

УДК 621.391

РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДА ДИНАМИЧЕСКОЙ БАЛАНСРОВКИ ОЧЕРЕДЕЙ НА МАРШРУТИЗАТОРАХ МУЛЬТИСЕРВИСНОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ



М.В. СЕМЕНЯКА

Харьковский национальный
университет радиоэлектроники



А.В. СИМОНЕНКО

Харьковский университет
Воздушных Сил им. Ивана Кожедуба



Али С. Али

Харьковский национальный
университет радиоэлектроники

Abstract - A method of queues dynamic balancing on the telecommunications network routers was proposed. Within the method flow-based queues balancing mathematical models use was adaptive and combined, queues load and requirements for quality of service accounted. Queue balancing models were reduced to solving optimization problems of linear and nonlinear programming. Method implements the so-called hierarchical queue (HQ), adaptively (based on the analysis of the interface) chooses a strategy of queues balancing, provides differentiated packet processing and distribution of bandwidth between queues, takes into account the priority and the length of a packet. Top level of the method is based on linear programming model, this allowed previous distribution of flows per queue with consolidating certain bandwidth for outgoing connection (physical interface). On the lower level of the method, each particular queue is divided into sub-queues and balancing of these sub-queues is based on more complex models of mixed integer and nonlinear programming. With the proposed method the queue balancing process was experimentally investigated with varying degrees of flow accounting parameters.

Анотація – Запропоновано метод динамічного балансування черг на маршрутизаторах телекомунікаційної мережі, в рамках якого здійснюється комбіноване адаптивне використання поточкових математичних моделей управління чергами з урахуванням їх завантаженості та вимог до якості обслуговування. В рамках запропонованих моделей задачу балансування черг було зведено до розв'язання оптимізаційних задач лінійного і нелінійного програмування. Метод реалізує так звані ієрархічні черги, адаптивно (на основі аналізу стану інтерфейсу) обираючи стратегію балансування черг і здійснюючи диференційовану обробку пакетів та розподіл пропускної здатності каналу зв'язку між чергами з урахуванням пріоритету і довжини пакетів того чи іншого потоку.

Аннотация – Предложен метод динамической балансировки очередей на маршрутизаторах телекоммуникационной сети, в рамках которого осуществляется комбинированное адаптивное использование поточковых математических моделей управления очередями с учетом их загруженности и требований к качеству обслуживания. В рамках предложенных моделей задача балансировки очередей была сведена к решению оптимизационных задач линейного и нелинейного программирования. Метод реализует так называемые иерархические очереди, адаптивно (на основе анализа состояния интерфейса) выбирая стратегию балансировки очередей и осуществляя дифференцированную обработку пакетов и распределение пропускной способности канала связи между очередями с учетом приоритета и длины пакетов того или иного потока.

Введение

Современные тенденции совершенствования и эволюции сетей связи определяет концепция сетей нового поколения (Next Generation Network, NGN), в которых повышение качества обслуживания (Quality of Service, QoS) пользователей является одним из важнейших направлений развития инфокоммуникационных технологий [1, 2]. Превалирующим в развитии средств обеспечения качества обслуживания яв-

ляется переход к динамическим стратегиям управления сетевыми ресурсами – буферным пространством сетевых узлов, пропускной способностью каналов связи и сетевым трафиком. С целью повышения гибкости управления сетевыми ресурсами динамический характер приобретают задачи управления, балансировки очередями (Congestion Management), от эффективности которых во многом зависят значения таких показателей качества обслуживания, как средняя задержка, джиттер и уровень потерь пакетов. В этой связи, на смену статическим решениям, заложенных в основу существующих механизмов балансировки очередей, должны прийти динамические математические модели и методы балансировки очередей, ориентированные на получение адаптивных решений по отношению к загруженности узла телекоммуникационной сети (ТКС) в целом, а также к требуемым значениям показателей QoS.

В свою очередь существующие математические модели управления буферным ресурсом сетевого узла, представленные, в частности, в работах [3-8], отличаются в силу положенных в их основу математических решений как своей программной сложностью, так и областью применения. Исходя из этого, актуальной представляется задача по разработке метода балансировки очередей на основе комбинированного адаптивного использования ранее разработанных математических моделей для обеспечения требуемого качества обслуживания в условиях динамически изменяющегося состояния сети и отдельных ее узлов и каналов связи.

I. Анализ известных решений

В настоящее время в современных ТКС нашло свое применение достаточно большое количество схем и механизмов управления очередями [1, 2] с точки зрения их обслуживания и превентивного ограничения длины очереди. Существующие основные механизмы управления очередями – FIFO Queueing (First In, First Out); приоритетные очереди (Priority Queuing, PQ); заказные очереди (Custom Queuing, CQ); взвешенные справедливые очереди (Weighted Fair Queuing, WFQ); взвешенные справедливые очереди, основанные на классах (Class-based Weighted Fair Queueing, CBWFQ); очереди с малой задержкой (Low Latency Queuing, LLQ) в своих решениях основаны на необходимости статического управления буферным ресурсом сетевого узла. Это в свою очередь приводит к несогласованности в решении следующих задач:

- 1) распределение поступающих на маршрутизатор информационных потоков в соответствии с их интенсивностью и приоритетностью по множеству очередей;
- 2) определение порядка обслуживания пакетов из различных очередей, что так или иначе связано с распределением пропускной способности канала связи (КС) между пакетами разных потоков.

Статический характер решений основан на ручной настройке администратором очередей, что в условиях динамически изменяющихся характеристик трафика и состояния интерфейсов сетевого узла приводит к невозможности реализации управления очередями в реальном времени. К примеру, протоколы динамической маршрутизации в зависимости от состояния сети пересчитывают порядок использования доступных маршрутов периодически (через 30-90 с) или по требованию. В задачах

же Congestion Management загрузка очередей необходимо контролировать более регулярно, речь идет о миллисекундах, т.е. придание решению задач управления очередями автоматического характера является первоочередной проблемой.

