

УДК 621.391

Удосконалення алгоритму керування інформаційними потоками у вузлах телекомунікаційних мереж

М. А. Скулиш

Севастопольський національний технічний університет, Україна

Высокие требования к качеству передачи информационных потоков побуждают ученых разрабатывать все новые и новые методы повышения эффективности работы телекоммуникационных сетей. В рамках данной статьи были исследованы показатели QoS, способы распределения сетевого ресурса и приведен новый алгоритм обслуживания очередей. Необходимость в новом алгоритме возникает в мультисервисных сетях, где необходимо передавать высокоприоритетные информационные потоки. Такой алгоритм исключает эрланговские потери пакетов потоков заданного типа и обеспечивает эффективную передачу трафиков низкого приоритета.

Ключевые слова: телекоммуникационные сети, мультисервисные сети, алгоритм обслуживания очередей, эрланговские потери, трафик низкого приоритета.

Високі вимоги до якості передачі інформаційних потоків спонукають науковців розробляти все нові й нові методи підвищення ефективності роботи телекомунікаційних мереж. В рамках даної статті були досліджені показники QoS, способи розподілу мережевого ресурсу та наведений новий алгоритм обслуговування черг. Потреба у новому алгоритмі виникає у мультисервісних мережах, де необхідно передавати високопріоритетні інформаційні потоки. Такий алгоритм виключає ерлангівські втрати пакетів потоків заданого типу та забезпечує ефективну передачу низько пріоритетних трафіків.

Ключові слова: телекомунікаційна мережа, мультисервісна мережа, алгоритм обслуговування черг, ерлангівські втрати, низько пріоритетний трафік

Requirements to quality of transmission of information treads induce the research workers to develop all new and new methods of increase the efficiency of work of telecommunication networks. In this article the QoS indexes were explored, methods of distributing of network resource. A new algorithm of maintenance of turns was presented in the article. A necessity in a new algorithm arises up in multiservice networks, where is nessusary to pass the high-expedited informative treads. Such algorithm exclude the Erlangs losses of packages of streams of the set type and provides the effective transmission of low-expedited traffics.

Keywords: telecommunication networks, multiservice networks, algorithm of service queues, Erlang loss, low priority traffic

1. Вступ

Телекомунікаційні технології, як і інші технологічні процеси, змінюються під впливом вимог навколишнього світу. Головна мета цих змін, при виправданих витратах на модернізацію або побудову мережі, є збільшення якості та збільшення спектру інфомаційно-телекомунікаційних послуг. Серед головних змін слід відмітити використання технології all-over-IP («все по IP») з конвергенцією різнорідного трафіку в рамках єдиної IP-мережі, підвищення вимог до швидкості передачі на всіх ділянках мережі, а також виникнення потреб користувачів у декількох видах послуг (телефонія, передача даних, відео) із заданим рівнем якості обслуговування [1].

2. Забезпечення якісного обслуговування абонентів

З огляду на обмеженість мережевих ресурсів (буферного простору на вузлах мережі, пропускної здатності трактів передачі, розрахункових потужностей та часу на прийняття керівних рішень) мультисервісна телекомунікаційна мережа повинна мати ефективні засоби забезпечення QoS. Термін «якість послуг» в рекомендації E.800 МСЕ розуміється як деяка інтегральна оцінка, що визначає ступінь задоволення користувача послугою зв'язку, що йому надається [5]. У мережах наступного покоління рішення задач QoS здійснюється на кожному з трьох рівнів у підсистемах адміністративного керування, керування технічною експлуатацією та динамічного керування [2]. Ефективність телекомунікаційних мереж, побудованих на принципах NGN, багато в чому залежить від коректності розв'язку задач QoS [3].

Якісне обслуговування забезпечується різними способами. Мета забезпечення якості обслуговування абонентів у ТКМ – надання гарантованих та диференційованих послуг у масштабах мережі одного оператора. В рамках даної статті розглядається IP-мережа. Диференціація послуг обумовлює наявність декількох рівнів якості обслуговування, кожному з яких відповідає власна архітектурна модель забезпечення функцій QoS.

