

УДК 004.713, 004.715

Ю.В. Каун

Луцький інститут розвитку людини ВМУРОЛ "Україна"

МЕХАНІЗМИ ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ОБСЛУГОВУВАННЯ В КОМУТАТОРАХ

В статті були проаналізовані характеристики трафік – контрактів, параметри якості обслуговування, механізми їх забезпечення, методи комутації, схеми буферизації та розподілу пам'яті, механізми черг, принципи використання буферної пам'яті.

Ключові слова: алгоритм, буфер, комутатор, комутація, механізм, пам'ять, якість обслуговування.

Рис. 4. Табл. 1. Літ 9.

Вступ

Великий успіх інформаційних мультимедійних технологій, Інтернету спричиняє неконтрольоване збільшення розміру трафіку, стимулює безпрецедентні вимоги до місткості базових мереж. Основними тенденціями розвитку мереж є: їх конвергенція, збільшення вузлів, уніфікація мережевих технологій, тяжкість трафіку, розширення функцій комунікаційних пристроїв.

Інфраструктура широкосмугових комунікацій базується на взаємозв'язаних мережах різного масштабу [1, с. 2]. Незалежно від протоколів верхнього рівня - на найнижчому знаходяться високошвидкісні комутатори (switches). Комутатор здійснює комутацію інформаційних потоків, з'єднуючи відповідні вхідні та вихідні порти. У найзагальнішому вигляді завдання комутації - з'єднання кінцевих вузлів через мережу транзитних вузлів - може представлятись послідовністю взаємозв'язаних часткових завдань: визначення інформаційних трафіків; визначення маршрутів для трафіків; повідомлення про знайдені маршрути вузлам мережі; просування - розпізнавання трафіків і локальна комутація на кожному транзитному вузлі; мультиплексування і демультіплексування трафіків, забезпечення необхідної якості обслуговування (Quality of service, QoS).

Сучасна комутація використовує високошвидкісні інтерфейси з продуктивністю внутрішньої комутаційної матриці в декілька сотень Гбіт/с [2, с. 107]. Головними вимогами до комунікаційного обладнання є збільшення пропускної спроможності та поліпшення інших його характеристик і мережі в цілому. Необхідність забезпечення можливості статистичного мультиплексування потоків, що проходять через комутаційні системні модулі, передача різних видів трафіку з несхожими вимогами до кількісних характеристик функціонування мережі є непростими завданнями. Комутатори складаються з інтегрованих модулів, які передають потік даних, буферизують, маршрутизують, управляють трафіком, окремими з'єднаннями, мережею.

При роботі в реальному масштабі часу на межі продуктивності потрібно буферизувати дані, що надходять з багатьох каналів, одночасно з комутацією та відправкою раніше прийнятих даних при забезпеченні необхідної якості обслуговування. Організація пам'яті та механізми обслуговування черг виходять на перший план.

Основною метою мережевого управління є забезпечення заданих показників якості обслуговування. Згідно ІТУ-Т E.800 QoS - це певна інтегральна оцінка, яка визначає ступінь задоволеності користувача наданою йому послугою зв'язку. Це визначення уточнене в рекомендації E.860: "Якість обслуговування - ступінь відповідності обслуговування, надаваного користувачеві постачальником, угоді між ними". Це надає ще більшу важливість угодам (трафік-контрактам) SLA між користувачами й постачальниками послуг. Фахівцями компанії Cisco уведено таке визначення терміну "якість обслуговування" - "Здатність мережі забезпечити необхідний сервіс заданому трафіку в певних технологічних рамках". Відповідно до змісту RFC 2475 під сервісом варто розуміти набір характеристик передачі пакетів в одному напрямку одним або декількома мережними маршрутами.

Постановка проблеми

Експоненційне збільшення розміру трафіку при відставанні базової місткості мереж, тобто екстенсивного збільшення мережевих ресурсів (буферного простору в вузлах комутації, пропускної здатності каналів передачі, розрахункових потужностей, часу прийняття керуючих дій), змушує шукати ефективні механізми забезпечення якості обслуговування як комплексного програмно-апаратного рішення в комутаторах.

Аналіз останніх досліджень і публікацій

Дану проблему вирішують шляхом вдосконалення апаратної складової – використанням буферної сортувальної пам'яті чи сортувальних мереж [3, с. 81-86].