Недостатки технологических решений определили актуальность проведения дополнительных научных исследований в области обслуживания очередей на узлах сети. Например, с целью более полного учета характеристик обслуживаемого трафика необходимо использовать именно потоковые модели (flow-based model), в рамках которых учитывается также интенсивность трафика наряду с другими важными параметрами – длиной пакета, его приоритетом и т.д.

В области управления очередями также накоплен немалый теоретический опыт [3-12], включающий в себя математические модели, методы и алгоритмы, которые по разным причинам пока еще не были положены в основу перспективных средств Congestion Management. Кроме того, среди множества работ, содержание которых все больше адаптируется под беспроводные технологии [9-12], заслуживает внимания подход, основанный на реализации идей балансировки очередей [3-8]. Чем более сбалансированной будет загрузка очередей, тем меньше будет их средняя длина и вероятность переполнения (вероятность потерь пакетов). В технологии MPLS парадигма обеспечения сбалансированной загрузки сетевых ресурсов получила название инжиниринга трафика (Traffic Engineering, TE), важное место в которой занимают решения по балансировке очередей (Traffic Engineering Queues, TEQ).

Идеи Traffic Engineering Queues получили свое развитие в работах [4-8, 13, 14]. В работах [4, 5] балансировка очередей основана на введении соответствующих метрик, а прикладная сетевая задача сведена к оптимизационной задаче линейного программирования. Основными достоинствами этих моделей являются совместное решение задачи распределения агрегированного потока трафика по очередям и закрепления за каждой очередью фиксированной пропускной способности, а также динамический характер решения задачи распределения пакетов того или иного потока трафика по очередям маршрутизатора. Задача балансировки решается методами линейного программирования, что существенно упрощает математическую, а следовательно, и аппаратную реализацию метода. Недостатками моделей является то, что процесс балансировки осуществляется для агрегированного потока трафика, а дифференциация потоков осуществляется за счет статической настройки соответствующих метрик обслуживания. Эти модели [4, 5] также не учитывают загрузку очередей интерфейса маршрутизатора по их длине.

В работе [8] предложена и исследована модель балансировки очередей, учитывающая загрузку очередей по пропускной способности, выделенной для обслуживания каждой очереди. К достоинствам модели стоит отнести полное соответствие полученных в ней решений концепции Traffic Engineering Queues. Процесс балансировки основан на обработке отдельных потоков агрегированного трафика. К недостаткам стоит отнести необходимость статического закрепления за каждой очередью пропускной способности, а также усложненную программно-алгоритмическую реализацию данной модели, так как уже необходимо было решать задачу смешанного целочисленного программирования (Mixed Integer Linear Programming, MILP).

В работах [6, 7] постановка задачи балансировки очередей несколько усложнена по сравнению с моделью, описанной в [8], т.к. в ней обеспечен учет текущей загрузки каждой очереди, что придало самой задаче нелинейный характер и заметно усложнило получение искомым решений.

В результате проведенного анализа для обеспечения наилучших показателей качества обслуживания в данной работе предлагается метод динамической балансировки очередей на маршрутизаторах мультисервисной телекоммуникационной сети. В связи с тем, что каждая из существующих моделей имеет свои характерные преимущества и связанную с этим область применения, предлагаемый метод будет основываться на их комбинированном использовании, реализуя так называемые иерархические очереди (Hierarchical Queues, HQ).

II. Метод динамической балансировки очередей на маршрутизаторах телекоммуникационной сети

В рамках предлагаемого метода процесс балансировки очередей предлагается организовать следующим образом: на первом этапе осуществляется сбор информации о состоянии рассматриваемого интерфейса (исходных данных для расчета). На основе мониторинга состояния маршрутизатора и характеристик поступающего трафика анализируются следующие данные: количество потоков агрегированного трафика, который поступает на вход маршрутизатора; интенсивность каждого потока; приоритет обслуживания того или иного потока; количество очередей, поддерживаемых на интерфейсе; пропускная способность исходящего канала связи. При этом *на верхнем уровне метода* управление очередями будет производиться на основе потоковой модели балансировки очередей, основанной на метриках [4], результатом чего будет предварительное распределение потоков по очередям с закреплением за этими очередями определенной пропускной способности исходящего канала связи (физического интерфейса).

В рамках этой модели [4] a_i ($i = \overline{1, M}$) – интенсивность трафика i -го класса, поступающего на обслуживание маршрутизатора ТКС; b_j ($j = \overline{1, N}$) – часть пропускной способности исходящего КС, которая выделена для обслуживания пакетов из j -й очереди. Для обеспечения условия сохранения потока и предотвращения перегрузки маршрутизатора введены следующие ограничения:

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

где b – пропускная способность исходящего интерфейса (канала связи).

Придать динамический характер процессу распределения пакетов удалось в модели за счет введения переменной x_{ij} , под которой подразумевалась часть i -го потока, который будет направлен для обслуживания в j -ю очередь. Согласно физическому смыслу переменной x_{ij} имеют место следующие дополнительные условия:

$$x_{ij} \geq 0 \quad (i = \overline{1, M}, j = \overline{1, N}), \quad \sum_{j=1}^N x_{ij} = a_i \quad (i = \overline{1, M}), \quad \sum_{i=1}^M x_{ij} \leq b_j \quad (j = \overline{1, N}). \quad (2)$$

Выполнение условия (2) гарантирует обслуживание всех пакетов трафика, поступающих на отдельный маршрутизатор, а также предотвращение перегрузки очередей маршрутизатора по пропускной способности. В рамках данной модели в качестве искомым выступают переменные x_{ij} , представленные вектором \vec{x} , а оптимальные решения по распределению потоков за очередями находятся в ходе минимизации следующей целевой функции:

$$F(x) = \vec{c}^t \vec{x}, \quad (3)$$

в которой \vec{c} – вектор весовых коэффициентов (метрик), который характеризует условную стоимость (c_{ij}) использования пакетами i -го потока ресурсов j -й очереди.

