Міністерство освіти і науки України

Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Механізми та моделі забезпечення QoS в мультисервісних IP-мережах

(тема)

Виконав:

студент II курсу, групи СПзм-20-1

Грошев А.С.
(прізвище, ініціали)

Спеціальність

123 «Комп'ютерна інженерія»
(код і повна назва спеціальності)

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма

Системне програмування
(повна назва освітньої програми)

Керівник: доц. Колтун Ю.М.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ _____ Коваленко А.А.
(підпис) (прізвище, ініціали)

2023 р.

Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання

Кафедра електронних обчислювальних машин

Рівень вищої освіти другий (магістерський)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва)

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системне програмування
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
"____" _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту Грошеву Андрію Сергійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Механізми та моделі забезпечення QoS в мультисервісних IP-мережах

затверджена наказом по університету від " 25 " березня 2022 р. № 33 Стз

2. Термін подання студентом роботи до екзаменаційної комісії 14 травня 2022 р.

3. Вхідні дані до роботи 1) тип мережі – пакетна, протокол обміну інформацією – IP, концепція функціонування – мультисервісна (МсМЗ); 2) зробити огляд стандартів ITU-T щодо забезпечення QoS в IP-мережах; 3) проаналізувати: базову модель, механізми, функції і технології підтримки QoS в мультисервісних IP-мережах; 4) проаналізувати технологічні моделі забезпечення QoS в IP-мережах; 5) запропонувати загальну методику оцінки мережних характеристик якості обслуговування для цього типу мереж.

4. Перелік питань, що потрібно опрацювати у роботі
1) загальний аналіз підходів до оцінки впливу мережних характеристик на QoS в;
в мультисервісних IP-мережах на основі стандартів ITU-T;
2) обґрунтування вимог до QoS додатків різних типів у IP МсМЗ;
3) аналіз базової моделі, механізмів і функцій забезпечення QoS в IP-мережах;
4) аналіз архітектурних і функціональних принципів реалізації моделі IntServ;
5) аналіз архітектурних і функціональних принципів реалізації моделі DiffServ;
6) обґрунтування загальної методики оцінки мережних характеристик QoS в IP МсМЗ;
7) висновки.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів)

Слайд-презентація – 14 слайдів

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Основні поняття і стандарти ITU-T щодо		
	забезпечення QoS в IP МсМЗ;	26.03.22 – 06.04.22	
2	Обґрунтування вимог до QoS додатків різних		
	типів у IP МсМЗ;	07.04.22-11.04.22	
3	Аналіз базової моделі, механізмів, функцій і		
	технологій підтримки QoS в IP МсМЗ;	12.04.22-21.04.22	
4	Аналіз архітектурних і функціональних		
	принципів реалізації моделей IntServ і DiffServ;	22.04.22-28.04.22	
5	Загальна методика оцінки мережних		
	характеристик QoS в IP МсМЗ	29.04.22-04.05.22	
6	Оформлення матеріалів кваліфікаційної роботи	05.05.22-09.05.22	
7	Подання кваліфікаційної роботи керівникові		
	та її попередній захист	10.05.22-11.05.22	
8	Подання кваліфікаційної роботи на рецензування	12.05.22-13.05.22	

Дата видачі завдання 26 березня 2022 р.

Студент _____
(підпис)

Керівник роботи _____ доц. Колтун Ю.М.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 86 с., 11 рис., 6 табл., 2 дод., 20 джерело.

QoS, IP-МЕРЕЖА, NGN, ЯКІСТЬ ОБСЛУГОВУВАННЯ, Рек. ITU-T E.800, Рек. ITU-T I.350, Рек. ITU-T Y.1540, Рек. ITU-T Y.1541, Рек. ITU-T Y.1291, НЕГАРАНТОВАНА ДОСТАВКА, ДИФЕРЕНЦІЙОВАНЕ ОБСЛУГОВУВАННЯ, DIFFSERV, ГАРАНТОВАНЕ ОБСЛУГОВУВАННЯ, INTSERV

Метою кваліфікаційної роботи є дослідження і аналіз основних понять, визначень, стандартизованих механізмів та моделей, які гарантують необхідний рівень якості обслуговування в мультисервісних IP-мережах.

У ході виконання кваліфікаційної роботи наведені та обґрунтовані основні поняття, визначення та мережні характеристики якості обслуговування, проаналізовані стандарти ITU-T, механізми і відповідні їм технології та моделі забезпечення потрібної QoS у мультисервісних IP-мережах. Запропонована та обґрунтована загальна методика оцінки мережних характеристик якості обслуговування в мережах такого типу.

THE ABSTRACT

Master’s thesis: 86 pages, 11 figures, 6 tables,2 appendices, 20 sources.

QoS, IP-NETWORK, NGN, QUALITY OF SERVICE, ITU-T Rec. ITU-T E.800, ITU-T Rec. ITU-T I.350, ITU-T Rec. ITU-T Y.1540, ITU-T Rec. ITU-T Y.1541, ITU-T Rec. ITU-T Y.1291, UNSATISFIED DELIVERY, DIFFERENTIATED SERVICE, DIFFSERV, GUARANTEED SERVICE, INTSERV

The purpose of the qualification work is to study and analyze the basic concepts, definitions, standardized mechanisms and models that guarantee the required level of quality of service in multiservice IP networks.

In the course of the qualification work, the basic concepts, definitions and characteristics of quality of service are given and justified, ITU-T standards, mechanisms and corresponding technologies and models for ensuring the required QoS in multiservice IP networks are analyzed. A general methodology for estimating the network characteristics the quality of service in such networks is proposed and substantiated.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

AF PHB (Assured Forwarding PHB) – PHB-политика гарантованої доставки пакетів;

ATM (Asynchronous Transfer Mode) – асинхронний режим передачі інформації;

CBR (Constant Bit Rate) – постійна бітова швидкість;

CBQ (Class-Based Queuing) – черга (потік) у відповідності із класом;

COPS (Common Open Policy Service) – протокол розповсюдження політик;

DCM (Dispersion Compensation Module) – модуль компенсації дисперсії;

DiffServ (Differentiated Service) – диференційоване обслуговування;

DSCP (Differentiated Services Code Point) – поле коду диференційованих послуг (DS-байт – байт диференційованої послуги);

GoS (Grade of Service) – рівень обслуговування;

ECN (Explicit Congestion Notification) – явне повідомлення про перевантаження;

EF PHB (Expedited Forwarding PHB) - PHB-політика негайної передачі пакетів;

FIFO (First-In, First-Out) – «першим прийшов, першим вийшов»;

IETF (Internet Engineering Task Force) – група інженерних проблем Internet;

IntServ (Integrated Service) – інтегроване обслуговування;

IPDV (IP packet delay variation) – варіація затримки IP пакета;

IPLR (IP packet loss ratio) – коефіцієнт втрати IP пакетів;

IPTD (IP packet transfer delay) – затримка доставки IP пакета;

ITU-T (International Telecommunication Union – Telecommunication Standardization Sector) – Міжнародна спілка електрозв'язку – сектор стандартизації телекомунікацій;

MPLS (Multi-Protocol Label Switching) – багатопротокольна комутація за мітками;

NFS (Network File System) – мережна файлова система;

NGN (Next Generation Network) - мережа наступного покоління;

NP (Network Performance) – мережні характеристики;

OTU (Optical Transponder Unit) –транспондер;

QoS (Quality of Service) – якість обслуговування;

PHB (Per-Hop Behavior) – ухвалення рішення про просування пакета в кожному проміжному вузлі мережі DiffServ;

RED (Random Early Detection) – довільне раннє виявлення;

RSVP (Resource Reservation Protocol) – протокол резервування ресурсів;

RTD (Round-Trip Delay time) – кругова затримка;

SLA (Service Level Agreement) – угода про рівень обслуговування;

SLS (Service Level Specification) – специфікація рівня сервісу;

TC (Traffic Class) – клас трафіку;

TCP (Transmission Control Protocol) – протокол управління передачею;

ToS (Type of Service) – тип обслуговування;

UDP (User Datagram Protocol) – протокол дейтаграм користувача;

VBR (Variable Bit Rate) – змінна бітова швидкість;

VoIP (Voice over IP) – голос поверх IP (голосовий зв’язок через мережу IP);

VPN (Virtual Private Networks) – віртуальна приватна мережа;

WFQ (Weighted Fair Queuing) – зважений алгоритм рівномірного обслуговування черг;

WRED (Weighted Random Early Detection) – зважений алгоритм раннього довільного виявлення.