Програмний шлях полягає в використанні комбінацій базових дисциплін обслуговування черг на мережевому рівні [4, с. 134-140] та поліпшенні алгоритму керування інформаційними потоками на базі WRR (Weighted Round Robin - зважений механізм кругового обслуговування) у вузлах телекомунікаційних мереж [5, с. 237-245].

Постановка завдання

Завданнями даного дослідження є аналіз параметрів QoS, механізмів їх забезпечення, аналіз методів комутації, схем буферизації та розподілу пам'яті, механізмів, формулювання вимог до механізмів обслуговування черг в мережевих вузлах і можливостей їх використання з буферною сортувальною пам'яттю.

Основний матеріал дослідження

Відповідно до рекомендації ІТУ-Т E.860, угода про рівень послуг (SLA) – це формальна угода між двома й більше об'єктами права, яка була досягнута після узгодження характеристик послуги, відповідальності й пріоритету кожної зі сторін. Постачальник послуг має гарантувати, що користувач одержить, щонайменше, обговорений рівень QoS. SLA є контрактом на обслуговування між провайдером і користувачем, що гарантує певні характеристики якості наданих послуг. У договорі SLA визначаються:

- плата за обслуговування залежно від обраного рівня сервісу;
- параметри QoS для даного рівня (середня та (або) максимальна затримка та її варіації, пропускна здатність, максимальний час відновлення мережі після аварій і т.ін.);
- методи вимірювання вищевказаних параметрів;
- штрафні санкції за незабезпечення необхідного рівня QoS;
- будь-які інші додаткові статті за взаємною згодою.

У цілому концепція укладання угоди по QoS є ієрархічною моделлю (рис. 1).

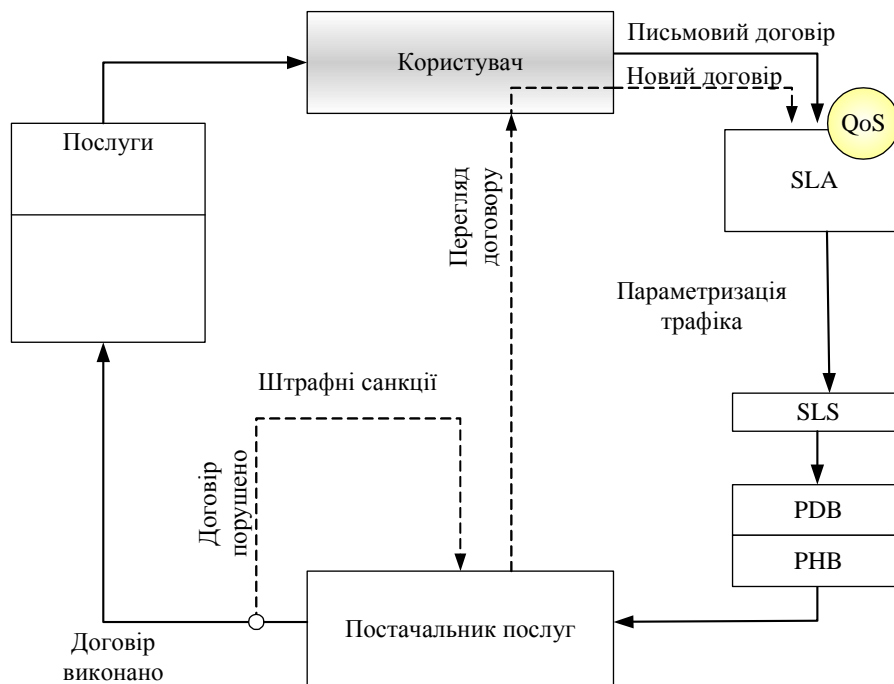


Рис. 1 Концепція укладання договору щодо якості обслуговування
Авторська розробка

Договір SLA у моделі займає найвищий рівень абстракції в специфікації послуги. Параметризація трафіка також виділяється в самостійну підмножину SLS (Service Level Specification). При статичному SLS користувач може передавати дані в будь-який час. При динамічному SLS користувач повинен використати сигнальний протокол для запиту необхідних ресурсів мережі й обробки запитів SLS. Блок PHB (Per-Hop Behavior) – є комбінацією функцій покрокової маршрутизації, класифікації, обробки черг і методів відкидання пакетів на кожному

кроці. Блок PDB (Per-Domain Behavior) реалізує концепцію покрокової маршрутизації на рівні домену.