На *нижнем уровне метода* предполагается, чтобы каждая конкретная очередь разбивалась на подочереди, что является характерным решением для НQ, причем балансировка уже этих подочередей будет осуществляться на основе моделей, описанных в [6, 8]. Основываясь на результатах решения верхней модели балансировки очередей, к использованию предлагается модель балансировки очередей, учитывающая загруженность по пропускной способности. В отличие от модели, основанной на метриках, в рамках данной потоковой модели в качестве искомым выступают переменные y_{ij} , под которыми понималась доля интенсивности i -го потока трафика, который будет направлен для обслуживания в j -ю очередь. В данном случае управляющие переменные имели булеву природу. С учетом условия предотвращения потерь пакетов на физическом интерфейсе маршрутизатора необходимые ограничения имели следующий вид:

$$y_{ij} \in \{0,1\} \quad (i = \overline{1, M}, j = \overline{1, N}), \quad \sum_{j=1}^N y_{ij} = 1 \quad (i = \overline{1, M}), \quad (4)$$

где величина M трактовалась как число отдельных потоков, поступающих в рамках агрегированного трафика на рассматриваемый интерфейс.

С целью недопущения перегрузки очереди по пропускной способности (по потоку) в рамках предлагаемой модели вводится следующее условие:

$$\sum_{i=1}^M y_{ij} a_i \leq \alpha \cdot b_j \quad (j = \overline{1, N}), \quad 0 \leq \alpha \leq 1, \quad (5)$$

где параметр α характеризует верхний динамически управляемый порог загруженности очередей на физическом интерфейсе маршрутизатора. Данный параметр выступает в качестве искомого, как и переменные (4). Условие (5) представляет собой видоизмененную форму условия (2).

В рамках данной предлагаемой потоковой модели параметры b_j ($j = \overline{1, N}$) характеризовали пропускную способность исходящего канала связи, выделенную для передачи пакетов из j -й очереди, которая предполагалась известной величиной.

Основываясь на пересмотре условия предотвращения перегрузки очередей по пропускной способности (по потоку) (5), в рамках рассматриваемой модели также видоизменяется и критерий оптимальности получаемых решений, в качестве которого выступала линейная форма

$$\min_{y, \alpha} \alpha. \quad (6)$$

В условии (6) параметр α по своему физическому смыслу соответствует максимальному коэффициенту загрузки той или иной очереди, от значения которого согласно результатам теории массового обслуживания [15, 16] напрямую зависят такие важные показатели качества обслуживания как средняя задержка и уровень потерь пакетов. Минимизацией этого параметра добиваемся улучшения этих показателей. Задача, связанная с минимизацией критерия (6), относится к классу задач смешанного целочисленного линейного программирования, т.к. часть переменных являются булевыми (4), а часть (α) вещественными числами.

В ходе анализа потоковой модели балансировки загрузки очередей по пропускной способности установлено, что эффективность ее работы возрастает при росте числа потоков, образующих поступающий на интерфейс агрегированный трафик, при одновременном снижении и выравнивании их интенсивности (рис. 1, а). В этих условиях динамически управляемый верхний порог загрузки очередей по трафику возрастал практически линейно (рис. 1, б). Чем линейнее эта зависимость, тем более точно предлагаемая модель реализует обобщенную схему разделения процессорного времени (Generalized Processor Sharing, GPS), обеспечивая наиболее справедливое (равномерное) обслуживание очередей [1].

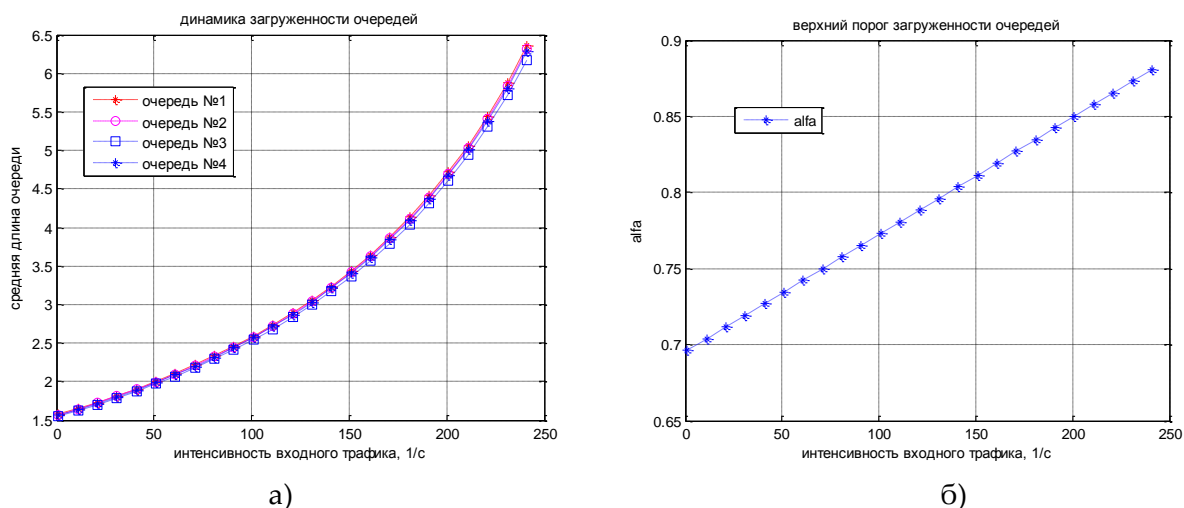


Рис. 1. Результаты анализа модели (4)-(6) для $M = 10$ и $N = 4$

В случае же, когда число потоков уменьшалось с одновременным ростом их интенсивности, ситуация по балансировке складывалась не столь идеальной, о чем свидетельствуют результаты, полученные на рис. 2.