ВОЛЗ – волоконно-оптичні лінії зв'язку;

ЕМВВС – еталонна модель взаємодії відкритих систем;

КП – комутація пакетів;

ТО – технічне обслуговування;

ТМЗК – телефонна мережа загального користування;

ВСТУП

Сучасне світове суспільство характеризується проникненням новітніх інфокомунікаційних технологій у всі області свого життєзабезпечення. Це проникнення здійснюється за рахунок наймасштабнішого і всебічного розвитку різноманітних послуг, що надаються користувачам. Ці послуги стають невід'ємною частиною спілкування людей, забезпечують доступ до популярних розваг, допомагають реалізувати свої амбіції у самоствердженні і діловому зростанні особистості, просто є засобом заробітку, і т.ін. З іншого боку такий розвиток послуг призводить до зростання складності у їх реалізації та підвищенню вимог до ресурсів мереж зв'язку, на базі яких вони впроваджуються. Це у свою чергу створило передумови до об'єднання мереж і мережних технологій на основі процесів інтеграції і конвергенції у єдину універсальну інфокомунікаційну структуру, що працює на базі загальної транспортної платформи [1].

Результатом втілення у життя цих передумов було створення концепції побудови мереж наступного покоління (Next Generation Networks, NGN). Основною ідеєю цієї концепції є побудова універсальної мережі, яка б дозволяла переносити будь-які види інформації, а також забезпечувати можливість надання необмеженого спектра сучасних послуг. Найважливішу роль у реалізації цієї концепції відіграли послуги передачі даних, голосу і відео, такі як: IP-телебачення, надання відео за запитом, IP-телефонія, відео- та аудіо-конференцзв'язок, т.ін [1, 2].

Звідси, узагальнюючи вище викладене, NGN можна визначити, як концепцію побудови мереж зв'язку, які забезпечують надання необмеженого набору послуг з гнучкими можливостями щодо їх управління, персоналізації і створенню нових послуг за рахунок уніфікації мережних рішень, що передбачає реалізацію загальної універсальної транспортної платформи з розподіленою комутацією і її конвергенцію з традиційними мережами

зв'язку. Тобто фізичну основу NGN становить універсальне транспортне середовище із розподіленою комутацією пакетів. Як правило, в якості транспортної платформи використовуються технології, що працюють за протоколом IP, тому що на цей час більшість додатків і сервісів орієнтовані саме на підтримку цього протоколу. В результаті стає можливим використовувати ту ж саму логіку послуги незалежно від способу доступу до транспортної платформи [2, 3].

Таким чином з практичної точки зору реалізації NGN, вони являють собою єдину телекомунікаційну структуру, що побудована на платформі деякої IP-мережі, та здатна надавати різноманітні послуги на основі передачі голосу, відео і даних. Такого типу телекомунікаційна структура узагальнено називається мультисервісною мережею зв’язку, а її привабливість для користувачів полягає у зниженні собівартості сервісів, що надаються, як за рахунок об’єднання самих мережних технологій, так і через постійний розвиток можливостей нових сучасних послуг. Але при цьому слід зазначити, що з технічної точки зору у разі організації мультисервісної мережі виникає багато проблем, таких як забезпечення необхідних параметрів за імовірністю втрати пакетів, часові затримки передачі, джиттер т.ін. Це говорить про те, що потрібно приділити значну увагу щодо забезпечення параметрів якості обслуговування (Quality of Service, QoS) [3].

Зазначені проблемні питання пов’язані з тим, що протокол IP первісно не забезпечував механізмів гарантованої доставки даних кінцевому користувачеві і відповідно не призначався для обміну інформацією у реальному часі. Адже пакети одного і того ж потоку даних маршрутизуються по мережі незалежно один від одного, а час обробки пакетів у вузлах може змінюватися в широких межах, внаслідок чого такі параметри передачі як затримка і варіація затримки пакетів (джиттер) також можуть змінюватися. А якісні показники мережних послуг, що забезпечують передачу інформації в реальному часі, як відомо, дуже залежать від величини часових затримок пакетів, в яких ця інформація переноситься [4, 5].

У зв'язку з цим метою цієї магістерської кваліфікаційної роботи є дослідження і аналіз основних понять, визначень, стандартизованих механізмів та моделей, які гарантують необхідний рівень QoS. Впровадження таких механізмів та моделей QoS у функціональність мультисервісних IP-мереж дозволяє отримати корисний ефект від конвергентних процесів у таких мережах в процесі надання сучасних послуг та постійно удосконалювати перспективну архітектуру мультисервісної NGN. Тобто вирішення проблем і питань питання забезпечення якості обслуговування є однією з найактуальніших задач.

1 ОСНОВНІ ПОНЯТТЯ І СТАНДАРТИ, ЩО ОПИСУЮТЬ КОНЦЕПЦІЮ ЗАБЕЗПЕЧЕННЯ QoS В МУЛЬТИСЕРВІСНИХ IP-МЕРЕЖАХ

1.1 Поняття якості обслуговування її терміни та зв’язок з характеристиками роботи мережі

Функціонально QoS являє собою набір вимог, що пред'являються до ресурсів мережі у разі здійснення транспортування потоків трафіку даних і забезпечує наскрізну гарантію передачі інформаційних потоків на основі системи правил контролю за засобами підвищення продуктивності мережі [6].

Як було зазначено у вступі, специфіка мереж з комутацією пакетів (КП), до яких відносяться і IP-мережі полягає в тому, що в одному інформаційному потоці може передаватися різнорідний трафік. При цьому кожен із типів трафіку характеризується рядом своїх, тільки йому притаманних, критичних та некритичних параметрів. Тобто показники якості обслуговування специфікують характеристики та властивості конкретних сервісів, проте вимоги для різних сервісів можуть відрізнятися. Наприклад, для послуги «телемедицина» точність доставки інформації більш важлива, ніж варіація затримки передачі, у той час як для послуг IP-телефонії значення та варіація затримки є ключовими параметрами і їх потрібно мінімізувати. Тому, щоб оцінити якість надання послуг у IP-мережі, у разі передачі голосового та відео трафіка вводиться поняття класів обслуговування. Але на цей час визначення якості обслуговування є суб'єктивним механізмом і відповідно до рекомендації ITU-T E.800, являє собою «…сумарний ефект показників якості послуги, який визначає ступінь задоволеності користувача послуги». Тобто неможливо абсолютно гарантувати, що на етапі проектування чи створення мережі будуть закладені мережні характеристики, які б стовідсотково дозволили забезпечити необхідну якість. З іншого боку, треба мати на увазі, що IP-мережі мають розвинені механізми забезпечення

QoS, використання яких дозволяє впливати на якість надання послуг зв'язку в процесі їх функціювання [2, 7].

Потрібно звернути увагу на відмінності термінів, що характеризують якість обслуговування і рівень обслуговування (Grade of Service, GoS). У рекомендації ITU-T E.600 під GoS розуміються технічні параметри (такі як імовірність втрат, час очікування сигналу відповіді станції і ін.), які у разі певних умов визначають відповідність деякої групи ресурсів навантаженню, що надходить [8].

Параметри рівня обслуговування використовуються при плануванні і/або проектуванні мережі та її елементів або надання оцінки щодо їх функціонування у процесі роботи. Тобто це технічні параметри, що формують погляд на якісні показники з боку оператора мережі. Якість обслуговування на відміну від GoS формує погляд на якісні показники з боку кінцевого користувача, тобто, як описує рекомендація E.800, визначає його ступінь задоволеності якістю послуги, що була надана мережею [8].

Якість обслуговування залежить від характеристик роботи мережі (Network Performance, NP), які характеризують ефективність обслуговування трафіку, тобто визначають продуктивність мережі. Насамперед продуктивність мережі характеризує ефективність обслуговування трафіку (пропускна здатність). Вона є найважливішою складовою NP і визначає здатність мережного вузла обслуговувати трафік із заданою інтенсивністю при заданій якості обслуговування та певному технічному стані (співвідношенні кількості працездатних і непрацездатних каналів/ліній). Здатність вузла обробляти трафік залежить від його надійності, якості передачі, а також від наявних ресурсів та можливостей [9].

Зокрема під якістю передачі розуміють рівень відтворення сигналу у вузлі мережі, який знаходиться у стані готовності його прийняти, а під ресурсами мережі – засоби комутації, маршрутизації, переприйому, зберігання інформації, адміністрування, тощо [9].