Кожен тип PDB має кількісні і якісні атрибути, які можуть використовуватися для оцінки параметрів обслуговування пакетів заданого типу PDB. Атрибути обчислюються й вимірюються практичним шляхом й описують такі параметри як швидкість примусового відкидання пакетів, пропускну здатність, верхнє й нижнє значення параметра затримки за деякий проміжок часу. Атрибути бувають довгострокові й короткострокові. Довгострокові характеризують агрегований трафік за досить великий проміжок часу, короткострокові описують допустимі флуктуації параметрів щодо середніх значень.

Існують наступні сервісні моделі QoS:

- Best Effort Service. Негарантована доставка. Абсолютна відсутність механізмів QoS. Використовуються усі доступні ресурси мережі без якого-небудь виділення окремих класів трафіку і регулювання. Вважається, що кращим механізмом забезпечення QoS є збільшення пропускну спроможності. Однак, деякі види трафіку (наприклад, голосовий) дуже чутливі до затримок пакетів і варіації швидкості їх проходження. Модель Best Effort Service навіть за наявності великих резервів допускає виникнення перевантажень у разі різких сплесків трафіку. Тому були розроблені і інші підходи до забезпечення QoS.

- Integrated Service (IntServ, RFC 1633) - модель інтегрованого обслуговування. Може забезпечити наскрізну (End - to - End) якість обслуговування, гарантуючи необхідну пропускну спроможність. IntServ використовує для своїх цілей протокол сигналізації RSVP. Дозволяє додаткам виражати наскрізні вимоги до ресурсів і містить механізми забезпечення цих вимог. IntServ можна коротко охарактеризувати як резервування ресурсів (Resource reservation).

- Differentiated Service (DiffServ, RFC 2474/2475) - Модель диференційованого обслуговування. Визначає забезпечення QoS на основі чітко певних компонентів, комбінованих з метою надання необхідних послуг. Архітектура DiffServ припускає наявність класифікаторів і формувачів трафіку на межі мережі, а також підтримку функції розподілу ресурсів в ядрі мережі в цілях забезпечення необхідної політики покрокового обслуговування (Per - Hop Behavior - PHB). Розділяє трафік на класи, вводячи декілька рівнів QoS. DiffServ складається з наступних функціональних блоків: граничні формувачі трафіку (класифікація пакетів, маркування, управління інтенсивністю) і реалізатори PHB політики (розподіл ресурсів, політика відкидання пакетів). DiffServ можна коротко охарактеризувати

Вирішення завдань QoS знаходиться в наступних областях: класифікація додатків з пріоритетизацією та диференціюванням трафіку; його профілювання; обмеження за потребою інтенсивності користувацького трафіку; управління чергами з послідовністю обробки пакетів в мережних вузлах; маршрутизація. Забезпечення QoS відбувається взаємопов'язано.

QoS (сервіс) описується наступними параметрами: параметри пропускну здатності (мінімальна, середня й максимальна швидкість передачі пакетів); параметри затримок передачі пакетів (середні й максимальні величини затримок і джитера); параметри надійності передачі (рівень втрат і викривлення пакетів).

Вимірювання зазначених параметрів здійснюється на деякому інтервалі часу. Чим він менше, тим жорсткіші вимоги ставляться до мережі, а отже, до всіх її елементів, оскільки забезпечення наскрізної QoS вимагає взаємодії всіх вузлів на шляху пакетів трафіка й визначається надійністю, функціональністю й продуктивністю "слабкої ланки".

Для опису номінальної пропускну здатності лінії зв'язку використовується термін смуга пропускання (bandwidth). Затримка при передачі пакета (packet delay) чи латентність (latency) є сумарною величиною, що об'єднує затримки: серіалізації, поширення, очікування в черзі, комутації, формування трафіка, мережіві. Складові затримки під час передачі пакетів наведені в таблиці 1.

З виникненням в мережі областей перевантажень, затримки очікування в черзі, переважають інші види затримок і призводять до виникнення тремтіння (packet jitter) – різниці у величині сумарної затримки при передачі різних пакетів одного потоку. Рівень втрат пакетів (packet loss) визначається як частка пакетів, відкинутих мережею під час передачі протягом визначеного інтервалу часу. Основними причинами втрат пакетів є їх ушкодження та перевантаження мережі (кількість пакетів, що надходять, набагато перевищує верхню межу розміру вихідної черги). Відкидання пакетів може викликатися недостатнім розміром вхідного буфера.

