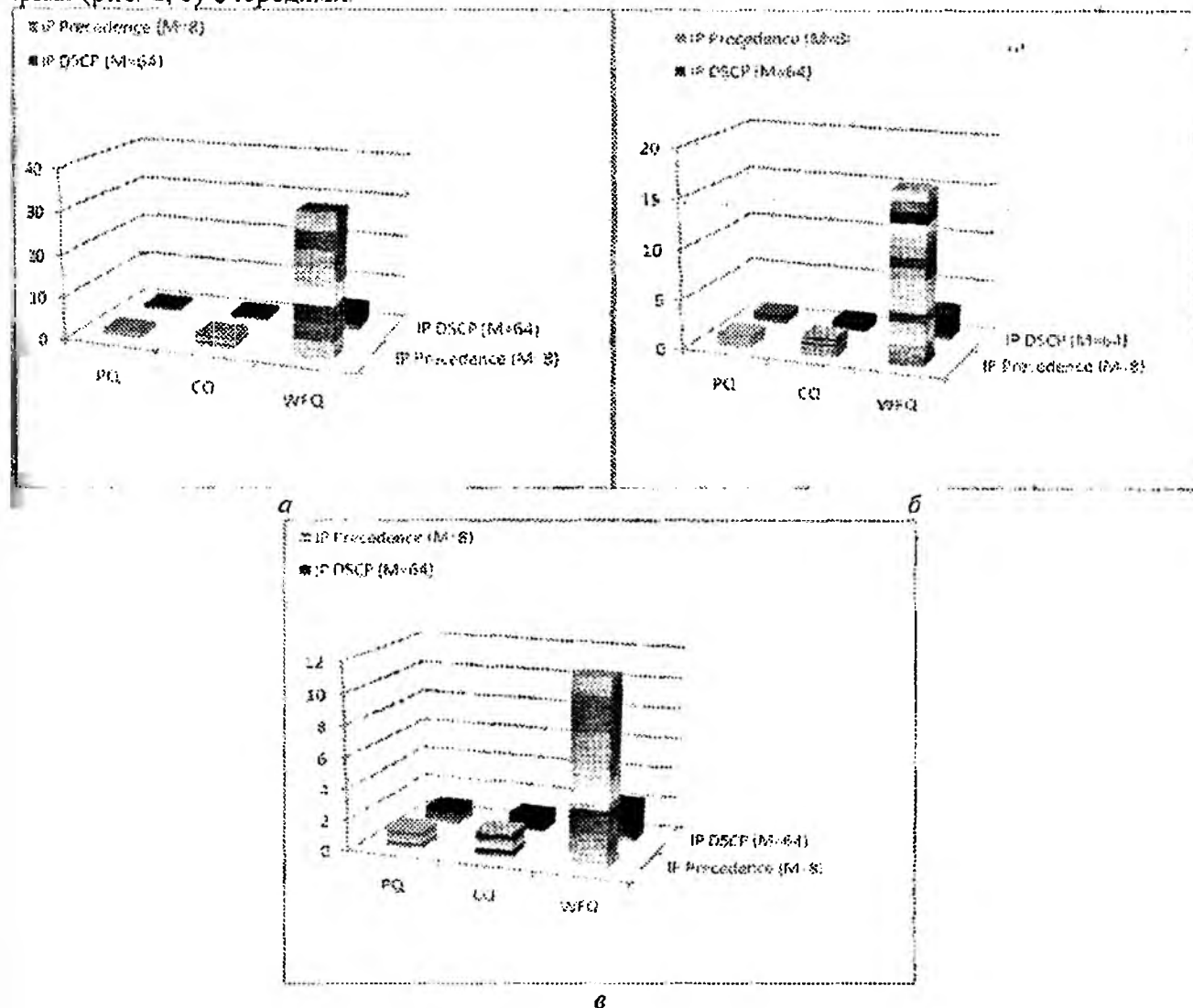


Рис 2

Как видно из полученных зависимостей (рис 2), преимущество предложенной модели (1) – (8) по сравнению с моделью [4] увеличивается, во-первых, с ростом общего числа поддерживаемых очередей в рамках того или иного алгоритма (PQ, CQ, WFQ), а во-вторых, со снижением числа классов трафика, а также количества очередей, с помощью которых разрешено обслуживание пакетов одного и того же класса трафика.

Выводы

Рассмотрен подход к динамическому и, что немаловажно, согласованному управлению очередями (буферным ресурсом) и пропускной способностью канала связи (канальный ресурс) в ТКС. Динамический характер решений достигается за счет адаптивного перерасчета порядка использования буферного ресурса сетевых узлов в зависимости от изменения требований к качеству обслуживания, интенсивности поступающего на узел информационного трафика, количества поддерживаемых очередей, пропускной способности исходящих трактов передачи. Предложенная модель, сохраняя высокую функциональность, определяет необходимость решения оптимизационной задачи, которая имеет менее высокую размерность, чем ранее известные подобные модели.

Следующим шагом на пути к совершенствованию предложенных решений является создание моделей, в рамках которых обеспечивается учет в явном виде требований к качеству обслуживания, адаптивно определяется число поддерживаемых очередей в зависимости от загрузки сетевого узла.

Список литературы: 1. *Справочник по телекоммуникационным технологиям*: Пер. с англ. М.: Издательский дом «Вильямс», 2004. 640 с. 2. *Вегенна Ш* Качество обслуживания в сетях IP: Пер. с англ. М.: Издательский дом «Вильямс», 2003. 386 с. 3. *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. 4. *Лемешко А. В., Симоненко А. В., Ватти Махмуд.* Поточковая модель управления очередями с динамическим распределением пропускной способности исходящего канала связи // *Наук. записки УНДІЗ.* 2008. №3(5). С. 34-39.

*Харьковский национальный
университет радиоэлектроники*

Поступила в редколлегию 16.11.2008