Відповідно до рекомендації ITU-T E.800 під надійністю розуміється набір характеристик, які описують такі властивості [9]:

- готовність – здатність вузла мережі обробити трафік у будь який час (крім запланованого періоду, протягом якого застосування вузла за призначенням не передбачається) і працювати безвідмовно протягом заданого інтервалу часу;

- безвідмовність – властивість вузла мережі безперервно зберігати працездатний стан протягом певного періоду часу;

- ремонтопридатність – властивість вузла мережі попереджувати і виявляти причини відмов та відновлюватися до працездатного стану шляхом проведення технічного обслуговування (ТО) чи/або ремонтних робіт;

- ТО і ремонт – можливість оператора надати засоби для ТО вузлів мережі (за певних умов експлуатації та визначеній процедурі ТО).

Кожна з цих властивостей може бути описана набором характеристик (показників, атрибутів). Наприклад, готовність до обслуговування визначається показниками, що характеризують працездатність обладнання, середовище розповсюдження, пропускну здатність станцій та вузлів мережі, тощо.

Зв'язок між QoS та NP детально описується у рекомендації ITU-T I.350. Зокрема показано, що якість обслуговування з точки зору користувача може бути виражена сукупністю характеристик, які описуються в термінах, що є зрозумілими як користувачеві, так і мережі, і які не залежить від її структурної та архітектурної організації. Ці характеристики ґрунтуються насамперед на оцінці сприйняття QoS користувачем і повинні бути гарантовані користувачеві оператором мережі та піддаватися об'єктивному виміру в точці доступу до її послуг. Характеристики роботи мережі визначають QoS, що сприймається користувачем, але далеко не завжди дозволяють змістовно, з точки зору того ж користувача, описати цю якість. Прикладами таких характеристик мережі можуть бути: трафік, втрати за викликами, за часом на ділянці мережі, коефіцієнт ефективних викликів за напрямом зв'язку та ін. Параметри QoS, що є корисними на етапі

проектування та побудови мережі, не завжди можуть застосовуватись для специфікації характеристик мережних з'єднань [9].

Крім того, згідно із рекомендацією ITU-T I.350 визначається три функції, що реалізуються мережею та її службами, кожна з яких може бути описана трьома параметрами. Таким чином формується так звана «матриця 3x3», що містить дев'ять родових первинних характеристик, які можуть бути використані для визначення специфічних параметрів QoS та NP (рисунок 1.1) [9]:

- швидкість отримання доступу;
- безпомилковість доступу;
- надійність доступу (імовірність відмови у доступі до ресурсу);
- швидкість перенесення інформації;
- безпомилковість перенесення інформації;
- надійність перенесення інформації;
- швидкість звільнення;
- безпомилковість звільнення;
- надійність звільнення.

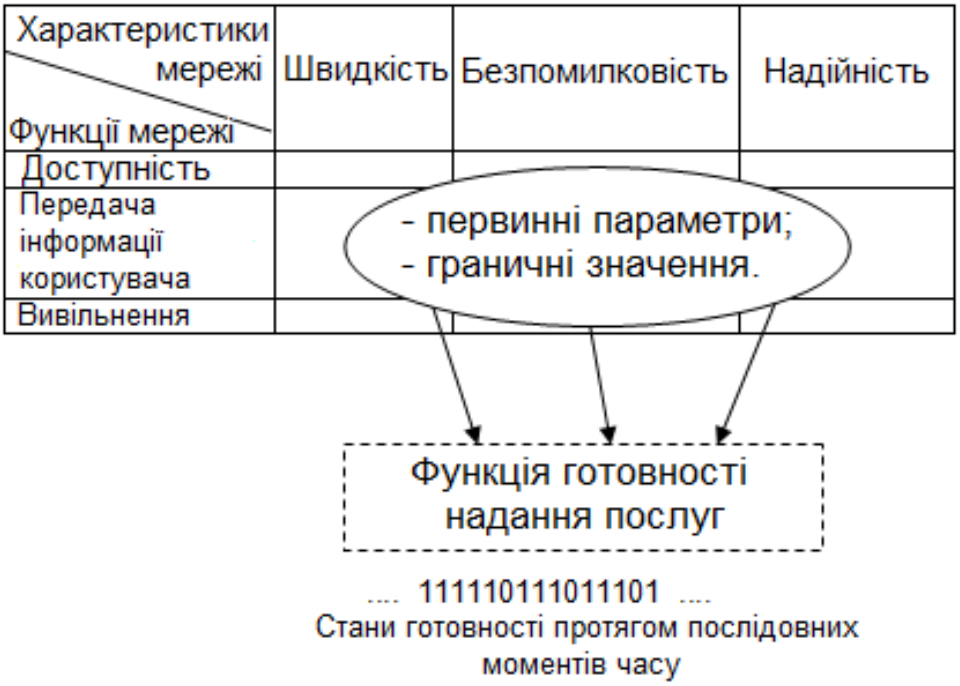

Рисунок 1.1 – Матричний метод 3x3 для визначення станів готовності служби

Служба мережі реалізує три функції зв'язку (рисунок 1.1) [9]:

- забезпечує доступ користувача до ресурсів служби;

- забезпечує перенесення (доставку) інформації по встановленому з'єднанню;

- забезпечує звільнення наданих раніше ресурсів після закінчення сеансу зв'язку.

Під доступністю розуміють можливість отримання ресурсів служби. Процедура доступу починається в момент появи запиту від користувача в інтерфейсі «користувач-мережа» і закінчується з появою хоча б одного біта інформації з його терміналу [9].

Процедура перенесення інформації користувача починається в момент завершення доступу та закінчується в момент передачі запиту звільнення, що позначає закінчення сеансу зв'язку.

Процедура звільнення починається в момент передачі сигналу запиту звільнення та завершується для кожного користувача після звільнення ресурсів служби, що виділялися під час сеансу зв'язку. Визволення включає як дії, що пов'язані з роз'єднанням існуючих з’єднань, так і з припиненням виконання протоколів верхніх рівнів. Якість послуги у разі реалізації функцій служби описується трьома параметрами [9]:

- швидкість – параметр, що характеризує проміжок часу, необхідний для виконання функції (швидкість виконання);

- безпомилковість – параметр, що характеризує ступінь правильності виконання функції (точність виконання);

- надійність – параметр, що визначає ступінь впевненості у виконанні функції протягом заданого періоду спостереження (незалежно від швидкості та безпомилковості виконання).

Для кожного наведеного параметра якості послуги має бути встановлений норматив, з яким можна було б порівнювати виміряні значення у процесі надання послуги.

1.2 Загальний аналіз підходів до оцінки впливу мережних характеристик на QoS в мультисервісних IP-мережах на основі стандартів ITU-T

Для розширення концепції QoS, що була описана у рекомендації ITU-T E.800, була розроблена рекомендація ITU-T G.1000, де зроблене розділення робочих характеристик обслуговування на функціональні компоненти та описаний їх зв'язок з мережними характеристиками, що визначені та описані в низці інших рекомендацій ITU-T, таких як I.350, Y.1540 та Y.1541 [10].

У доповнення до рекомендації ITU-T G.1000, що визначає структуру зв'язків між робочими характеристиками (продуктивністю, надійністю, втратами, затримкою та ін) та характеристиками мережі, рекомендація ITU-T G.1010 містить специфікації вимог з боку додатків, що орієнтовані на кінцевого користувача [10].

На підставі цих рекомендацій ITU-T, що присвячені стандартизації QoS в IP-мережах, передбачаються наступні етапи вирішення питань забезпечення якості обслуговування у таких мережах [10]:

- створення узгодженого загального набору робочих характеристик IP-мереж та опрацювання відповідних норм для нього;

– впровадження мережних механізмів, які забезпечуватимуть задані показники QoS;

- введення нормованих значень показників QoS в протоколи сигналізації;

- розробка архітектури підтримки параметрів якості обслуговування з використанням мережних механізмів.

Вирішенню завдань, що відповідають першому з наведених етапів відповідають рекомендації ITU-T Y.1540 та Y.1541. У першій (Y.1540) описуються стандартні мережні характеристики передачі пакетів в IP-мережах, а у другій (Y.1541) – визначені норми для параметрів, що описані в Y.1540, між двома граничними мережними інтерфейсами (точками підключення кінцевих терміналів). Також у рекомендації ITU-T Y.1541

наведена специфікація шести класів QoS в залежності від додатків, що використовуються [10].

Потрібно зазначити, що ці рекомендації є важливими як для операторів мереж і виробників мережного обладнання, так і для кінцевих користувачів. Мережні оператори будуть використовувати їх у разі планування, розгортання та визначення оцінки IP-мереж відповідно з вимогами кінцевих користувачів до якості обслуговування. Виробники обладнання будуть орієнтуватися на ці рекомендації під час створення обладнання, яке має відповідати специфікаціям мережних операторів. Кінцеві користувачі (перш за все корпоративні) зможуть застосувати рекомендації Y.1540 та Y.1541 у разі оцінки характеристик IP-мереж, що реально функціонують, з позицій відповідності цих характеристик вимогам споживачів їх послуг.