Складові затримки під час передачі пакетів
 Авторська розробка

| Різновид затримки | Характер величини | Визначення | Місце виникнення |
|---|-------------------|---|--|
| Затримка серіалізації (serialization delay). | постійна | Час, необхідний для передачі всіх біт пакета у фізичне середовище. Є функцією розміру пакета і швидкості каналу (ширини смуги пропускання) | На виході будь-якого фізичного інтерфейсу. Звичайно дуже мала для каналу ТЗ і вище |
| Затримка поширення (propagation delay). | постійна | Час, необхідний для досягнення бітові інформації приймаючого пристрою на іншому кінці каналу. Залежить від середовища поширення (точніше від швидкості світла в даному фізичному середовищі) і від відстані | У будь-якому фізичному каналі. Звичайно дуже мала в каналах мереж рівня LAN і на коротких ділянках WAN мереж |
| Затримка в черзі (queuing delay). | непостійна | Час, витрачений пакетом на перебуванні в черзі в очікуванні подальшої передачі (вихідна черга) або в очікуванні можливості перетнути комутаційне поле (вхідна черга) (в очікуванні комутації) | Можлива на будь-якому вихідному інтерфейсі. Вхідні черги мало специфічні для маршрутизаторів, однак поширені в комутаторах мереж рівня LAN |
| Час пересилання або обробки (forwarding or processing delay). | непостійна | Час, необхідний для прийняття вхідного пакета і його обробки, доти поки пакет не буде поставлений у чергу для подальшої передачі | У будь-якому комутаторному обладнанні, включаючи маршрутизатори, комутатори LAN, Frame Relay, ATM |
| Затримка, пов'язана з формуванням трафіку (shaping delay). | непостійна | За умови здійснення формування трафіку, це час, на який пакети, що підлягають передачі, затримуються щоб уникнути втрат пакетів у середовищі Frame Relay або ATM | Скрізь, де включений механізм формування трафіку (shaping), найчастіше в маршрутизаторах, що передають пакети в мережі Frame Relay або ATM |
| Мережна затримка (network delay). | непостійна | Затримка, внесена компонентами абонентської мережі | У межах мережі провайдера послуг |

Задача розподілу ресурсів каналів між потоками даних вирішується встановленням порядку обслуговування пакетів і організацією черг. Черги та алгоритми їх обробки управляють перевантаженнями комутатора. Черги - області пам'яті, де групуються однопріоритетні пакети. Передача пакетів з черги здійснюється за алгоритмом її обслуговування. Застосування алгоритмів зводиться до забезпечення найкращого обслуговування високопріоритетного трафіку з гарантуванням низькопріоритетному відповідну уваги.

Механізми обслуговування черг класифікують за ознаками: реалізований принцип розподілу ресурсів (без розподілу ресурсів, пріоритетний розподіл шляхом застосування однойменного обслуговування, пропорційний розподіл шляхом кругового обслуговування черг і рівномірний розподіл шляхом реалізації максимінної схеми); надання гарантій за будь-якими параметрами мережного з'єднання (у термінах смуги пропускання або гарантованої затримки); принцип формування черг (формування черг за потоками або за класами); режим виконання (розподілений на процесорах VIP-плат або нерозподілений на центральному процесорі маршрутизатора).

Найчастіше в комутаторах застосовуються такі механізми обробки черг:

- FIFO (First In – First Out) є найпростішим механізмом обслуговування, за яким пакети передаються на вихід у порядку, в якому вони надійшли на вхід. Він є алгоритмом за

замовчуванням у всіх пристроях з комутацією пакетів. Переваги: простота реалізації та відсутність конфігурування. Недолік: неможливість диференційованої обробки пакетів різних потоків.

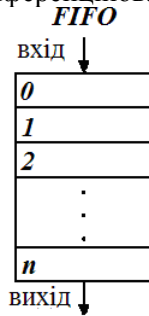


Рис. 2. Організація найпростішим механізмом обслуговування FIFO
 Авторська розробка

- пріоритетне обслуговування (Priority Queuing, PQ); припускає наявність чотирьох вихідних підчерг з високим, середнім, звичайним і низьким пріоритетом обслуговування. У межах черги пакети обробляються за FIFO. Пріоритетне обслуговування черг забезпечує високу якість обслуговування та мінімальний рівень затримок пакетів із черги з найвищим пріоритетом. Недоліки: необхідність ретельного контролю трафіку на етапі доступу в мережу з метою належного надання пріоритету; відсутність верхньої межі для кожного з рівнів пріоритету; високий ризик придушення низькопріоритетних потоків потоками з найвищим пріоритетом.