Типове рішення задач QoS охоплює наступні області [4]:

- класифікація прикладного програмного забезпечення з призначенням пріоритетів та диференціюванням трафіку;
- профілювання інформаційного трафіку;
- обмеження інформаційних потоків, що надходять від користувачів;
- керування чергами із встановленням послідовності обробки пакетів на мережених вузлах;
- маршрутизація мереженого трафіку.

Класифікувати основні послуги, що надаються в мультисервісних мережах в рамках забезпечення QoS можна по трьох основних характеристиках [5]:

- чутливість до величини пропускної здатності, що надається;
- чутливість до затримок;
- чутливість до втрат.

Кількісно ступінь чутливості прикладного програмного забезпечення оцінюється по відповідних показниках якості обслуговування інформаційних трафіків. До основних показників QoS відносяться показники продуктивності (швидкісні варіанти), показники часової прозорості (часові показники), показники семантичної прозорості (показники надійності та достовірності).

До основних показників продуктивності сети відносять ефективну, пікову, стійку та мінімальну швидкості передачі, що вимірюють як правило в біт/с. Мінімальне значення продуктивності зазвичай гарантується поставщиком послуг, який в свою чергу, повинен мати гарантії від мережного провайдера. Параметри пов'язані ефективною швидкістю передачі можуть бути визначені через дескриптор трафіку IP-мережі, що описаний в рекомендаціях МСЕ-T.1221.

Другий важливий критерій класифікації застосувань за типом трафіку їх чутливість до затримок пакетів. В рекомендації МСЕ-T.1540 затримка є основним параметром, що характеризує доставку пакетів IP-мережі, і виражається через параметр затримки IPTD (IP packet transfer delay), що

визначається як час доставки пакету між джерелом та отримувачем для всіх пакетів як успішних так і пакетів з помилками.

3. Огляд засобів розподілу каналних та буферних ресурсів в ТКМ

У випадках, коли мережа близька до стану перевантаження, для рішення задач QoS для окремого потоку інформаційного трафіку використовуються засоби керування чергами. Умовно їх можна поділити на алгоритми обслуговування черг та алгоритми превентивного обмеження черг. На практиці найбільше поширення отримали наступні алгоритми:

- PQ (Priority Queuing) – алгоритм обслуговування черг з абсолютним пріоритетом;
- CBQ (Class-Based Queuing) – алгоритм обслуговування черг на основі класів трафіків, що передбачає виділення всім класам номінальної пропускної здатності;
- WFQ (Weighted Fair Queuing) – зважений алгоритм рівномірного обслуговування черг на основі розрахунку порядкового номеру пакету, який передбачає збільшення або зменшення розміру черги в залежності від рівня пріоритету;
- DWFQ (Flow-Based Distributed WFQ) – розподілений зважений алгоритм рівномірного обслуговування черг на основі потоку;
- CBWFQ (Class-Based Distributed WFQ) – розподілений зважений алгоритм рівномірного обслуговування черг на основі класу, що дозволяє вказати потрібну мінімальну пропускну здатність для кожного класу трафіку;
- WRR (Weighted Round Robin) – зважений алгоритм кругового обслуговування;
- DRR – алгоритм кругового обслуговування з дефіцитом.

Алгоритми обслуговування черг в IP-мережах та ATM-мережах є одним з рішень задач QoS шляхом завчасного встановлення порядку використання каналних та буферних ресурсів мережі тобто сам процес розподілу має чітко виражений статистичний характер.

Пріоритетне обслуговування черг дозволяє гарантувати надання всієї смуги пропускання трафіку, необхідному для рішення критично важливих задач при ігноруванні решти трафіків. Для використання замовленого обслуговування черг гарантується надання визначеної смуги пропускання трафіку, необхідному для виконання критично важливих задач і в той самий час враховуючи решту трафіків.

Алгоритми превентивного обмеження черг приймаються для регулювання довжини черги. Є наступні алгоритми: алгоритм довільного раннього виявлення – RED (Random Early Detection), зважений алгоритм довільного раннього виявлення – WRED (Weighted RED), зважений алгоритм довільного раннього виявлення на основі потоку – Flow WRED, алгоритм явного повідомлення про перевантаження ECN (Explicit Congestion Notification), алгоритм вибіркового відкидання пакетів - SPD (Selective Packet Discard) .