Розглянемо докладніше ці рекомендації щодо до основних мережних характеристик, які пов'язані із забезпеченням QoS у мультисервісних IP-мережах.

У рекомендації ITU-T Y.1540 розглядаються такі найважливіші за рівнем впливу на забезпечення QoS мережні характеристики: продуктивність мережі, надійність мережі (мережних елементів), затримка, втрати пакетів, варіація затримок (джиттер) [10].

1) Під продуктивністю мережі розуміється ефективна швидкість передачі даних користувача, яка вимірюється в бітах на секунду. Необхідно зазначити, що значення цієї характеристики не є тотожною максимальній пропускній здатності мережі, яку досить часто помилково називають смугою пропускання. Мінімальне значення продуктивності мережі, як правило, гарантує провайдер послуг, який у свою чергу повинен мати відповідні гарантії від мережного провайдера [10].

Слід зазначити, що у рекомендації Y.1540 не наведено нормативних характеристик продуктивності мережі для різних додатків. Але параметри, що пов'язані з ефективною швидкістю передачі, можуть визначатися через дескриптор трафіку IP-мережі, який описаний в рекомендації ITU-T Y.1221 – це, зокрема, стверджується у рекомендації Y.1541 [10].

2) Користувачі завжди сподіваються отримати високий рівень надійності від систем зв'язку. Як було вище зазначено, у відповідності з рекомендацією ITU-T E.800 надійність можна визначити через декілька параметрів, але найчастіше використовується коефіцієнт готовності ($К_Г$). У загальному випадку, для мереж зв’язку і мережного обладнання, коефіцієнт готовності – це імовірність того, що мережа або обладнання у будь-який (довільний) момент часу перебуватиме в робочому стані. Цей параметр обчислюється як [10, 11]:

$$К_Г = \tau_{нв} / (\tau_{но} + \tau_в) \quad (1.1)$$

де $\tau_{нв}$ – середній час напрацювання на відмову (середнє напрацювання між відмовами);

$\tau_в$ – середній час відновлення працездатності (середній час до відновлення).

В найкращому випадку $К_Г$ має дорівнювати 1, що говорить про сто відсоткову готовності мережі. Насправді він оцінюється деякою кількістю «дев'яток». Наприклад, «три дев'ятки» говорять про те, що $К_Г$ становить 0,999 – це відповідає 9 годинам часу простою (недоступності) мережі на рік. Наприклад, готовність телефонної мережі загального користування (ТМЗК) оцінюється величиною «п'ять дев'яток» (0,99999), що відповідає 5,5 хвилинам простою на рік [10].

Потрібно звернути увагу на те, що забезпечення коефіцієнта готовності 0,99999 в мультисервісних IP-мережах, що побудовані на традиційному обладнанні (маршрутизатори, сервери), буде становити серйозну проблему. Це буде тому, що обробка інформаційних потоків в IP-мережах у значній частині базується на програмному забезпеченні (ПЗ) (а не на апаратному, як це має місце у ТМЗК). До того ж, статистика відмов мережного обладнання показує, що надійність ПЗ приблизно вдвічі нижча за надійність апаратного забезпечення [10].

3) У загальному випадку сеанс зв'язку складається з трьох фаз: встановлення з'єднання, передачі та роз'єднання з'єднання. У рекомендації

ITU-T Y.1540 із цих трьох фаз розглядається лише друга – фаза доставки пакетів IP. Такий підхід випливає із технологічних властивостей IP-мереж – вони не орієнтовані на встановлення з'єднань. Y.1540 визначає такі основні часові параметри, що характеризують доставку IP-пакетів [2, 3, 10]:

- затримка доставки IP пакета (IP packet Transfer Delay, IPTD);
- середня затримка доставки IP пакета.

Параметр IPTD визначається як час (t_2 - t_1) між двома подіями – введенням пакета у вхідну точку мережі в момент часу t_1 та виведенням пакета з вихідної точки мережі в момент часу t_2,, де ($t_2 > t_1$) та (t_2 - t_1) $\leq T_{max}$. Тобто, у загалі параметр IPTD визначає час доставки пакета між джерелом і одержувачем для всіх пакетів – як успішно переданих, так і уражених помилками [10].

Середня затримка доставки пакета IP визначається як середня арифметична величина затримок пакетів у вибраному наборі переданих і прийнятих пакетів. Значення середньої затримки залежить від трафіку, що передається в мережі, і доступних мережних ресурсів, зокрема, від пропускної здатності. Зростання навантаження та зменшення доступних мережних ресурсів ведуть до зростання черг у вузлах мережі та, як наслідок, до збільшення середніх затримок доставки пакетів [10].

Прикладами чутливого до затримок трафіку є передача мовної та відеоінформації, тоді як передача звичайних даних (текстова інформація) є менш чутливою до них. У разі якщо затримка доставки пакета перевищує певні значення часу T_{max}, то такі пакети будуть відкинуті. У додатках, що працюють у реальному часі (наприклад, IP-телефонії) – це призведе до погіршення якості мови. Обмеження, що пов'язані із середньою затримкою пакетів IP, мають суттєве значення для успішного впровадження технологій VoIP, відеоконференцій та інших додатків, щ працюють у реальному часі [10].

4) Варіація затримки IP пакета (IP packet Delay Variation, IPDV) або джиттер характеризується параметром V_k. Для IP-пакета з індексом k цей параметр визначається між вхідною та вихідною точками мережі у вигляді

різниці між абсолютною величиною затримки X_k у разі доставки пакета з індексом k, та певною еталонною (або опорною) величиною затримки доставки IP пакета ($d_{l,2}$) для тих же мережних точок: $V_k = X_k - d_{l,2}$ [10].

Еталонна затримка доставки пакета IP ($d_{l,2}$) між джерелом та одержувачем визначається як абсолютне значення затримки доставки першого IP пакета між цими мережними точками. Таким чином, варіація затримки IP пакета або джиттер – це різниця у часі проходження по мережі послідовних пакетів одного з'єднання, тобто він проявляється в тому, що послідовні пакети прибувають до одержувача у нерегулярні моменти часу. Чим більше джиттер, тим сильніше буде відрізнятися затримка при передачі одного пакета від затримки при проходженні іншого. У системах IP-телефонії це, наприклад, веде до спотворень звуку (тріскіт чи клацання) і в результаті – мова стає нерозбірливою [10].

Вплив джиттеру зменшують шляхом включення в приймальну частину буфера шлюзу статичної або динамічної пам'яті, за рахунок чого відновлюється вихідна послідовність пакетів. Пакети, джиттер яких перевищує час їх «утримання» у буферній пам'яті, не сприймаються приймальним пристроєм і відкидаються. Таким чином, буфер зменшує вплив джиттеру за рахунок збільшення як загального часу утримання, так і втрати пакетів. Тут регулювання часу утримання (розміру буфера) являє собою компроміс між ними.

5) Коефіцієнт втрати IP пакетів (IP packet Loss Ratio, IPLR) визначається як відношення сумарного числа втрачених пакетів до загального числа прийнятих у вибраному наборі переданих та прийнятих пакетів. Втрати пакетів у IP мережах виникають у тому разі, коли значення затримок у разі їх передачі перевищує нормоване значення, що визначене вище як T_{max}. Якщо пакети губляться, то при передачі даних можливо здійснити їх повторну передачу за запитом приймаючої сторони. У системах VoIP пакети, що прийшли до одержувача із затримкою, яка перевищує час T_{max}, будуть відкинуті, що веде до погіршення розбірливості мови. Серед

причин, які викликають втрати пакетів, необхідно звернути увагу на зростання черг у вузлах мережі, що виникають під час перевантаження [10].

Рекомендація ITU-T Y.1541 визначає чисельні значення параметрів, які специфікованих у рекомендац ії Y.1540 та яких мають додержуватися в IP мережах на міжнародних трактах, що з'єднують термінали користувачів. Існуючі норми на ці параметри поділені за різними класами якості обслуговування, які визначені залежно від додатків та мережних механізмів, що застосовуються для забезпечення гарантованої QoS. У таблиці 1.1 наведені норми на визначені вище мережні характеристики [10].

