

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДИНАМИЧЕСКОГО УПРАВЛЕНИЯ КАНАЛЬНЫМ И БУФЕРНЫМ РЕСУРСОМ НА УЗЛАХ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ**Введение**

Как с точки зрения теории, так и с позиции практики известно, что количественные значения основных показателей качества обслуживания (Quality of Service, QoS) во многом определяются эффективностью решения задач по управлению очередями на узлах телекоммуникационной сети (ТКС) [1, 2]. Именно на IP-маршрутизаторах, АТМ-коммутаторах и на маршрутизирующих коммутаторах LER/LSR в MPLS-сетях решаются задачи формирования очередей в соответствии с приоритетом поступающих на обслуживание пакетов, обслуживание очередей с выдачей пакетов в канал связи, распределение пропускной способности канала связи (КС) между очередями, т.е. пакетами того или иного приоритета.

Характерной чертой существующих решений по управлению буферным (очередями) и канальным (пропускной способностью канала связи) ресурсом, реализованных в рамках как простейших механизмов PQ, CQ, FQ, так и более сложных – CBQ, WFQ, CBWFQ и др., является статическая стратегия их распределения [1, 2]. Стоит отметить, что протоколы маршрутизации, отвечающие за распределение трафика в сети, уже давно классифицируются как протоколы *динамической* маршрутизации, позволяя периодически или по требованию модифицировать маршруты следования пакетов того или иного трафика в зависимости от изменения топологии ТКС и других ее характеристик. Соответствующие таймеры перерасчета маршрутных таблиц, для сравнения, составляют десятки секунд: для протокола RIP – 30 с, для протокола IGRP – 90 с. Однако порядок распределения буферного пространства и канальной емкости по-прежнему преимущественно устанавливается административно, т.е. вне реального времени, хотя динамика изменения состояния отдельного сетевого узла несравненно выше, чем динамика загрузки ТКС.

В связи с этим актуальной представляется задача о придании динамического характера решениям по управлению канальным и буферным ресурсом в зависимости от характеристик поступающего на узел трафика, величины доступного ресурса, требуемых показателей качества обслуживания и т.д. Решение данной задачи связано, прежде всего, с пересмотром самих принципов распределения ресурсов путем разработки новых математических моделей управления очередями и канальной емкостью, которые должны прийти на смену достаточно примитивным и преимущественно эвристическим схемам обслуживания пакетов на сетевых узлах.

Анализ известных решений в области управления ресурсом сетевого узла

К настоящему моменту времени известно достаточно много математических моделей управления канальным и (или) буферным ресурсом, в рамках которых с той или иной степенью учтены требования, предъявляемые к перспективным решениям в этой области. В работе [3] предложена потоковая модель динамического обслуживания очередей со статическим закреплением пропускной способности исходящего канала связи за каждой из очередей. В работах [4, 5] в рамках предложенных математических моделях уже обеспечивается и динамический характер решений задачи распределения и канального ресурса, согласованный с результатами решения задачи обслуживания очередей. При этом в работе [4] за распределение канального и буферного ресурса отвечают различные управляющие переменные, а в работе [5] решение задачи по распределению пропускной способности канала связи осуществляется по факту решения задачи обслуживания очередей, что не позволяет выделить тому или иному типу трафика большей величины канальной емкости, чем его интенсивность., а

ведь это является одним из основных условий достижения заданных значений временных показателей QoS и показателей надежности доставки пакетов. Поэтому в настоящей работе будет предложено дальнейшее развитие математической модели управления очередями и пропускной способностью исходящего канала связи на сетевом узле, но уже с учетом конечной емкости буфера на сетевом узле.

Модель динамического управления канальным и буферным ресурсом на узлах ТКС

В ходе последующего изложения математической модели управления канальным и буферным ресурсом на сетевом узле предполагается, что количество отдельных трафиков или агрегированных по классам или приоритетам потоков известно и равно M , что соответствует принятым на практике решениям в рамках известных методов маркировки пакетов. Кроме того, оговоримся, что число очередей на сетевом узле также фиксировано (N) или назначается административно. Например, в алгоритме приоритетного обслуживания (Priority Queuing, PQ) может быть выделено до четырех очередей ($N = 4$), а в алгоритме взвешенного справедливого обслуживания (Weighted Fair Queuing, WFQ) по умолчанию число очередей равно 256 [3].