В основному всі перераховані засоби відносяться до засобів локального керування трафіком та мережевими ресурсами. В їх основі покладені переважно децентралізовані алгоритми та механізми керування ресурсами, що реалізуються

на окремих вузлах ТКМ. Варто також відмітити, що недоліком всіх описаних засобів керування та боротьби з перевантаженнями є реалізація статичної стратегії розподілу каналних ресурсів. Крім того, в процесі керування не координуються рушення, отримані на окремих мережевих вузлах.

До другої категорії засобів QoS інформаційних трафіків відносяться засоби загальносистемного керування – маршрутизація, розподіл каналних та буферних ресурсів в інтересах інформаційних потоків користувачів. В цьому випадку керування носить розподілений характер, передбачаючи високий рівень узгодженості у використанні мережевих ресурсів. Важливість та першочерговість удосконалення засобів загальносистемного керування полягає у тому, що саме вони цілком визначають ефективність роботи засобів першої групи.

4. Удосконалення алгоритму обслуговування черг

В статті представлений оригінальний алгоритм керування чергами інформаційних заявок, що базується на одному з відомих алгоритмів WRR (Weighted Round Robin - зважений механізм кругового обслуговування) черг. Розробка алгоритму переслідує мету забезпечення максимальної доступності високопріоритетної послуги зв'язку при збереженні ефективного розподілу мережного ресурсу, який забезпечується зваженим механізмом кругового обслуговування черг.

Зважений механізм кругового обслуговування являє собою розширення планувальника кругового обслуговування, відповідно до якого кожному потоку трафіку призначається своя вага. Алгоритм WRR обробляє потік трафіку пропорційно до його ваги. Найбільш оптимальним чином WRR-планувальник погоджується з механізмом комутації ATM (Asynchronous Transfer Mode – режим асинхронної передачі), відповідно до якого пакет представляється у вигляді ячеек, а алгоритм WRR використовується для обробки черг, що складаються ячеек. Кожній черзі виділяється частина смуги пропускання інтерфейсу у відповідності до ваги потоку трафіку, на залежно від розміру пакету.

Задача оптимізації передачі інформаційних потоків у вузлі ТКМ є особливо актуальною враховуючи динамічну структуру широкосмугової цифрової мережі інтегрального обслуговування. Крім того, при представленні послуг високошвидкісної передачі даних актуальною також є задача підвищення стійкості її роботи при перевантаженнях мережі високо пріоритетними трафіками (відео даними та голосовим трафіком).

Вимоги до якості обслуговування абонентів диктують збільшення ймовірності обслуговування заявок кожного виду трафіку та зменшення часу затримок трафіків при проходженні через ТКМ.

Для розв'язку поставленої задачі за допомогою математичної моделі необхідно знайти раціональну комбінацію трафіку передачі даних та голосового трафіку при незмінному головному пріоритеті відео трафіку.

5. Постановка задачі

Розглянемо наступний варіант системи масового обслуговування. До комутаційного центру надходять заявки трьох видів: заявка на передачу відео інформації, заявка на обслуговування голосового трафіку, заявка на обслуговування Інтернет трафіку. Така модель відображає характер поведінки системи з пріоритетами і дозволяє проаналізувати плюси і мінуси подібних систем, і розрахувати найбільш прийнятний інтервал затримки заявки на обслуговування перед тим, як вона буде прийнята або відкинута.

Черги з пріоритетами відносяться до дисциплін, при яких вхідним повідомленням присвоюється ряд параметрів (або випадково, або спираючись на деякі властивості повідомлень). Система позначень, що використовується, буде мати наступний вигляд: повідомлення з пріоритетом q_2 має переважне обслуговування порівняно з повідомленням з пріоритетом q_1 , якщо $q_2 > q_1$. В системах з фіксованим пріоритетом, які обговорюються зараз, це означає, що повідомлення r -ї групи пріоритету має більш високий пріоритет, ніж повідомлення $(r-1)$ -ї групи. Це правило пріоритетів порушується при процедурі «перший надійшов – перший обслужений» (FIFO).