Таблиця 1.1 - Норми характеристик IP мереж з розподілом їх за класами QoS

Мережні характеристики	Класи якості обслуговування					
	0	1	2	3	4	5
Варіація затримки IP пакета, IPDV	50 мс	50 мс	Н	Н	Н	Н
Коефіцієнт помилок IP пакетів, IPER	$1x10^{-4}$	$1x10^{-4}$	$1x10^{-4}$	$1x10^{-4}$	$1x10^{-4}$	Н
Затримка доставки IP пакета, IPTD	100 мс	400 мс	100 мс	400 мс	1 с	Н
Коефіцієнт втрати IP пакетів, IPLR	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	Н

Примітка: Н – не нормовано

Значення характеристик, що наведені в таблиці 1.1, являють собою, відповідно, верхні межі для величин середніх затримок, джиттера, втрат і помилок пакетів. У рекомендації Y.1541 надані специфікації набору характеристик, які зв'язані з вимірюванням реальних значень мережних характеристик: періоду спостережень, довжини тестових пакетів, числа пакетів тощо. Зокрема, при здійсненні оцінки передачі пакетів мови в IP-телефонії мінімальний період спостереження має бути в межах 1 - 20 с у разі

середньої швидкості передачі 50 пак./с. Рекомендований інтервал вимірювань для затримки, джиттера і втрат має становити не менше 60 с [10].

Також у рекомендації Y.1541 встановлюється відповідність між класами QoS та IP-орієнтованими додатками [10]:

- клас 0 – додатки реального часу, що є чутливими до джиттеру, для яких характерний високий рівень взаємодії з користувачами (відеоконференції, VoIP);
- клас 1 – додатки реального часу, що також чутливі до джиттеру, і також мають можливості щодо забезпечення взаємодії з користувачами;
- клас 2 – транзакції даних, для яких характерний високий рівень забезпечення взаємодії з користувачами (наприклад, сигналізація);
- клас 3 - транзакції даних, що також мають можливості щодо забезпечення взаємодії з користувачами;
- клас 4 – додатки, для яких допустимий низький рівень втрат (потокове відео, масиви даних, короткі транзакції);
- клас 5 – традиційні застосування IP-мереж.

1.3 Вимоги до QoS додатків різних типів у мультисервісних IP-мережах

Впровадження будь яких механізмів QoS передбачає забезпечення з боку мережі з'єднання з певними обмеженнями продуктивності, основними параметрами якої, як було визначено вище є смуга пропускання, затримка, джиттер і втрати пакетів (рекомендація ITU-T Y.1540). Особливо параметри QoS важливі тоді, коли по мережі передається одночасно трафік різного типу, наприклад, трафік веб-додатків та мовний, оскільки, як це також зазначалось, різні типи трафіку закономірно висувають різні вимоги до параметрів QoS. Одночасно врахувати всі параметри QoS для всіх видів трафіку дуже складно, тому всі види трафіку, що існують у мережі, класифікують, відносячи кожен вид до одного з найпоширеніших типів, а потім намагаються досягти одночасного виконання певного набору вимог для цих типів трафіку [5].

В якості основних критеріїв класифікації за основу прийняті три параметри трафіку (рисунок 1.2) [5]:

- відносна передбачуваність швидкості передачі даних;

- чутливість трафіку до затримок пакетів;

- чутливість трафіку до втрат та спотворень пакетів.

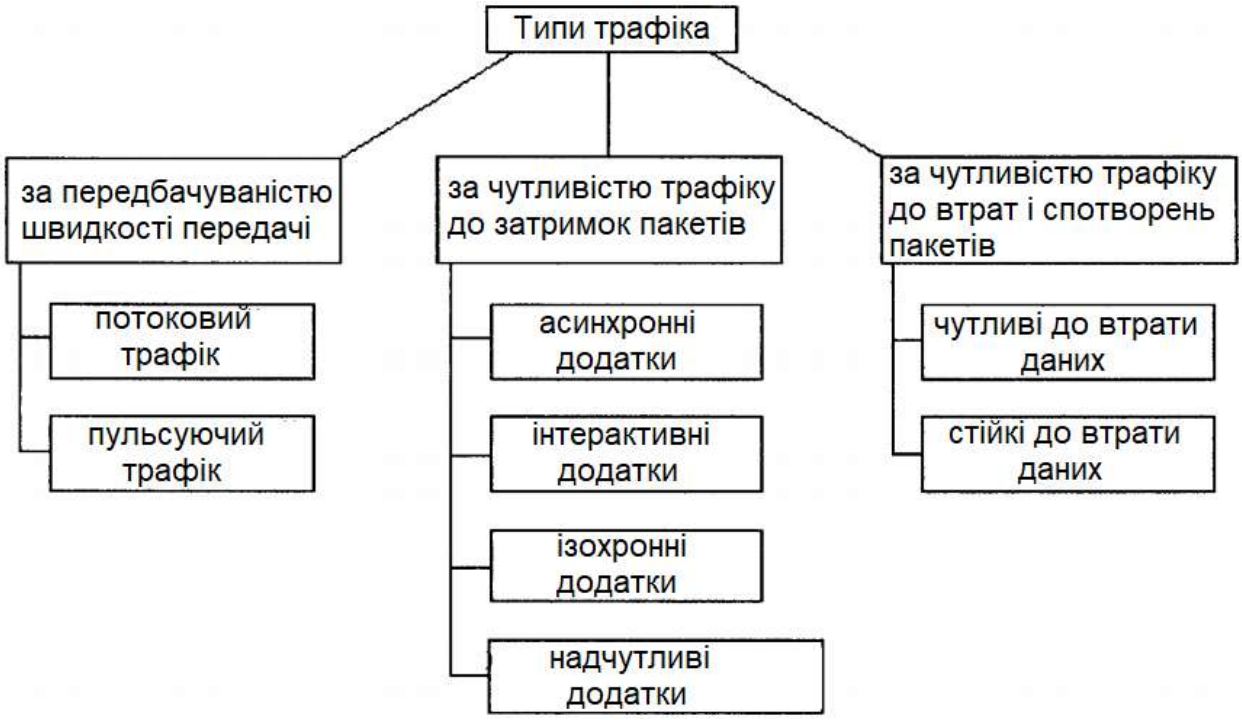

Рисунок 1.2 – Класифікація мережного трафіка в мультисервісній IP-мережі

1) Трафік з відносною передбачуваністю швидкості передачі даних включає наступні типи трафіку [5]:

- додатки з потоковим трафіком породжують рівномірний потік даних, що надходить у мережу з постійною бітовою швидкістю (Constant Bit Rate, CBR). У пакетній мережі трафік таких додатків являє собою послідовність пакетів однакового розміру (що дорівнює B біт), які просуваються один за одним через один і той же інтервал часу T. Такого типу трафік може бути обчислений шляхом усереднення за одним періодом: $CBR = B/T$ біт/с;

- додатки з пульсуючим трафіком відрізняються високим ступенем непередбачуваності, коли періоди мовчання змінюються пульсацією,

протягом якої пакети «щільно» просуваються один за одним. В результаті трафік має змінну бітову швидкість (Variable Bit Rate, VBR). Практично будь-який трафік, навіть трафік потокових додатків, має ненульовий коефіцієнт пульсації (для пульсуючого трафіку – від 2:1 до 100:1, для потокового – приблизно 1:1).

2) До трафіку, що є чутливим до затримок пакетів, відносяться такі додатки як [5]:

- асинхронні додатки, до яких належать ті, що практично не мають обмежень на час затримки (еластичний трафік), наприклад, електронна пошта;

- інтерактивні додатки – це ті, на функціональності яких затримки не позначаються негативно, наприклад, текстовий редактор, у якому редагується файл, що був завантажений віддалено;

- ізохронні додатки, до яких належать ті, що мають поріг чутливості до варіацій затримок, перевищення якого різко знижує функціональність додатку, наприклад, передача голосу;

- функціональність надчутливих до затримок додатків затримка зводить нанівець, наприклад, додатки, що здійснюють управління технічним об'єктом у реальному часі.

3) До трафіку, що є чутливим до втрат та спотворень пакетів, відносяться такі додатки як [5]:

- додатки, що є чутливими до втрати даних, – це програми, що передають алфавітно-цифрові дані (текстові документи, коди програм, числові масиви, тощо). Усі традиційні мережні додатки (файловий сервіс, сервіс баз даних, електронна пошта, тощо) відносяться до цього типу додатків;

- додатки, що є стійкими до втрати даних, до яких належать ті, що передають трафік з інформацією про інерційні фізичні процеси. Їх стійкість до втрат пояснюється тим, що невелику кількість відсутніх даних можна визначити на основі вже прийнятих. До цього типу належить більшість додатків, що працюють з мультимедійним трафіком (аудіо- і відео-). Проте

відсоток втрачених пакетів у таких додатках не повинен бути великим (не більше ніж 1%).