Предлагаемая модель будет носить потоковый характер, т.к. современный, преимущественно мультимедийный сетевой трафик также имеет потоковую природу. Кроме того, при решении задач обеспечения заданных показателей QoS необходимо оперировать, прежде всего, с интенсивностью трафика, а уже потом с отдельными пакетами, накапливающимися в буфере, что характерно для существующих алгоритмов FQ (Fair Queuing), WFQ (Weighted Fair Queuing) и др. [1, 2]. В этой связи 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 канала связи, которая выделена j -й очереди ($j = \overline{1, N}$), что типично, например, для алгоритма CBWFQ (Class-Based Weighted Fair Queuing). При этом необходимо выполнить условие отсутствия перегрузки канала связи:

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

где b – пропускная способность исходящего КС.

Кроме того, с целью предотвращения перегрузки сетевого узла необходимо обеспечить выполнение следующего условия:

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

Выполнение условия (2) состоит в том, что суммарный (агрегированный) поток пакетов, поступающих на сетевой узел, не должен превышать пропускную способность исходящего канала связи. Требование (2) обеспечивается, во-первых, за счет эффективной маршрутизации трафика в сети, во-вторых, путем использования средств борьбы с перегрузкой на сетевом узле, т.е. алгоритмов произвольного раннего обнаружения перегрузки и ограничения длины очереди (RED – Random Early Detection, WRED – Weighted RED).

Придать динамический характер этому обслуживанию очередей в рамках предлагаемой модели удалось путем введения переменной x_{ij} , под которой подразумевалась доля i -го трафика, который будет направлен для обслуживания в j -ю очередь. Согласно физическому смыслу переменной x_{ij} имеют место следующие дополнительные условия:

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

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

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

Выполнение условия (4) гарантирует отсутствие потерь пакетов при их обслуживании на рассматриваемом сетевом узле. Условия (5) вводятся для предотвращения перегрузки по пропускной способности отдельных очередей сетевого узла в процессе управления. По аналогии с моделью, рассмотренной в работе [4], в качестве искомого вектора выберем вектор

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

в ходе расчета которого удастся обеспечить согласованность в решении задач обслуживания очередей и динамического распределения за ними пропускной способности исходящего канала связи.

Формализация условий предотвращения перегрузки очередей на узле ТКС

Важно отметить, что даже в ходе выполнения условий (5) ввиду случайного и нестационарного характера сетевого трафика на узле могут возникать очереди и связанные с ними задержки пакетов. С целью введения верхней границы подобных задержек на узлах ТКС общую буферную емкость, как правило, ограничивают. Таким образом, для каждой очереди определим ее текущую загрузженность и максимальную вместимость, обозначив их соответственно через $\bar{n}_j$ и $n_j^{\max}$ ($j = \overline{1, N}$), и дополним условия предотвращения перегрузки отдельных очередей по их пропускной способности (5) условиями предотвращения перегрузки очередей по их длине. В самом общем виде искомые условия будут иметь следующий вид:

$$\bar{n}_j \leq n_j^{\max} \quad (j = \overline{1, N}) \quad (7)$$

и задача теперь сводится лишь к выбору (обоснованию) аналитического выражения для расчета средней длины очереди в процессе обслуживания.

Для формулировки искомых условий необходимо задаться моделью трафика (его характеристиками – интенсивность, длина пакета и т.д.) и моделью обслуживания пакетов в рамках отдельно взятой очереди, в качестве которой на практике, как правило, реализуется модель FIFO (First In, First Out). Кроме того, выбирая гипотезу относительно характеристик обслуживаемого трафика, можно описать отдельную очередь известной системой массового обслуживания (СМО), в рамках которой уже можно корректно рассчитать ее вероятностно-временные характеристики, в том числе время обслуживания и среднюю (текущую) длину очереди. Так, например, при моделировании процесса обслуживания в отдельной очереди различными типами систем массового обслуживания средняя длина очереди, опуская индекс очереди, будет выражаться следующими аналитическими зависимостями [6]:

СМО М/М/1:
$$\bar{n} = \frac{\rho}{1-\rho};$$

$$\text{СМО } M/M/1/n^{\max} : \quad \bar{n} = \frac{\rho^2 \left[1 - (n^{\max} + 1)\rho^{n^{\max}} + n^{\max} \rho^{n^{\max} + 1} \right]}{(1 - \rho^{n^{\max} + 2})(1 - \rho)};$$

$$\text{СМО } M/D/1: \quad \bar{n} = \frac{\rho}{1 - \rho} - \frac{\rho^2}{2(1 - \rho)};$$

$$\text{СМО } M/D/1/n^{\max} : \quad \bar{n} = \frac{\sum_{n=0}^{n^{\max}} n \left[\sum_{j=0}^n \frac{e^{j\rho} (-j\rho)^{n-j}}{(n-j)!} - \sum_{j=0}^{n-1} \frac{e^{j\rho} (-j\rho)^{n-j-1}}{(n-j-1)!} \right]}{\sum_{j=0}^{n^{\max} + 1} \frac{e^{j\rho} (-j\rho)^{n^{\max} - j + 1}}{(n^{\max} - j + 1)!}};$$

$$\text{СМО самоподобного трафика:} \quad \bar{n} = \frac{\rho^{1/2(1-H)}}{(1 - \rho)^{H/(1-H)}},$$

$$\text{где } \rho = \frac{\sum_{i=1}^M a_i x_{ij}}{b_j}, \quad 0.5 \leq H \leq 1 - \text{параметр самоподобия (Херста)}.$$