Нехай має місце P різних класів пріоритету. Повідомлення пріоритетного класу p ($p=1, 2, \dots, P$) надходять пуассонівським потоком із інтенсивністю λ_p повідомлень в секунду: кожне повідомлення з класу пріоритету p має час обслуговування розподілений експоненційно з параметром $1/\mu_p$.

Для системи, що моделюється в рамках даної статті, будемо говорити про три пріоритетні групи: потоки відео інформації мають найвищий пріоритет $pV=3$, голосові заявки є пріоритетними над заявками Інтернет трафіку і $pVo=2$, заявки Інтернет трафіку є непріоритетними їх пріоритет $pD=1$.

6. Схема обробки заявок за удосконаленням алгоритмом

Новий алгоритм обслуговування заявок передбачає наступну схему обробки заявок. Нехай заявки, що надійшли, утворюють черги по типам заявок, тобто для прикладу, який розглядається, утворюються три черги: черга голосових заявок, черга відео заявок та черга заявок на передачу Інтернет даних. При обслуговуванні черг за базовий механізм береться зважений механізм кругового обслуговування, тобто відповідно до класу трафіку черга отримує вагу на основі якої для обслуговування трафіку виділяється смуга пропускання у відповідному КЦ.

З метою підвищення якості обслуговування та мінімізації втрати пакетів інформації вводиться контроль за часом перебування в черзі пріоритетних заявок. Тобто задається інтервал часу Δt_m (t – пріоритетні групи) протягом якого заявки пріоритетного, наприклад голосового (Δt_{Vo}), трафіку чекають в черзі голосового трафіку. Коли час Δt_{Vo} минає, пакети з простроченим часом некваптивного обслуговування терміново ставляться на перші позиції в чергу Інтернет трафіку для здійснення термінової передачі.

Аналогічно з заявками на передачу відео інформації. Задається час Δt_V , інтервал часу протягом якого заявки відео трафіку чекають в черзі свого класу

трафіку. При закінченні часу Δt_V , пакети з простроченим часом неквапливого обслуговування терміново ставляться в чергу трафіку з меншим пріоритетом.

Узагальнений алгоритм обслуговування трафіку для моделі системи з m пріоритетними групами.

Утворюється m черг відповідно до пріоритетної групи.

Для пріоритетних трафіків задаються $(t-1)$ інтервалів часу Δt_m .

На початку кожної ітерації передачі інформації для кожного пакету пріоритетних черг оцінюється час перебування заявки в черзі,

Якщо час перебування в черзі більше ніж Δt_m , тоді ці пакети класифікуються як прострочені та ставляться в першу позицію найкоротшої черги, що має пріоритет менший t . Такі дії виконуються послідовно для всіх прострочених пакетів всіх $(t-1)$ пріоритетних груп.

Такий алгоритм дозволяє ефективно передавати заявки пріоритетних типів трафіку, при цьому не завдаючи суттєвої шкоди для передачі непраіоритетних заявок. До того ж можливі різні варіанти настройки даного алгоритму. Це стосується черги яку буде обрано для переставлення прострочених заявок, визначимо їх як черги наслідування прострочених заявок. Є ще один варіанти перестановок: обирається група непраіоритетних черг наслідування, час очікування заявок яких необмежений, при оцінці місця для перестановки саме серед цієї групи обирається найкоротша черга в яку будуть переставлені пріоритетні заявки.

За технологією АТМ, що найбільш відповідає вимогам широкосмугових цифрових мереж інтегрального обслуговування (ШЦМЮ), передача інформаційних потоків здійснюється блоками фіксованої довжини – ячейками. Кожен інформаційний блок передається окремо, в його заголовку прописано адреса відправника, адреса отримувача, пріоритет, поле типу сервісу (Type of Service) тощо. Таким чином інформаційний блок АТМ є самостійною одиницею для передачі.

Система масового обслуговування, що моделюється, є системою з втратами. Тому важливо визначити кількість місць у черзі (обмеження буфера). Для ефективної роботи затримка заявки в комутаційному центрі не повинна перевищувати деяке максимальне значення затримки трафіку, щоб зберегти актуальність інформації, особливо це стосується інтерактивного прикладного програмного забезпечення, трафіків голосу та відео.