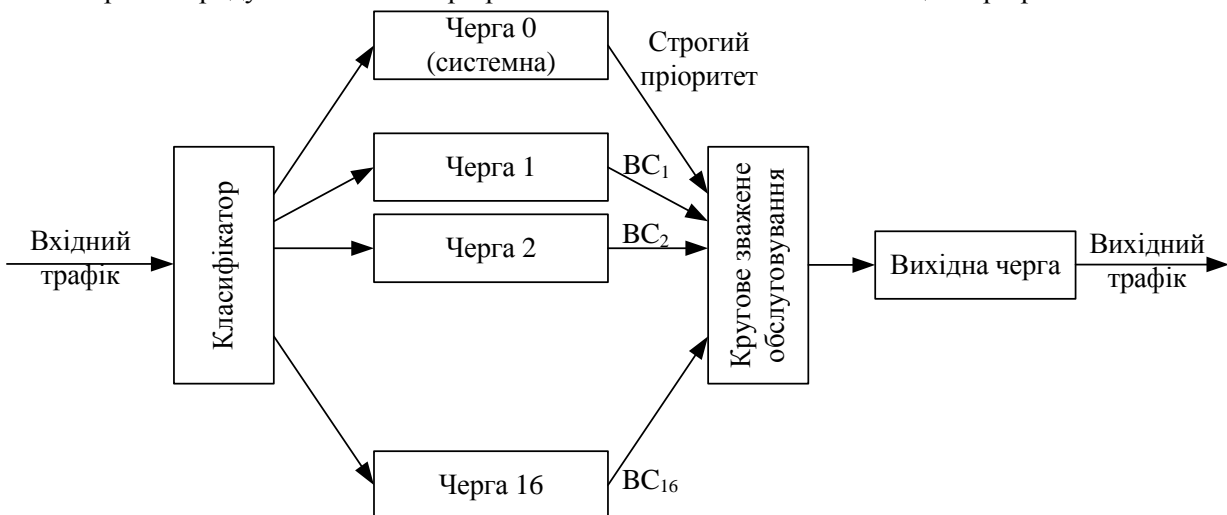


Рис. 3 Механізм замовленого обслуговування черг CQ.
 Авторська розробка

- замовлене обслуговування (Custom Queuing, CQ); надає адміністратору можливість вручну задати розподіл ресурсів, тобто розподілити смугу пропускання. Трафік поділяють на 16 черг (рис. 3). Додаткова нульова системна черга призначена виключно для обробки високопріоритетних пакетів (пакети підтримки з'єднання і управляючі пакети).

У рамках CQ черги 1–16 обслуговуються за круговим принципом, починаючи з першої. З кожною чергою пов'язаний свій лічильник байт (Byte Count, BC), який визначає кількість байтів, що має бути передана з цієї черги.

- справедливе (Fair Queuing, FQ); моделює узагальнену схему поділу процесорного часу (Generalized Processor Sharing, GPS) шляхом обчислення порядкового номера кожного отриманого пакета (службова позначка, яка визначає відносний порядок обробки пакетів). Значення лічильника циклів (Round Number, RN) відповідає кількості виконаних циклів побайтового планувальника кругового обслуговування в заданий момент часу. Лічильник циклів безпосередньо визначає значення порядкового номера пакета.

- на основі класу (Class Based Queuing, CBQ); забезпечує ієрархічний розподіл ресурсів і організації черг. Смуга пропускання каналу поділяється між декількома класами трафіку в заданих співвідношеннях. Кожному з класів у період перевантаження гарантується частка загальної смуги.

Використовується два планувальника: загальний планувальник (general scheduler) і планувальник розподілу ресурсів (link-sharing scheduler) (рис. 4).