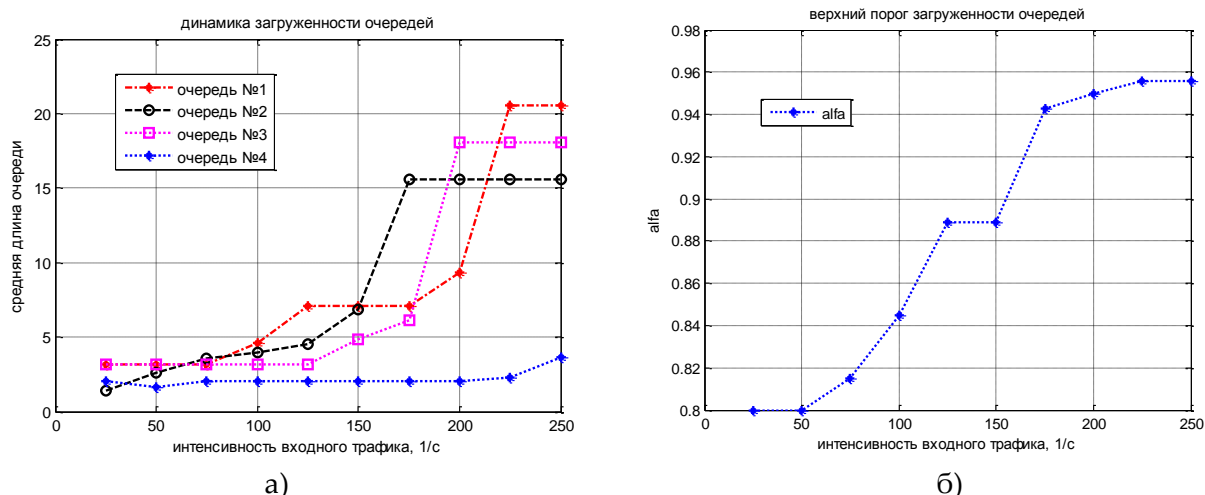


Рис. 2. Результаты анализа модели (4)-(6) для $M = 6$ и $N = 4$

По результатам можно сделать вывод, что модель (4)-(6) обеспечивает лучшую балансировку при условии (*условие №1*), что агрегированный трафик, поступающий на интерфейс, образуется достаточно большим количеством (по отношению к числу очередей) отдельных потоков приблизительно равного приоритета и невысокими (и желательно соизмеримыми между собой) интенсивностями. В этом случае к обслуживанию очередей предлагается потоковая модель балансировки загрузки очередей по пропускной способности. Кроме того, данную модель можно использовать лишь в случае (*условие №2*) низкой и средней загрузки интерфейса, т.к. в ней в явном виде отсутствуют условия предотвращения перегрузки очередей по их длине. Данную модель балансировки целесообразно использовать в указанных условиях ввиду ее более простой алгоритмическо-программной (а значит и аппаратной) реализации по сравнению с моделью, учитывающей загрузку очередей по их длине [6], так как в первой модели решается более простая MILP-задача, а во второй – более сложная MILNP (Mixed Integer Non-linear Programming) задача.

Таким образом, когда условия применения модели (4)-(6) не выполняются, то необходимо обеспечивать балансировку очередей на основе использования потоковой модели балансировки очередей, учитывающей длину очереди [6]. Особенно актуально ее применять, во-первых, в области высокой загрузки интерфейса, когда очень важно предотвратить перегрузку очередей по их длине, во-вторых, при существенном расхождении в приоритетах обслуживаемых потоков трафика, и в третьих, при условии, что агрегированный поток, поступающий на интерфейс, образует относительно небольшое число (по сравнению с числом очередей) потоков высокой интенсивности (по сравнению с ПС КС, закрепленной за каждой отдельной очередью).

В рамках предлагаемой модели по-прежнему в качестве искомым выступают булевы переменные y_{ij} с соблюдением условий (4).

Сумма пропускных способностей, выделенных для обслуживания пакетов той или иной очереди, не должна превышать физической ПС исходящего канала связи. Во избежание перегрузки каждой отдельной очереди по ПС (по потоку) вводятся условия (аналог (5)):

$$\sum_{i=1}^M a_i y_{ij} \leq b_j \quad (j = \overline{1, N}), \quad \sum_{j=1}^N b_j \leq b. \quad (7)$$

Условие (7) вводится для предотвращения перегрузки пропускной способности КС, выделяемой для передачи пакетов той или иной очереди маршрутизатора в ходе их балансировки. В отличие от модели (4)-(6), но по аналогии с моделью, основанной на метриках, в качестве искомого вектора выберем вектор

$$\bar{y} = \begin{bmatrix} y_{ij} \\ \dots \\ b_j \end{bmatrix} \quad (i = \overline{1, M}; j = \overline{1, N}), \quad (8)$$

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

В ходе выполнения условий (7) ввиду случайного и нестационарного характера сетевого трафика на маршрутизаторе возникают очереди и связанные с ними дополнительные задержки пакетов. С целью введения верхней границы подобных задержек на маршрутизаторах ТКС общую буферную емкость, как правило, ограничивают десятками пакетов. Таким образом, для каждой очереди определим ее текущую загруженность и максимальную емкость, обозначив их соответственно через \bar{n}_j и n_j^{\max} ($j = \overline{1, N}$). Кроме того, дополним условия предотвращения перегрузки отдельных очередей по их пропускной способности (7) условиями предотвращения перегрузки очередей по их длине. В общем виде искомые условия будут иметь вид:

$$\bar{n}_j \leq n_j^{\max} \quad (j = \overline{1, N}), \quad (9)$$

и задача теперь сводится лишь к выбору (обоснованию) аналитического выражения для расчета средней длины очереди в процессе обслуживания. Для формулировки искомого условия необходимо задаться моделью трафика (его характеристиками – интенсивностью, длиной пакета и т.д.) и моделью обслуживания пакетов в рамках отдельно взятой очереди.

В связи с тем, что, в общем случае, выбор управляющих переменных y_{ij} и b_j в рамках ограничений (7)-(9) можно произвести множеством случаев, то целесообразно задачу, связанную с расчетом вектора (8), сформулировать в виде оптимизационной. Основным требованием к целевой функции является учет физики протекающих на узле процессов обслуживания пакетов (7)-(9), а также соответствие получаемых решений принципам концепции Traffic Engineering Queues, касающихся обеспечения сбалансированной загрузки буферных ресурсов. С этой целью вышеизложенную модель (4), (7)-(9) важно дополнить следующими условиями:

$$f(p_j, d_j) \cdot \bar{n}_j \leq \beta \cdot n_j^{\max} \quad (j = \overline{1, N}), \quad (10)$$

где β – верхний динамически управляемый предел загруженности очередей на

маршрутизаторе ТКС, $f(p_j, d_j)$ – некоторая функция от характеристик j -го потока, например, его приоритета p_j или длины пакета d_j .