Отже, розміри черг повинні бути такими, щоб виконувалась умова, що час який буде чекати заявка, що останньою стала в чергу, повинен бути не більше $\Delta t_{m \max}$ (часу максимальної затримки рафіку t -ї пріоритетної групи), для знаходження кількості ячеек в черзі необхідно дотримуватись значення границі:

$$N_{slot}(m) = \lim_{t_{N_{slot}(m)} \rightarrow \Delta t_{m \max}} v_m t_{N_{slot}(m)} = v_m \Delta t_{m \max},$$

де $N_{slot}(m)$ - кількість позицій в черзі t -ї пріоритетної групи при комутаційному центрі (довжина черги), $t_{N_{slot}(m)}$ - час обробки обслуговуючими пристроями комутаційного центру всіх $N_{slot}(m)$ пакетів з черги t -ї пріоритетної групи, v_m - швидкість обслуговування заявок t -ї пріоритетної групи.

Наприклад, для заявок відео т рафіку

$$N_{slot}(V) = v_V \Delta t_{V \max} .$$

де $N_{slot}(V)$ - кількість позицій в черзі при комутаційному центрі (довжина черги) для відео т рафіку.

Після того, як розмір черг був встановлений є потреба роз'яснити поведінку системи в цілому. При надходженні заявок до комутаційного центру вони ставляться в чергу відповідно до пріоритетної групи. Як вже було з'ясовано для пріоритетних заявок задаються інтервали часу Δt_m та $\Delta t_{m \max}$, на основі даних про них формуються черги. Відомо, що $\Delta t_m < \Delta t_{m \max}$. Заявки, що знаходяться в черзі довше Δt_m вважаються простроченими. Для того, щоб на кожній ітерації не виявляти прострочені заявки (виконання зайвих дій), є можливість ставити мітки тим заявкам, що потрапили в хвіст черги і явно стануть простроченими (отримають метку lim), якщо будуть послідовно обслуговуватися в своїй черзі. Ячейки черги, що заповнюються заявками, які гарантовано стануть простроченими, можна розрахувати за формулою:

$$N_{slot}^{lim}(m) = v_m \Delta t_m ,$$

Тобто всі заявки, що потрапили на позиції черги в інтервалі $(N_{slot}^{lim}(m), N_{slot}(m)]$, отримають мітку lim (прострочені). Якщо запропонувати перенесення цих заявок одразу після їх надходження до черг наслідування, тоді збережеться цілісність задачі, суттєво спростяться розрахунки часу затримки повідомлень, та збільшиться кількість втрачених заявок з пріоритетних груп, тому що останні позиції черг пріоритетних заявок будуть порожніми.

На рис. 1 показана робота даного алгоритму для системи, що моделюється, де до обслуговуючого пристрою утворюються три черги з відео повідомлень (V), голосових повідомлень (Vo) та повідомлень Інтернет трафіку (D). Група черг наслідування складається з черги Інтернет трафіку. Повідомлення відео трафіку, що отримали мітку lim одразу займають першу позицію в черзі трафіку даних.

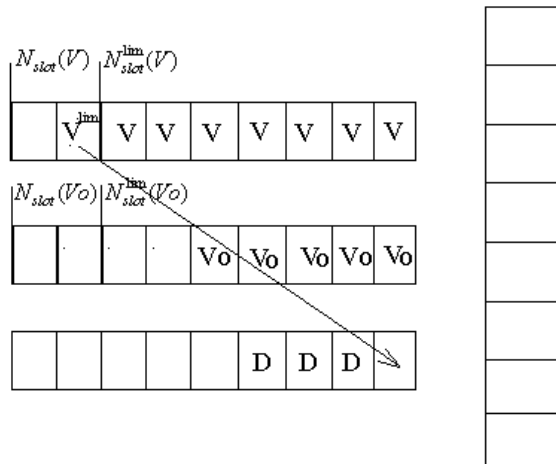


Рис. 1 Принцип роботи нового алгоритму обслуговування черг.

Група наслідування складається з черг неперіоритетних заявок або низько пріоритетних заявок. Якщо Інтернет трафік класифікований деяким чином, і заявки різних класів Інтернет трафіку мають різні пріоритети, тоді всі або деякі з них можуть потрапити до групи наслідування. В рамках даної роботи робиться припущення, що події пов'язані з потраплянням заявок до однієї з черг групи наслідування вважаються рівноймовірними.