Таким чином, з вищевикладеного видно, що організація мультисервісних IP мереж супроводжується серйозними проблемами у сфері забезпечення необхідної QoS. У значній мірі це пов'язано з тим, що протокол IP, незважаючи на свою універсальність, спочатку не призначався для обміну інформацією в реальному часі. Транспортні протоколи, такі як TCP та UDP, що реалізуються в обладнанні користувачів і функціонують поверх протоколу IP, також не забезпечують у повній мірі високої QoS трафіку, який є чутливим до затримок.

Разом з тим обґрунтовано необхідність отримання від мережі необхідних гарантій якісної доставки чутливої до затримок інформації (такої як мова, відео та інші мультимедійні додатки подібного типу) у реальному часі з мінімально можливою затримкою. З цією метою для пакетних мереж, у тому числі і для IP, ITU-T розроблено безліч стандартів, які регламентують забезпечення в таких мережах необхідної QoS.

Зокрема ці стандарти описані у відповідних рекомендаціях, серед яких у цьому розділі значна увага приділена тим з них, що описують:

- термінологію та поняття QoS та якісні показники функціонування мережі (рекомендації ITU-T E.800 та I.350);

- базові мережні характеристики передачі пакетів у IP-мережах (рекомендація ITU-T Y.1540);

- нормовані значення цих базових мережних характеристик відповідно до різних класів QoS (рекомендація ITU-T Y.1541).

У наступному розділі проаналізуємо базову модель підтримки QoS у пакетних мережах, яка описує мережні механізми забезпечення QoS в мультисервісних IP-мережах відповідно до рекомендації ITU-T Y.1291.

2 БАЗОВА МОДЕЛЬ, МЕХАНІЗМИ, ФУНКЦІЇ І ТЕХНОЛОГІЇ ПІДТРИМКИ QOS В МУЛЬТИСЕРВІСНИХ IP-МЕРЕЖАХ

2.1 Базова модель підтримки якості обслуговування в мультисервісних IP-мережах

Крім визначення мережних характеристик та специфікацій норм на якість обслуговування для них, що були розглянуті при аналізі рекомендацій Y.1540 та Y.1541 у попередньому розділі, ITU-T також зробив ідентифікацію та стандартизацію мережних механізмів, що забезпечують QoS в мультисервісних IP-мережах. Це призвело до появи рекомендації ITU-T Y.1291, у якій описується базова модель мережних механізмів, що забезпечують QoS у пакетних мережах [10].

Такі мережні механізми мають застосовуватися у комбінації з параметрами QoS, які формуються в залежності від додатків, що підтримуються мережею. Під час здійснення розробки цієї моделі ITU-T було враховувано, що різні послуги матимуть різні вимоги до мережних характеристик. Так у попередньому розділі вже наводився приклад послуги телемедицини, для якої більш важливе значення має точність доставки, ніж джиттер або сумарна середня затримка, але для послуг IP-телефонії параметри джиттеру і затримки є ключовими характеристиками, тому що їх потрібно якомога сильніше мінімізувати. Тому, з урахуванням тенденції постійного розширення числа додатків, що мають різні вимоги до характеристик QoS, модель підтримки якості обслуговування, має включати широкий набір загальних мережних механізмів, як тих, що існують, так і перспективних, що знаходяться у розробці [10].

Запропонована ITU-T у рекомендації Y.1291 базова модель підтримки якості обслуговування визначає набір мережних механізмів, які називаються конструктивними блоками. Зокрема в Y.1291 визначені конструктивні блоки, які відповідають трьом логічним площинам (рисунок 2.1) [10]:

- площині контролю;

- площині даних (інформаційній площини);

- площині адміністративного управління.

Далі розглянемо докладніше особливості механізмів якості обслуговування, що входять до складу цих площин.

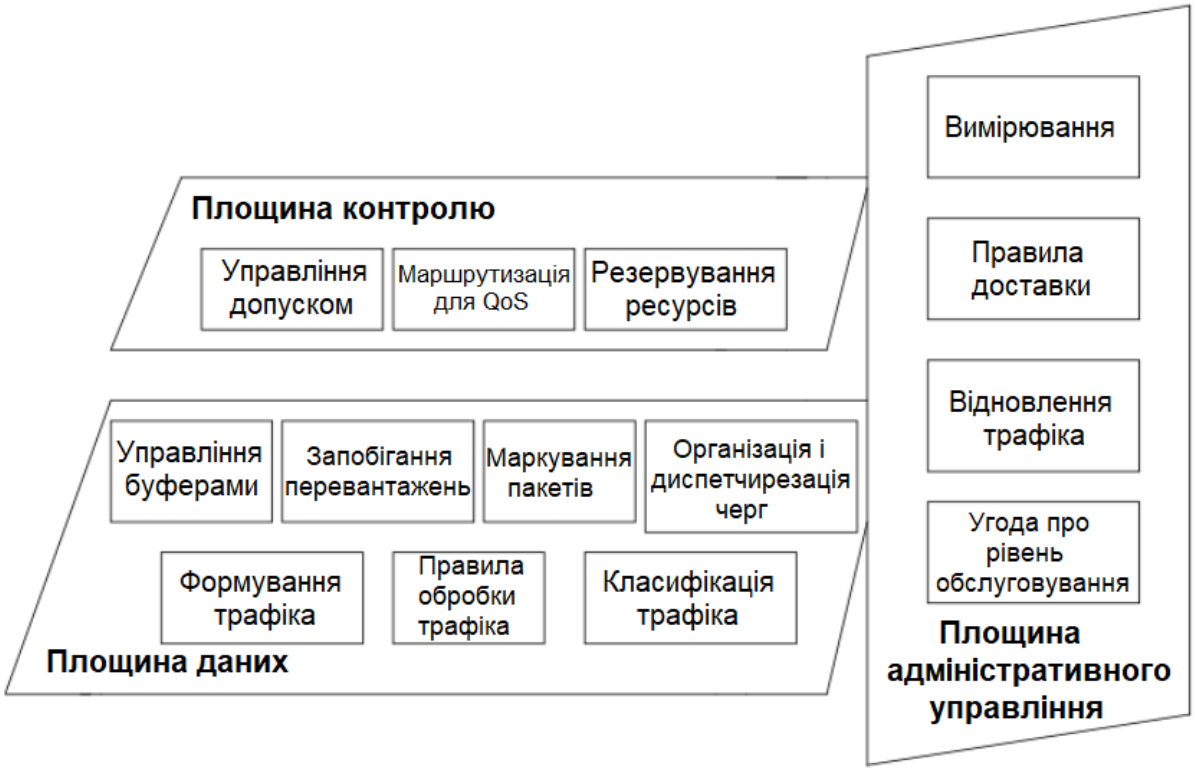

Рисунок 2.1 – Базова модель підтримки якості обслуговування в мультисервісних IP-мережах

2.1.1 Механізми забезпечення QoS, що реалізуються на площині контроля

На першій площині контролю механізми якості обслуговування взаємодіють із шляхами, якими передається трафік користувачів, і включають до свого складу [10]:

- управління допуском;

- маршрутизацію QoS;

- резервування ресурсів.

1) Механізм управління допуском здійснює контроль за новими заявками на передачу трафіку через мережу. Він визначає, чи може новий трафік привести до перевантаження мережі або до погіршення рівня QoS для вже наявного в мережі трафіку. Зазвичай управління допуском здійснюється за певним набором правил адміністрування, контролю та управління мережними ресурсами. Ці правила можуть бути специфіковані відповідно до потреб мережі провайдера або базуватися на угоді, що укладена між провайдером і користувачем, і включати до свого складу різні характеристики якості обслуговування. Для задоволення вимог певних служб (наприклад, у разі надзвичайних ситуацій) відповідному трафіку може бути наданий найвищий пріоритет у разі здійснення доступу до мережі [10].

2) Маршрутизація QoS забезпечує вибір шляху, що буде відповідати вимогам до QoS для конкретного потоку даних. Шлях, що обирається, може не співпадати з найкоротшим шляхом. Процес визначення шляху передбачає знання вимог до QoS з боку потоку даних та наявність інформації про доступні мережні ресурси. На цей час запропоновано велику кількість можливих методів і алгоритмів визначення найоптимальнішого шляху за критерієм забезпечення якості обслуговування. Як правило, у процесі визначення найоптимальнішого шляху в маршрутизації QoS, для того, щоб зробити процес обчислень прийнятним для інженерних розрахунків, враховується або одна мережна характеристика, або дві (продуктивність і затримка, вартість і продуктивність, вартість і затримка, тощо) [10].