Стоит отметить, что поток с более высоким приоритетом должен традиционно обслуживаться лучше, чем трафик с низким приоритетом. В результате, значение функции $f(p, d)$ должно быть тем больше, чем выше приоритет и чувствительность к задержке данного потока. В этой связи, в выражении (10) в качестве функции характеристик потока можно использовать следующее выражение:

$$f(p_j, d_j) = \frac{p_j}{v \cdot d_j} \quad (j = \overline{1, N}), \quad (11)$$

где v – некоторый нормировочный коэффициент, который должен сглаживать различие в порядке значений приоритета (1÷8) и длины пакета в байтах.

Тогда выражение (10) с учетом равенства (11), а также ограничений, накладываемых на коэффициент β , примет вид

$$\frac{p_j}{v \cdot d_j} \cdot \bar{n}_j \leq \beta \cdot n_j^{\max} \quad (j = \overline{1, N}), \quad 0 \leq \beta \leq 1. \quad (12)$$

Если в одной очереди обслуживаются потоки с различными (но по определению близкими) значениями длины и (или) приоритета пакета, то в выражении (12) целесообразно использовать их усредненные значения. Тогда задача обслуживания очередей может быть сведена к задаче балансировки их длины в ходе использования следующей формы:

$$\min_{y, b, \beta} \beta, \quad (13)$$

что соответствует минимизации верхнего порога загруженности очередей непосредственно по их длине на маршрутизаторе ТКС, взвешенного относительно таких характеристик потока, как длина пакета и его приоритет, что способствует сбалансированной загруженности всех очередей в соответствии с требованиями технологии Traffic Engineering Queues.

В отличие от ранее известных решений, предложенных в данной работе, в модели (4), (7)-(13) обеспечивается отсутствие перегрузки каждой конкретной очереди не только по ПС (по потоку), но по ее длине (9) и (12). Важной особенностью предлагаемого решения является также то, что балансировку в рамках предлагаемой модели планируется осуществлять с учетом приоритета и длины образующих ту или иную очередь пакетов.

Сама технологическая по своей сути задача обслуживания очередей в общем случае была сведена к оптимизационной задаче смешанного целочисленного нелинейного программирования (MILNP), связанной с минимизацией линейной функции (13) при наличии в т.ч. нелинейных ограничений (12), а также булевой природы некоторых рассчитываемых переменных (8).

В результате исследования предложенной модели с ростом интенсивности агрегированного трафика, поступающего на интерфейс, очереди по своей длине за-

гружались по-разному – с учетом приоритета и длины пакетов (рис. 3, а). Однако максимальный порог загрузки очередей по их длине возрастал линейно (рис. 3, б). Это было обусловлено тем, что в данной модели, в отличие от модели (4)-(6), динамически также перераспределялась и пропускная способность КС между очередями интерфейса, что придавало высокую гибкость самому процессу динамической балансировки очередей.

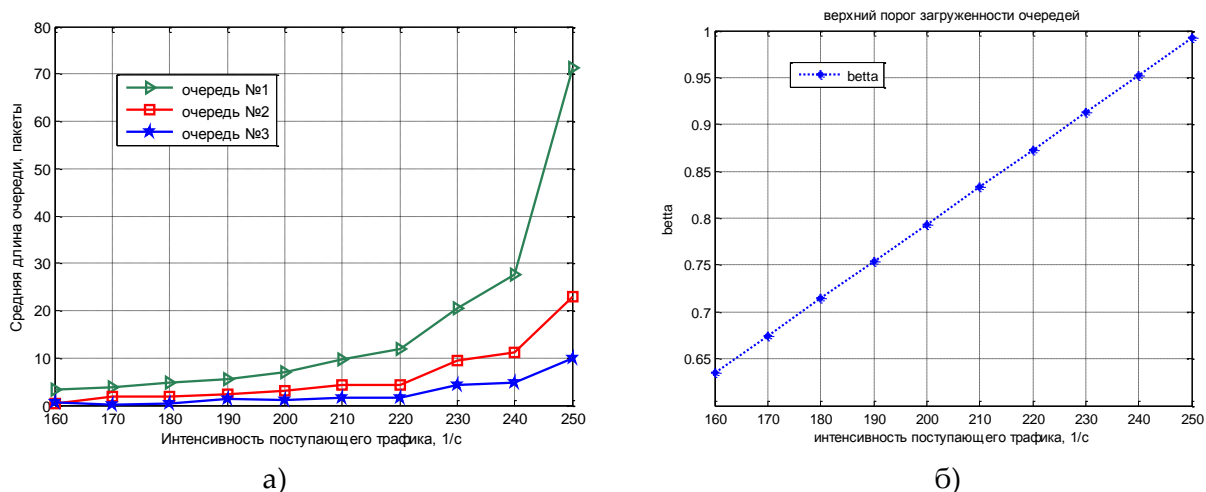


Рис. 3. Результаты анализа модели (4), (7)-(13) для $M = 8$ и $N = 3$

В целом потоковая модель балансировки очередей (4), (7)-(13) более функциональна, чем модель (4)-(6), но ее использование с точки зрения алгоритмическо-программной и аппаратной реализации несколько усложнено ввиду необходимости решения достаточно сложной в вычислительном аспекте оптимизационной MILNP-задачи. Поэтому и есть смысл использовать модель (4)-(6) при выполнении условий №1 и №2, т.к. в этом случае эффективность решений по балансировке очередей, обеспечиваемых и первой (4)-(6), и второй (4), (7)-(13) моделями, в целом приблизительно идентична.