Вхідний трафік класифікується та поділяється на декілька черг за задалегідь заданими правилами фільтрації. Загальний планувальник (WRR) визначає, з якої черги обслуговуватиметься наступний пакет. Черга вибирається, щоб гарантувати кожному класу трафіку щонайменше задану частку від сумарної пропускної здатності каналу (гарантована смуга). У рамках окремих черг використовується механізм FIFO.

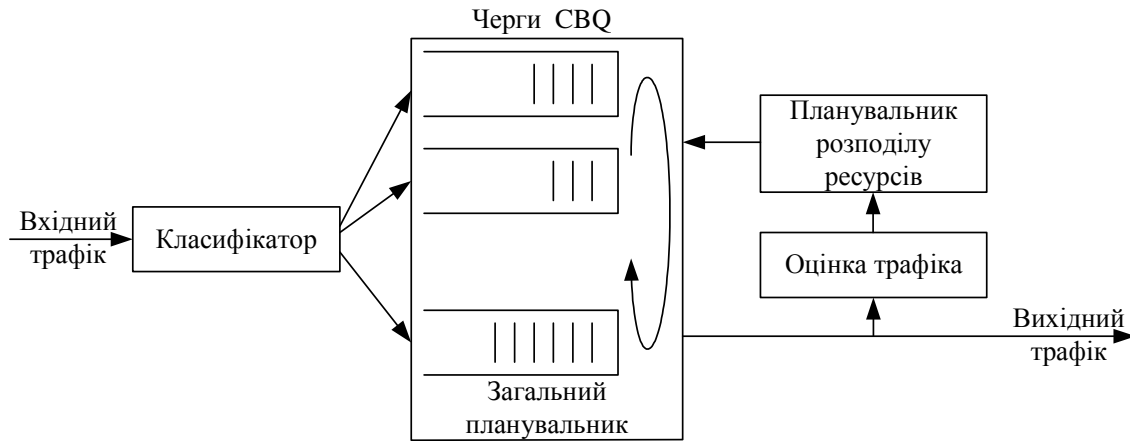


Рис. 4 Механізм СВQ
Авторська розробка

Вихідний трафік контролюється шляхом оцінки часу між сусідніми пакетами одного класу трафіку. Трафік вважається таким, що перевищує ліміт, якщо смуга, яку він використовує більше, ніж йому гарантується, відповідно трафік є таким, що укладається в ліміт, якщо він використовує смугу менше гарантованої. Якщо черга деякого класу трафіку виявиться порожньою, планувальник розподілу ресурсів розділить гарантовану даному класові смугу між іншими класами. При перевантаженні планувальник розподілу ресурсів переводить клас, що перевищив ліміт, у розряд пасивного і загальний планувальник тимчасово не обслуговує пакети з цієї черги.

- Їх модифікації (зважене справедливе (Weighted Fair Queuing, WFQ); зважене справедливе на основі класу (Class Based WFQ, CBWFQ); обслуговування з малою затримкою (Low Latency Queuing, LLQ); зважене кругове (Weighted Round-Robin, WRR) і його модифікації; кругове з дефіцитом (Deficit Round-Robin, DRR) і його модифікації).

Методи комутації та структуру комутатора поділяють на [7, с. 119]: з розподіленою пам'яттю; з розподіленим середовищем передачі (кільце, шина, подвійна шина); з повнозв'язною топологією; з просторовим поділом (з єдиним і множинними шляхами).

Ситуація, при якій декілька блоків вхідних даних одночасно мають бути зкомутувані на один вихідний порт називається вихідним конфліктом і знижує максимальну продуктивність. Комутаційне поле є внутрішньо блокуючим, якщо набір з N блоків даних, адресованих на N різних вихідних портів, здатний викликати конфлікти при передачі через це поле [8, с. 20]. Для всіх архітектур комутаційного поля необхідна буферизація. Найпоширенішими варіантами розміщення буферів в комутаторі можна вважати 4 наступних: організація черг на вході; організація черг на виході; внутрішня організація черги; вживання рециркуляційних буферів. Перевага надається організації черг на виході комутатора. Зазначимо, що тип комутатора однозначно визначає ефективні в поєднанні з ним типів буферної пам'яті, а тип буферної пам'яті задає можливі типи комутаторів.

Число і розмір буферів також важливі при розробці комутатора. Необхідно виділити п'ять схем розподілу буферної пам'яті [9, с. 109]: схема повного розподілу (CP); повнодоступна схема (CS); повнодоступна схема з індивідуальними обмеженнями на довжини вихідних черг (SMQ); неповнодоступна схема (SMA) є комбінацією CP і CS; неповнодоступна схема з індивідуальними межами (SMQMA).