При этом каждому типу трафика, а значит и каждой очереди может соответствовать своя модель обслуживания, не обязательно соответствуя перечню вариантов СМО, приведенных выше. Кроме того, нетрудно заметить, что условие (7) является более строгим, чем требование (5).

Формулировка оптимизационной задачи по управлению канальным и буферным ресурсом сетевого узла

В связи с тем, что, в общем случае, выбор x_{ij} и b_j в рамках ограничений (1), (3), (4) и (7) можно произвести множеством случаев, то целесообразно задачу, связанную с расчетом искомых переменных (6), сформулировать в виде оптимизационной задачи со следующей минимизируемой целевой функцией:

$$T(x) = \bar{c}^t \bar{x}, \quad (8)$$

где координаты вектора весовых коэффициентов

$$\bar{c} = \begin{bmatrix} c_{ij} \\ \dots \\ c_j \end{bmatrix} \quad (9)$$

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.

Важно уточнить, что, как правило, выбирается $c_{ij} > c_{in}$ в том случае, если в рамках реализуемой дисциплины обслуживания j -я очередь должна использоваться для обслуживания трафика i -го класса лишь при заполнении n -й очереди. Аналогично, если $c_i > c_j$ (т.е. удельная стоимость выделения пропускной способности исходящего канала связи для i -й очереди выше, чем для j -й очереди), то пропускная способность канала будет первоочеред-

но выделяться именно j -й очереди. Таким образом, путем ранжирования численных значений координат c_{ij} и c_j вектора весовых коэффициентов (9) можно добиться реализации в рамках предложенной модели различных типов очередей – заказных, приоритетных, справедливых и др.

Кроме того, т.к. часть искомым переменных булевы (3), то каждый поток будет направляться лишь в одну очередь, причем некоторые потоки (или отдельные трафики) могут обслуживаться одной очередью. Как правило, подобная ситуация характерна для трафиков с соизмеримыми требованиями к качеству своего обслуживания. Ранее, в работе [3] рассматривался случай *статического* закрепления пропускной способности канала связи за очередями, а переменные x_{ij} характеризовали интенсивность i -го трафика, который будет направлен для обслуживания в j -ую очередь. В таком случае пакеты одного и того же трафика могли обслуживаться больше чем одной очередью, в случае если неосновная очередь «простаивала», что характерно для механизма СВQ. Для случая, рассмотренного в данной статье, при освобождении той или иной очереди, закрепленная за ней пропускная способность канала связи динамически будет перераспределена в интересах другой очереди, что является неоспоримым преимуществом приведенного решения.

Таким образом, предложенная математическая модель управления каналным и буферным ресурсом (1)-(9) ввиду нелинейности ограничений (7) представлена в виде задачи нелинейного программирования, в которой часть искомым переменных булевы. Для решения сформулированной задачи известны точные и приближительные методы [7].

Примеры решения задачи управления очередями с динамическим распределением пропускной способности исходящего канала связи

Для демонстрации особенностей работы предлагаемой модели приведем пример решения задачи управления очередями с динамическим распределением пропускной способности исходящего канала связи. В качестве исходных данных будут выступать: число потоков равно пяти ($M = 5$), а число очередей – трем ($N = 3$). При этом, пусть первый трафик имел наивысший приоритет, второй приоритет имели второй и третий трафики, а четвертый и пятый трафики имели третий приоритет. Приоритетность же очередей уменьшалась с возрастанием их порядкового номера. В ходе расчетов, для примера, за основу была принята модель обслуживания пакетов на сетевых узлах типа М/М/1. В рамках рассмотренного примера характеристики обслуживаемых трафиков и пропускная способность исходящего канала связи приведены в табл.1. В этой же таблице в соответствии с присвоенными приоритетами выбраны значения вектора весовых коэффициентов (9).