Исходя из выше описанных математических моделей, а также условий их применения, на рис. 4 представлена структурная схема метода динамической балансировки очередей на маршрутизаторах телекоммуникационной сети, в рамках которого, во-первых, адаптивно (на основе анализа состояния интерфейса) выбирается стратегия балансировки очередей, а во-вторых, осуществляется дифференцированная обработка пакетов и распределение ПС канала связи между очередями с учетом приоритета и длины пакетов того или иного потока.

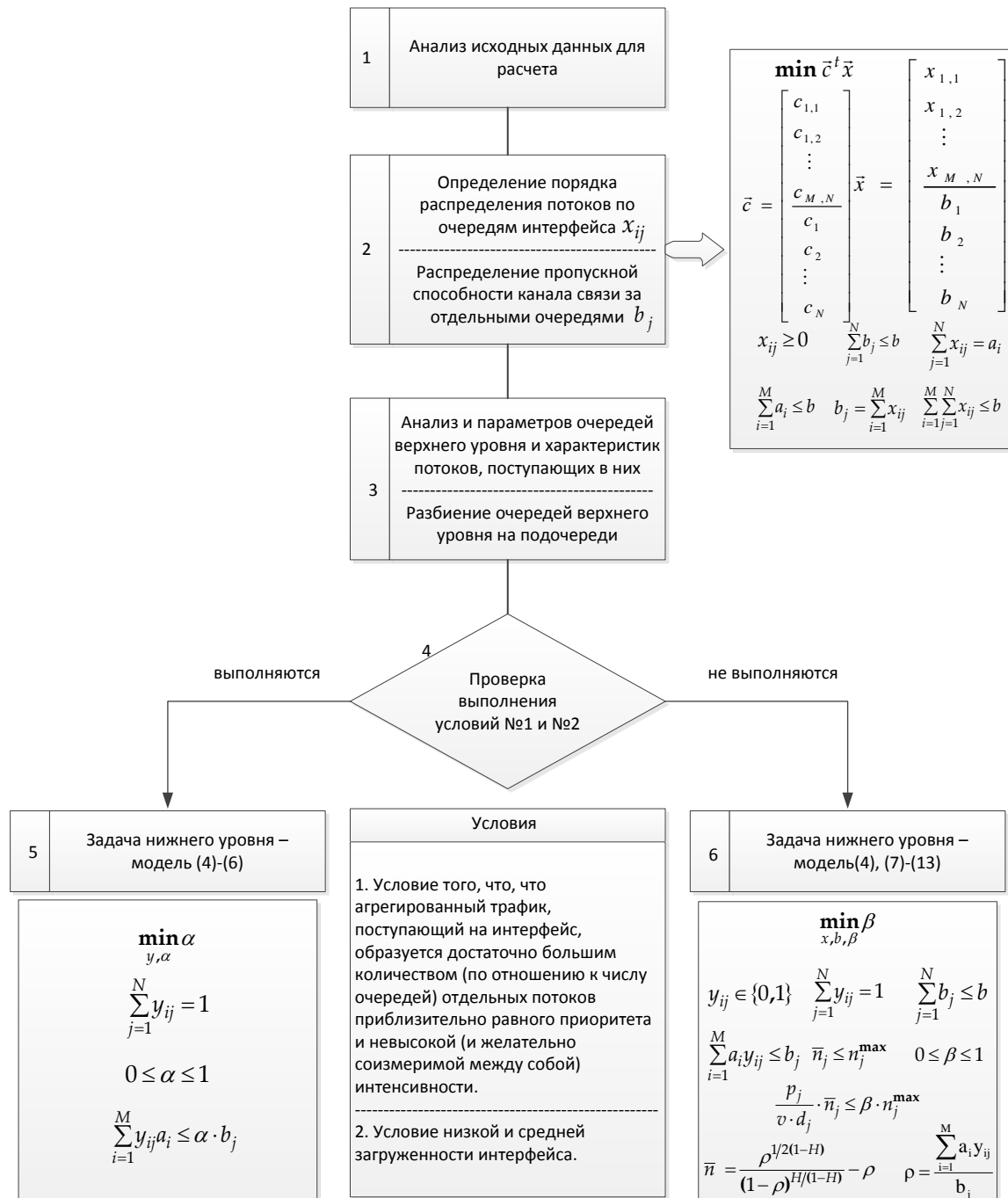


Рис. 4. Структура метода динамической балансировки очередей на маршрутизаторах мультисервисной телекоммуникационной сети

III. Исследование и оценка эффективности предложенного метода динамической балансировки очередей на маршрутизаторах мультисервисной телекоммуникационной сети

Для оценки эффективности решений, получаемых при использовании предложенного метода балансировки очередей на интерфейсах маршрутизаторов мультисервисных ТКС был проведен эксперимент, методика которого представлена в

[17]. В качестве варьируемых характеристик поступающего на обслуживание трафика, отдельных интерфейсов и сети в целом выступали следующие показатели:

- структура ТКС, а именно количество транзитных маршрутизаторов на пути передачи пакета от отправителя до получателя;
- число, интенсивность и приоритет потоков, образующих трафик, поступающий на интерфейс маршрутизатора;
- число очередей на интерфейсе маршрутизатора ТКС;
- пропускная способность каналов связи ТКС.

Проведен анализ зависимости средней межконцевой задержки пакетов от числа переприемов в маршруте передачи пакетов обслуживаемых потоков трафика и их интенсивности. Сравнивались варианты применения для управления очередями механизмов FIFO, WFQ и предлагаемого метода. Выбор сравниваемых механизмов был продиктован высоким уровнем их автоматичности работы, т.к. такие механизмы как CQ, PQ и LLQ во многом зависят от административных настроек, производимых вручную. Результаты расчетов по предлагаемому методу балансировки очередей эмулировались с помощью механизма CBWFQ в ходе классификации трафика и задания за очередями ПС с помощью команды «bandwidth».