У мікропроцесорній техніці використовують різні типи пам'яті, які можна класифікувати за способом доступу до даних [7, с. 309]:

- Пам'ять із впорядкованим доступом. Із пам'яті дані вибираються в порядку, який визначається її внутрішньою структурою.

- Пам'ять з довільним доступом. До/з довільної комірки в кожному такті може записуватись/зчитуватись дане, використовуючи адресу комірки на адресному вході пам'яті.
- Пам'ять з асоціативним доступом. Пошук даних здійснюється шляхом порівняння асоціативних ознак, які ставляться у відповідність кожному елементу даних, що зберігається в пам'яті.
- Пам'ять з програмованим доступом, коли записані до пам'яті дані зчитуються в наперед заданому порядку.
- Сортувальна пам'ять, коли дані записуються в пам'ять в порядку їх зчитування.

Для вибору оптимального механізму обслуговування черг він повинен: мати засоби диференціювання пакетів і визначення рівня обслуговування кожного пакета; гарантувати QoS шляхом розподілу ресурсів для кожного окремого потоку трафіку чи визначенням відносного пріоритету потоків; забезпечувати захист і рівномірну обробку всіх потоків трафіку з однаковим пріоритетом; легко реалізовуватись і підтримувати управління доступом для потоків, що потребують гарантованого надання ресурсів.

З апаратної точки зору використання буферної сортувальної пам'яті дозволяє розвантажити програмну складову та забезпечити необхідну масштабованість і відмово стійкість.

Висновки

В даному дослідженні були проаналізовані параметри QoS, механізми забезпечення заданих показників QoS, методи комутації, схеми буферизації та розподілу пам'яті, сформульовані вимоги до механізмів обслуговування черг в мережевих вузлах. Новизною є комплексне використання відповідних алгоритмів обслуговування черг та пристроїв планування комутації при використанні сортувальної пам'яті в комутаторах. При побудові комутаторів доцільно використовувати відповідні алгоритми обслуговування черг та пристрої планування комутації з використанням сортувальної пам'яті.

Перспективи подальших досліджень. Подальші дослідження направлені на синтез структур запам'ятовуючих пристроїв на основі буферної сортувальної пам'яті в комбінації з механізмами обслуговування черг.

1. Chao H. J. Broadband Packet Switching Technologies: A Practical Guide to ATM Switches and IP Routers / H. J. Chao, C. H. Lam, E. Oki. John Wiley & Sons Inc, 2001. - 459 с.
2. Cisco Systems, Inc. Руководство по технологиям объединенных сетей, 4-е изд.. : Пер. с англ. - М.: Издательский дом "Вильямс", 2005. - 1040 с.: ил.
3. Мельник А.О. Порівняльний аналіз типів пам'яті комп'ютера / А.О. Мельник, Д.Х. Аль Равашдех // Вісник Національного університету "Львівська політехніка: комп'ютерні системи та мережі. - Львів : Львівська політехніка, 2007. -№063.-с.81-86.
4. Верхола Б. Дисципліни обслуговування черг на мережевому рівні OSI / Б. Верхола. //Комп'ютерні технології друкарства. Львів. -2010. №24.-с. 134-140
5. Скулиш М. А. Удосконалення алгоритму керування інформаційними потоками у вузлах телекомунікаційних мереж / М. А. Скулиш //Вісник Харківського національного університету №863, 2009. –с. 237-245
6. Каун Ю.В. Постановка задач комутації в мережах / Ю.В. Каун, М.Я Максимович // Наукові нотатки. -Луцьк. ЛНТУ. - 2009. - №26. -с. 116-122.
7. Мельник А.О. Архітектура комп'ютера. / А.О. Мельник. - Луцьк: Волинська обласна друкарня, 2008. – 470 с.
8. Дунець Р. Б. Топології комп'ютерних систем / Р. Б. Дунець. - Львів: Ліга-Прес , 2007 - 46 с
9. Каун Ю.В. Буферна пам'ять комутаторів / Ю.В. Каун // Міжвузівський збірник "Комп'ютерно-інтегровані технології: освіта, наука, виробництво". –Луцьк. ЛНТУ. - 2011. - №5. -с. 108-113.