3) Для забезпечення функціонування механізму резервування ресурсів загальною необхідною умовою є наявність ресурсів у мережі. Цей механізм застосовувся у пакетних мережах з асинхронним режимом передачі інформації (Asynchronous Transfer Mode, ATM) для формування постійних віртуальних з'єднань. У IP-мережах найбільш розповсюдженим механізмом резервування є механізм, що забезпечується однойменним протоколом резервування ресурсів (Resource Reservation Protocol, RSVP) [10].

2.1.2 Механізми забезпечення QoS, що реалізуються на площині даних

На площині даних механізми QoS направлені на роботу безпосередньо з трафіком користувачів і включає наступні їх типи [10]:

- управління буферами;
- запобігання перевантажень;
- маркування пакетів;
- організацію та диспетчеризацію черг;
- формування трафіку;
- правила обробки трафіку;
- класифікацію трафіку.

1) Механізм управління буферами (або чергами) полягає в управлінні пакетами, що стоять у черзі у вузлах, очікуючи на передачу. Основними задачами щодо управління чергами є мінімізація середньої довжини черги при одночасному забезпеченні високого використання каналу, а також забезпечення справедливого розподіл буферного простору між різними потоками даних. Схеми здійснення управління чергами розрізняються переважно за критерієм відкидання пакетів і місцем у черзі, звідки здійснюється скидання пакетів (початок або кінець черги). Найбільш простим критерієм для скидання пакетів є досягнення чергою максимальної своєї довжини, яка регламентується тим або іншим алгоритмом обробки черги, що підтримується буфером [10].

Найпоширенішими сьогодні є механізми активного управління чергами, прикладом яких є алгоритм довільного раннього виявлення перевантаження (Random Early Detection, RED). У разі використанні алгоритму RED пакети, що надходять у буфер, відкидаються на підставі здійснення оцінки середньої довжини черги. Імовірність відкидання пакетів зростає із зростанням середньої довжини черги. Модифікацією цього алгоритму є так званий зважений алгоритм раннього довільного виявлення (Weighted Random Early Detection, WRED), що дозволяє налаштовувати різні

характеристики RED в залежності від значення поля IP-пріоритету або класу трафіку. Алгоритм WRED на основі потоку передбачає можливість призначення штрафу з ненульовою імовірністю для тих потоків, які намагаються заволодіти надто великою часткою доступних ресурсів. Подібним до алгоритмів RED і WRED є алгоритм явного повідомлення про перевантаження (Explicit Congestion Notification, ECN), який дозволяє попередити джерело на основі протоколу TCP про те, що у мережі починається перевантаження, шляхом маркування (а не відкидання) пакетів [5, 10, 12].

2) Механізми запобігання перевантажень підтримують рівень навантаження в мережі нижче за її пропускну здатність. Звичайний спосіб запобігання перевантажень полягає у зменшенні трафіку, що надходить у мережу. Як правило, команда зменшити трафік впливає насамперед на низькопріоритетні джерела. Одним із прикладів механізмів запобігання перевантаженням є механізм вікна, що реалізується у протоколі TCP [10].

Також потрібно зазначити, що алгоритми управління чергами теж відносяться до механізмів боротьби з перевантаженнями в мультисервісних IP-мережах. Найбільш простим і поширеним механізмом обслуговування черг є алгоритм «першим прийшов, першим вийшов» (First-In, First-Out, FIFO). Він достатньо ефективний, але не передбачає пріоритетної обробки чутливого до затримок трафіку шляхом його переміщення на початок черги. Також цей алгоритм не проводить ніяких дій щодо запобігання перевантажень або зменшення розміру черги для зниження часу затримки. Згадані вище алгоритми RED, WRED та ECN належать до значно ефективніших механізмів боротьби з перевантаженнями [5].

Звідси також можна бачити, що механізм управління буферами можна вважати окремим випадком механізмів запобігання перевантаженням.

3) Механізм маркування пакетів полягає в тому, що пакети можуть бути промарковані відповідно до певного класу обслуговування. Маркування зазвичай проводиться у вхідному прикордонному вузлі, де у спеціальне поле заголовка (Type of Service, ToS в заголовку IP або DS-байт в заголовку

DiffServ) вводиться певне значення. Крім того, маркування застосовується для тих пакетів, які можуть бути видалені у випадку перевантаження мережі [10].

4) Метою групи механізмів організації та диспетчеризації черг є вибір пакетів передачі з буфера в канал. Більшість дисциплін обслуговування (або планувальників) ґрунтуються на схемі «перший прийшов – перший обслуговується». Для забезпечення більш гнучких процедур виведення пакетів із черги було запропоновано низку схем, що ґрунтуються на формуванні кількох черг. Серед них перш за все необхідно назвати схеми пріоритетного обслуговування. Інший приклад гнучкої організації черги – це застосування механізму на основі зваженого алгоритму рівномірного обслуговування черг (Weighted Fair Queuing, WFQ), який обмежену пропускну здатність на виході вузла розподіляє між кількома потоками (чергами) залежно від вимог до пропускної здатності кожного потоку [10].

Ще одна схема організації черги ґрунтується на класифікації потоків за класом обслуговування (Class-Based Queuing, CBQ). Потоки класифікуються відповідно до класів обслуговування, а потім розміщуються в буфері в різних чергах. Кожній черзі виділяється певний відсоток вихідної пропускної здатності залежно від класу, і далі черги обслуговуються за циклічною схемою [10].

5) Механізм формування або керування характеристиками трафіку передбачає контроль швидкості передачі пакетів та обсягу потоків, що надходять на вхід мережі. В результаті проходження через спеціальні буфери-формувачі зменшується пачечність вихідного трафіку і його характеристики стають більш передбачуваними. Відомі два механізми обробки трафіку: «відро з дірками» (Leaky Bucket) та «відро з жетонами» (Token Bucket) [10].

Алгоритм Leaky Bucket регулює швидкість пакетів, що залишають вузол. Незалежно від швидкості вхідного потоку, швидкість на виході вузла є постійною величиною. Коли відро переповнюється, зайві пакети відкидаються. На протилежність «відру з дірками» алгоритм Token Bucket не регулює швидкість на виході вузла і не скидає пакети. Швидкість пакетів на виході вузла може бути такою самою, як і на вході, якщо тільки у відповідному накопичувачі

(«відрі») є жетони. Жетони генеруються з певною швидкістю та накопичуються у «відрі». Алгоритм характеризується двома параметрами: швидкістю генерації жетонів та розміром пам'яті (розміром «відра») для них. Пакети не можуть залишити вузол, якщо у відрі немає жетонів. І навпаки, відразу група пакетів може залишити вузол, витративши потрібну для цього кількість жетонів [10].

6) Механізм «Правила або політика обробки трафіку» приймає рішення про відповідність трафіку, що надходить від одного транзитного вузла до іншого транзитного вузла, заздалегідь узгодженим правилам обробки або контрактам. Як правило, ті пакети, що не відповідають цим правилам, відкидаються. Відправники можуть бути сповіщені про відкинуті пакети та виявлені причини, а також про дотримання відповідності в майбутньому, що обумовлене угодами про рівень обслуговування (Service Level Agreement, SLA) [10].

7) Класифікація трафіку може бути проведена на потоковому або пакетному рівні. На вході в мережу у вузлі доступу (прикордонному маршрутизаторі) пакети класифікуються для того, щоб виділити пакети одного потоку, який характеризується загальними вимогами до QoS. Потім трафік піддається процедурі нормування (механізм Traffic Conditioning). Нормування трафіку передбачає вимірювання його параметрів та порівняння результатів з параметрами, що прописані у трафік-контракті (угода SLA). Якщо умови SLA порушуються, частина пакетів може бути відкинута. Магістральні маршрутизатори, що становлять ядро мережі, забезпечують пересилання пакетів відповідно до потрібного рівня якості обслуговування [10].

2.1.3 Механізми забезпечення QoS, що реалізуються на площині адміністративного управління

Остання площина адміністративного управління містить механізми якості обслуговування, що мають відношення до експлуатації, адміністрування та управління мережею відповідно до здійснення доставки трафіка від користувачів.

У склад цих механізмів входять [10]:

- вимірювання;
- задані правила доставки;
- відновлення трафіку;
- угода про рівень обслуговування (SLA).