Число переприемов варьировалось от 2 до 16, ограничиваясь лишь количественными характеристиками доступного сетевого оборудования (маршрутизаторов). Приоритеты пакетов в генерируемых потоках распределялись равномерно. Сравнению подлежали случаи, когда маршрутизаторы взаимодействовали между собой по интерфейсам Serial (до 2 Мбит/с) и Ethernet (10 Мбит/с).

Результаты сравнительного анализа показали, что в случае использования ненагруженных (до 4 Мбит/с) высокоскоростных интерфейсов (Ethernet) вне зависимости от числа переприемов лучшую межконцевую задержку (хоть и не на много, до 2-3%) обеспечивал механизм FIFO (рис. 5). Это объяснялось тем, что очередей на интерфейсах транзитных маршрутизаторах не возникало, а основная задержка обработки пакетов вносилась работой того или иного механизма обслуживания очередей. У FIFO эта задержка минимальна, так как число поддерживаемых и обслуживаемых очередей там равна 1. Также в этом механизме не производилась дифференциация пакетов разных потоков, что также несколько снижало время обработки пакетов. Особого расхождения в значениях межконцевой задержки при реализации предлагаемого метода и механизма WFQ не было (расхождение до 3%).

С ростом загрузки интерфейса (более 4 Мбит/с) алгоритм FIFO начал уступать по качеству обслуживания и механизму WFQ, и предлагаемому методу. Причем наблюдалось улучшение средней межконцевой задержки в среднем от 7-10% (при средних нагрузках и числе переприемов 2-4) до 17-26% (в условиях, близких к перегрузке при числе переприемов 6-16) при реализации предлагаемого метода по сравнению с механизмом WFQ (рис. 6). Таким образом, предлагаемый метод балансировки особо эффективно проявил себя при росте загрузки интерфейса и увеличении числа переприемов в маршруте передачи пакетов.

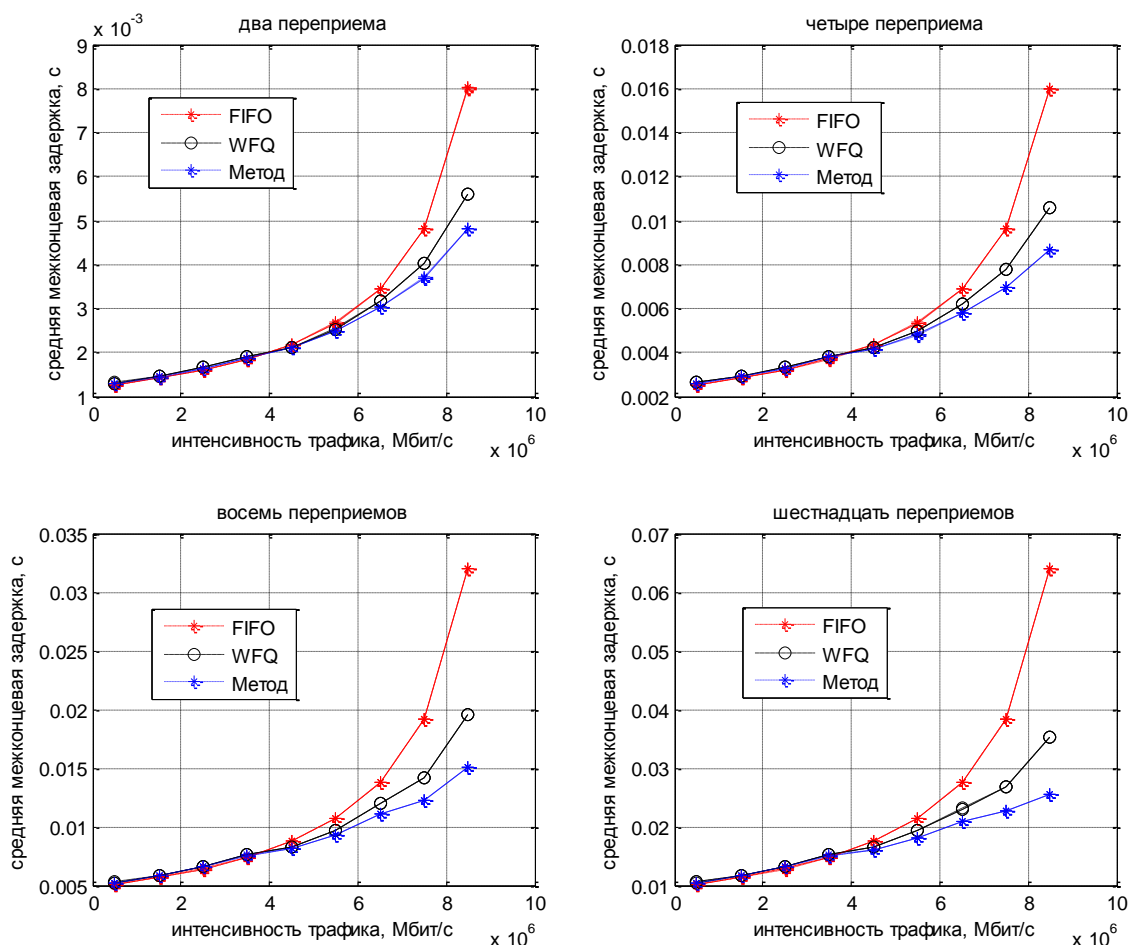


Рис. 5. Анализ влияния загруженности интерфейса и числа переприемов пакетов в маршруте на среднюю межконцевую задержку для различных методов управления очередями

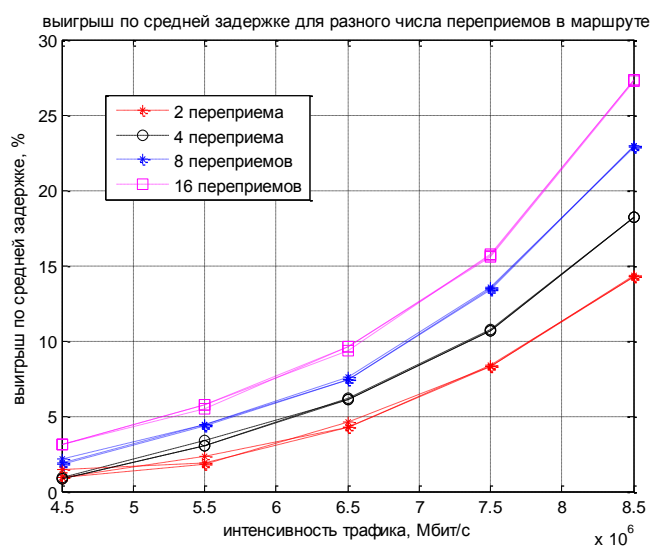


Рис. 6. Выигрыш предлагаемого метода по сравнению с механизмом WFQ по показателю средней межконцевой задержки