7. Модифікований удосконалений алгоритм передачі обробки інформації в КЦ із збереженням послідовності високопріоритетних пакетів, які передаються

Слід зауважити, що удосконалений алгоритм, наведений вище, підходить лише для датаграмного режиму передачі голосової інформації, де допускається перемішування (зміна послідовності передачі) пакетів. Для режиму віртуального з'єднання для збереження послідовності пакетів, що передаються, даний алгоритм потребує деякої модифікації.

На рис. 2 показана логіка модифікованого удосконаленого алгоритму обслуговування черг.

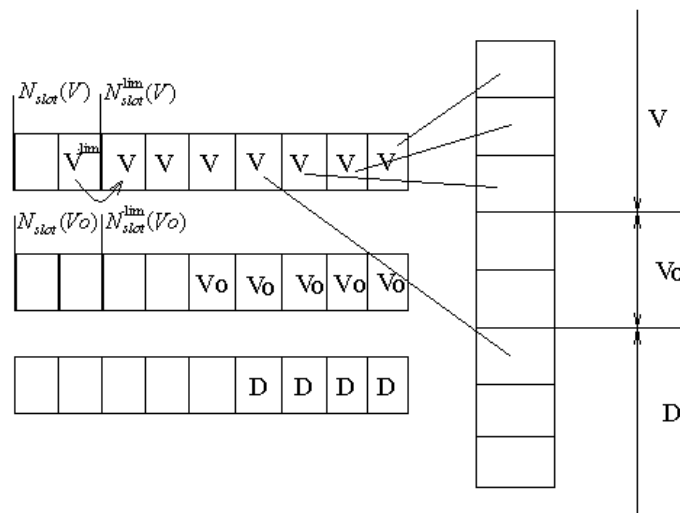


Рис. 2 Принцип роботи модифікованого нового алгоритму обслуговування черг.

Основна ідея полягає у тому, що при надходженні надлишкового пакету від високопріоритетного потоку заявок, що не зможе бути обслугований у час Δt_m (m – тип високопріоритетного потоку), в наступному циклі обслуговування заявок з m -ї черги будуть взяті не лише ті пакети, що обслуговуються обслуговуючими пристроями виділеними для даної пріоритетної групи, але й наступні по черзі пакети у кількості пакетів надлишкових пакетів, що надійшли за час попереднього циклу обробки, ці пакети будуть передані обслуговуючим пристроями черг групи наслідування.

8. Висновки

Високі вимоги до якості передачі інформаційних потоків спонукають науковців розробляти все нові й нові методи підвищення ефективності роботи телекомунікаційних мереж. В рамках даної статті були досліджені показники QoS, способи розподілу мережевого ресурсу та наведений оригінальний алгоритм обслуговування черг. Потреба у подібному алгоритмі виникає у мультисервісних мережах, де необхідно передавати високопріоритетні інформаційні потоки. Такий алгоритм виключає ергалгівські втрати пакетів потоків заданого типу та забезпечує ефективну передачу низькопріоритетних трафіків.

ЛІТЕРАТУРА

1. В.А. Фрейнкман. Интеллектуальные услуги и оборудование интеллектуальных сетей: взгляд за год // Вестник связи. – 2005. – №6.
2. Концепція конвергенції телефонних мереж і мереж з пакетною комутацією в Україні. – К.: Державний комітет зв'язку та інформатизації України, 2003. – 47 с.
3. Шварцман В.О. Интеграция в электросвязи. – М.: Агентство ИРИАС, 2001. – 167 с.
4. Кучерявый Е.А. Управление трафиком и качество обслуживания в сети Интернет. – М.: Наука и Техника, 2004. – 336 с.
5. Вегенша Ш. Качество обслуживания в сетях IP / Пер. с англ.. - М.: Издательский дом «Вильямс», 2003. - 386 с.
6. Харитонов и др. Качество обслуживания и эффективное использование ресурсов в мультисервисных сетях // Вестник связи. – 2004. – №12. – С. 48-55.

Надійшла 31.03.2009.

© Скулиш М. А., 2009