Исходные данные для расчета

Табл.1

Интенсивности трафика различного класса (тыс. пакетов/с) (суммарный трафик 284 тыс. пакетов/с)										Пропускная способность исходящего канала связи (тыс. пакетов/с)		Максимальная емкость очереди (пакетов)						
a_1	a_2	a_3	a_4	a_5						b	$n_1^{\max}$	$n_2^{\max}$	$n_2^{\max}$					
60	54	54	58	58						300	12	25	30					
Весовые коэффициенты c_{ij} и c_j																		
$c_{1,1}$	$c_{1,2}$	$c_{1,3}$	$c_{2,1}$	$c_{2,2}$	$c_{2,3}$	$c_{3,1}$	$c_{3,2}$	$c_{3,3}$	$c_{4,1}$	$c_{4,2}$	$c_{4,3}$	$c_{5,1}$	$c_{5,2}$	$c_{5,3}$	c_1	c_2	c_3	
20	30	30	50	20	20	50	20	20	50	50	20	50	50	20	10	20	30	

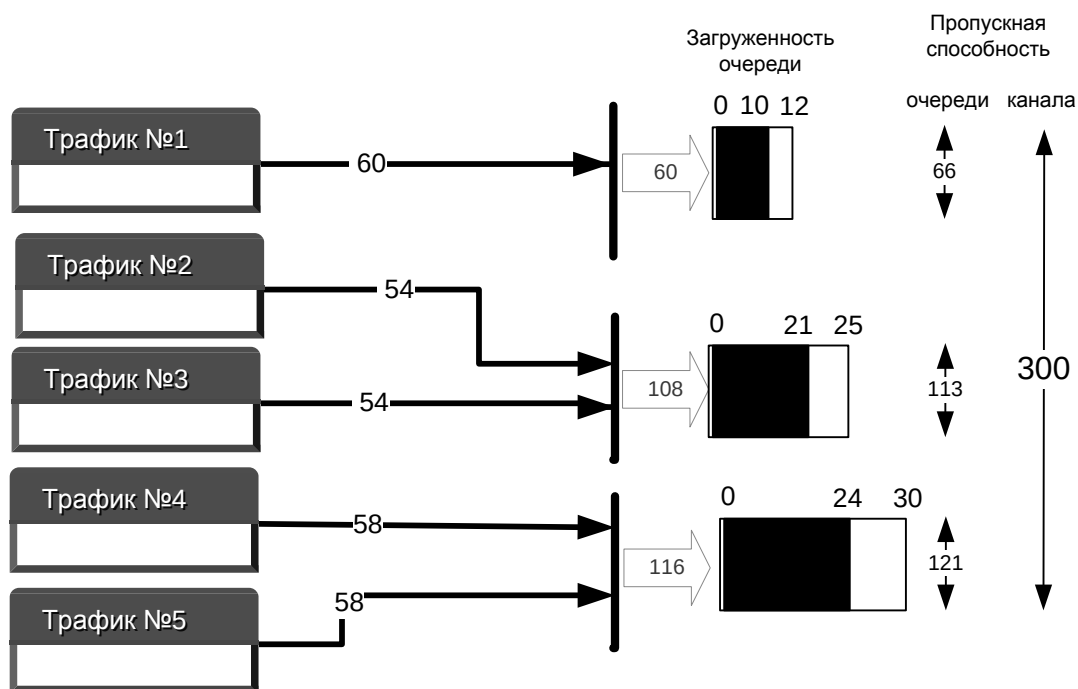


Рис.1. Порядок решения задачи управления очередями с динамическим распределением пропускной способности исходящего канала связи

Выводы

Таким образом, в статье предложена математическая модель динамического управления канальным и буферным ресурсом на узлах телекоммуникационной сети. Новизна модели состоит в том, что в ней в отличие от ранее известных подходов произведен учет конечного объема буфера на узлах ТКС, что наложило отпечаток на содержание введенные ограничения. Они носили нелинейный характер и определялись всецело характером трафика и выбранной моделью обслуживания пакетов на сетевых узлах. Расчет искомых переменных, отвечающих за управление очередями и пропускной способностью канала связи, осуществлен в ходе сформулированной на выходе модели оптимизационной задач, относящейся к классу задач нелинейного программирования.

Список литературы: 1. *Справочник по телекоммуникационным технологиям*: Пер. с англ. М.: Издательский дом «Вильямс», 2004. 640 с. 2. *Вегенша 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 // *Наукові записки УНДІЗ.* 2008. №3(5). С. 34-39. 5. *Симоненко А.В., Ахмад Хайлан, Али Али* Модель динамического управления очередями и пропускной способностью канала связи на маршрутизаторах мультисервисной сети // *Радиотехника: Всеукр. межвед. науч.-техн. сб.* 2008. Вып. 155. С. 164-168. 6. *Петров М.Н.* Вероятностно-временные характеристики в сетях и системах передачи интегральной информации: Научное издание КГТУ, Красноярск, 1997. 220 с. 7. *Матрешин Н.П., Макеева В.К.* Математическое программирование. Изд. 2-е, перераб. и доп. Х.: Вища школа, 1978. 160 с.