Выводы

На основе проведенного анализа известных подходов балансировки очередей в работе предложен метод адаптивной балансировки очередей на маршрутизаторах мультисервисной ТКС. Особенность метода состоит в том, что в нем обеспечивается адаптивный выбор стратегий балансировки, основанных на предложенных в работе моделях, с учетом загруженности очередей и требований к качеству обслуживания. В рамках предлагаемого метода реализуются так называемые иерархические очереди (HQ); адаптивно (на основе анализа состояния интерфейса) выбирается стратегия балансировки очередей, а также осуществляется дифференцированная обработка пакетов и распределение ПС канала связи между очередями с учетом приоритета и длины пакетов того или иного потока.

При этом на верхнем уровне метода (рис. 4) управление очередями производится на основе модели (1)-(2), результатом чего будет предварительное распределение потоков по очередям с закреплением за этими очередями определенной пропускной способности исходящего канала связи (физического интерфейса). На нижнем уровне метода (рис. 4) каждая конкретная очередь разбивается на подочереди, что является характерным решением для HQ, причем балансировка уже этих подочереди будет осуществляться на основе моделей (4)-(6) и (4), (7)-(13)

Решение достаточно сложных оптимизационных задач смешанного целочисленного линейного и нелинейного программирования (MILP, MILNP), возникавших в ходе исследований, осуществлялось с использованием инструментария пакета Optimization Toolbox системы Matlab 7, представленного программой `fminconset`. Предложенный метод особенно эффективно решает задачу балансировки очередей в условиях высокой загруженности интерфейса маршрутизатора, обеспечивая улучшение средней межконцевой задержки в среднем от 7-10% (при средних нагрузках и числе переключений 2-4) до 17-26% (в условиях, близких к перегрузке при числе переключений 6-16) при реализации предлагаемого метода по сравнению с механизмом взвешенных справедливых очередей – WFQ.

Список литературы:

1. Вегенша III. Качество обслуживания в сетях IP: Пер. с англ. – М.: Издательский дом «Вильямс», 2003. – 386 с.
2. Справочник по телекоммуникационным технологиям: Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 640 с.
3. Li Y., Panwar S., Liu C.J. On the Performance of MPLS TE Queues for QoS Routing // Simulation series. – 2004. – Vol. 36; part 3. – P. 170–174.
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 С. Али Модель динамического управления очередями и пропускной способностью канала связи на маршрутизаторах мультисервисной сети // Радиотехника: Всеукраїн. міжвед. науч.-техн. сб. – 2008. – Вып. 155. – С. 164–168.

6. Али С. Али, Симоненко А.В. Поточковая модель динамической балансировки очередей в MPLS–сети с поддержкой Traffic Engineering Queues [Электронный ресурс] // Проблемы телекоммуникаций. – 2010. – № 1 (1). – С. 59 – 67. – Режим доступа до журн.: http://pt.journal.kh.ua/2010/1/1/101_ali_balancing.pdf.

7. Лемешко А.В., Симоненко А.В. Математическая модель динамического управления канальными и буферными ресурсами на узлах телекоммуникационной сети // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2009. – Вып. 156. – С. 36-41.

8. Лемешко А.В., Али С. Али, Семеняка М.В. Разработка и исследование потоковых моделей динамической балансировки очередей на маршрутизаторах мультисервисной телекоммуникационной сети // Системи обробки інформації. – 2011. – Вип. 8(98).– С. 218-223.

9. Teixeira M.A., Guardieiro P.R. A new and efficient adaptive scheduling packets for the up-link traffic in WiMAX networks // EURASIP Journal on Wireless Communications and Networking. – 2011. – Vol. 2011, No. 1. – P. 113-123.

10. Annadurai C. Review of Packet Scheduling Algorithms in Mobile Ad Hoc Networks // International Journal of Computer Applications. – 2011. – Vol. 15, No. 1. – P. 7-10.

11. Jandaeng C., Suntiamentut W., Elz N., Review PSA. The Packet Scheduling Algorithm for Wireless Sensor Networks // GRAPH-HOC International journal on applications of graph theory in wireless ad hoc networks and sensor networks. – 2011. – Vol. 3, No. 3. – P. 1-12.

12. Ala'a Z. Al-Howaide, Ahmad S. Doulat, Yaser M. Khamayseh. Performance evaluation of different scheduling algorithms in WiMax // IJCSEA International Journal of Computer Science, Engineering and Applications. – 2011. – Vol. 1, No. 5. – P. 81-94.

13. Li Y. Panwar S. Liu C.J. On the Performance of MPLS TE Queues for QoS Routing // Simulation series. – 2004. – Vol. 36; part 3. – P. 170–174.

14. Huerta M. Buffer Capacity Allocation: A method to QoS support on MPLS networks / M. Huerta, J. J. Padilla, X. Hesselbach, Ramon Fabregat, O. Ravelo // Proc. EATIS2006 – Euro American Conference on Telematics and Information Systems, 2006. – P. 14–28.

15. Клейнрок Л. Вычислительные системы с очередями. – М.: Мир, 1979. – 600 с.

16. Клейнрок Л. Теория массового обслуживания. – М.: Машиностроение, 1979. – 432 с.

17. Лемешко А.В., Семеняка М.В., Али С. Али. Особенности проведения лабораторного эксперимента по исследованию процессов управления очередями в телекоммуникационных сетях // Цифрові технології. – 2012. – Вип. 10. – С. 5-17.