1) Вимірювання забезпечують контроль параметрів трафіку, наприклад, швидкості потоку даних у порівнянні з узгодженою в SLA швидкістю. За результатами вимірювань можуть бути реалізовані певні процедури – такі, як скидання пакетів та застосування механізмів Leaky Bucket та Token Bucket, що були розглянуті вище [10].

2) Задані правила (політика) доставки. Під правилами доставки тут розуміється набір правил, що використовуються для здійснення контролю та адміністративного управління доступом до мережних ресурсів. На основі таких правил постачальники послуг можуть здійснювати реалізацію механізмів у площині управління та площині даних. Можливими застосуваннями таких заданих правил доставки є: маршрутизація за цими правилами, фільтрація пакетів на основі цих правил (маркування або відкидання пакетів), реєстрація заданих потоків, правила обробки, що пов'язані з безпекою [10].

3) Під механізмом відновлення трафіку в рекомендації ITU-T Y.1291 розуміється реакція мережі, що пом'якшує можливі наслідки у разі виконання умов, за якими здійснюється відмова в обслуговуванні. Відновлення трафіку розглядається на різних рівнях еталонної моделі взаємодії відкритих систем (EMBBC). На фізичному рівні, наприклад, у разі використання технології надійність забезпечується автоматичною захисною комутацією. На канальному рівні транспортних мереж відновлення трафіку забезпечується спеціальними механізмами, розвиненими, у випадку застосування технології SDH, для кільцевих та коміркових структур. Відновлення на мережному рівні (протокол IP) здійснюється за допомогою технології багатопротокольної комутації за мітками (Multi-Protocol Label Switching, MPLS) [10].

4) Одним із основних понять у концепції забезпечення необхідного рівня QoS у сучасних мультисервісних IP-мережах є угода про рівень обслуговування (SLA) або так званий «трафік-контракт», що укладається між користувачем і провайдером послуг / мережним провайдером. Необхідність заключення таких трафік-контрактів викликана зростаючими вимогами до операторів з боку клієнтів, які потребують все більш надійної та своєчасної передачі інформації. У контракті визначаються основні характеристики (профіль) трафіку, що формується в обладнанні користувача, та параметри QoS, що надаються провайдером. Зокрема, у SLA мають бути присутніми такі характеристики: потрібна швидкість протягом сеансу, припустима затримка пакетів у потоці, допустима імовірність втрати пакетів у потоці, правила перевірки відповідності дійсних параметрів трафіку угоді SLA, дані для здійснення маршрутизації пакетів (адреси пунктів призначення). Угода SLA може включати також і цінові характеристики. [9, 10].

Контракт SLA може бути статичним, тобто укладеним на тривалий період (місяць, рік, тощо), або динамічним – визначається для кожного сеансу. В останньому випадку для запиту необхідного рівня QoS має використовуватися сигнальний протокол (наприклад, протокол RSVP). У разі укладання угоди SLA передусім передбачається визначити чітко регламентовані зобов'язання постачальника послуг щодо забезпечення їх якості (наприклад, час надання послуги – цілодобово або лише в робочі дні, час реакції на можливий інцидент, час виїзду технічного майстра до замовника; час закриття інциденту, тощо) і, навіть, штрафні санкції у разі порушення регламенту виконання тих або інших робіт чи надання сервісів [10].

Таким чином можна бачити, що мультисервісні IP-мережі мають розвинені мережні механізми забезпечення QoS (або, згідно із термінологією ITU-T у рекомендації Y.1291, блоки QoS). Ці механізми можуть бути специфіковані відповідно до мережних вузлів (наприклад, управління буферами вузлів) або до мережних сегментів (маршрутизація QoS). Зазначимо, що термін «мережний сегмент» може відноситися до: між кінцевого з'єднання, ділянки доступу, між вузлової ділянки або ділянки, що з'єднує дві та більше IP-мережі [10].

2.2 Функції і технології QoS

Узагальнюючи аналіз мережних механізмів забезпечення QoS в мультисервісній IP-мережі, можна стверджувати, що всі функції якості обслуговування, які вони забезпечують, направлені на здійснення диференційованого і гарантованого обслуговування трафіку, що у свою чергу досягається передачею контролю за завантаженістю мережі та використанням її ресурсів оператору. Тобто, інакше кажучи, якість обслуговування представляється певною сукупністю вимог, які пред'являються до мережних ресурсів у разі передачі потоку даних. При цьому забезпечується наскрізна гарантія передачі даних та контроль за засобами підвищення продуктивності мультисервісної IP-мережі, що засновані на розглянутих вище механізмах з дотриманням визначеної системи правил [2, 6].

Зокрема мережні механізми QoS мають забезпечити реалізацію наступних функцій [13]:

- управління ресурсами мережі – здійснюється управління смугою пропускання, мережними пристроями, використовуються можливості роботи у глобальній мережі та ін.;

- ефективного використання мережевих ресурсів – використання інструментів менеджменту та тарифікації дозволяє регулювати трафік з метою збільшення економічного ефекту;

- специфічних послуг – управління та контроль параметрів QoS дозволяє провайдерам послуг забезпечувати своїм клієнтам різні рівні обслуговування;

- спільного існування різних критичних додатків, тобто мається на увазі, що мережа повинна використовуватися ефективно одночасно для різних додатків, критичних до використовуваних ресурсів.

Така здатність IP-мережі забезпечувати реалізацію різних рівнів обслуговування, що потребують ті чи інші мережні додатки, поряд із здійсненням контролю за використанням мережних ресурсів та параметрами продуктивності, приклади яких вже неодноразово наводилися вище – може

бути реалізована із застосуванням трьох технологій обслуговування: негарантованої доставки даних, гарантованого і диференційованого обслуговування [2, 6, 14].

1) Технологія негарантованої доставки даних (Best-Effort Service), яка передбачає, що мережні ресурси розподіляються між різними додатками на рівних умовах залежно від обсягу трафіку, але при цьому буде відсутнє жорстке закріплення ресурсів за будь-яким двоточковим з'єднанням. Тобто це говорить про те, що немає ніякої гарантії у доставці пакетів у правильній послідовності, немає ніякої різниці між будь якими видами трафіку, і що цей трафік буде доставлений у потрібний час або взагалі буде доставлений, тощо [10].

Принцип негарантованої доставки досить ефективний для сервісів, де дані можна передавати у відносному режимі часу (передача файлів, електронна пошта). Крім того, з урахуванням, що в транспортних мережах, які побудовані на базі волоконно-оптичних ліній зв'язку (ВОЛЗ), існує надлишок мережних ресурсів (смуга пропускання), цей принцип у певному ступені дозволяє забезпечити сьогодні вимоги сервісів, орієнтованих на передачу трафіку VoIP та інших додатків, що працюють у реальному часі. Однак, як тільки нестача мережних ресурсів стає відчутною, то це швидко призводить до зростання затримок пакетів і збільшення імовірності їх втрат. Для додатків, що працюють у реальному часі, вже не можуть бути забезпечені необхідні показники QoS. По-перше, це пояснюється основним принципом функціонування IP-мереж, який полягає у передачі даних у дейтаграмному режимі, тобто без здійснення управління процесом передачі та без встановлення з'єднання. По-друге, з появою нових додатків, зокрема тих, що функціонують у реальному часі (інтерактивна передача мови, відеоконференції та відеотелефонія та ін.), досить жорстко стає питання про забезпечення гарантованої QoS в мережах [10].

Таким чином можна зробити висновок, що технологія Best-Effort не відноситься до технологій, що забезпечують QoS, тому що гарантії щодо якості обслуговування є відсутніми, а також відсутні гарантії щодо забезпечення доставки пакетів [2, 6].

2) У технологію гарантованого обслуговування закладені принципи інтегрованого резервування ресурсів, а в основі її реалізації лежить використання моделі надання інтегрованих послуг (Integrated Services, IntServ), що розроблена робочою групою Integrated Services Working Group комітету IETF. Ця модель орієнтована на всебічну підтримку додатків, що функціонують у реальному часі передачі даних, які є чутливими до затримок. Механізми, що реалізують модель інтегрованих послуг, мають забезпечувати взаємодію всіх мережних пристроїв для підтримки гарантованої якості обслуговування, для чого здійснюється попереднє резервування мережних ресурсів уздовж всієї траєкторії руху визначеного потоку пакетів [6, 10].

3) Технологія диференційованих послуг (Differentiated Services, DiffServ), в основі реалізації якої лежить використання однойменної моделі, є логічним продовженням робіт IETF над архітектурою моделі IntServ. Її появі передували недоліки, що випливли у процесі застосування моделі IntServ: жорсткі гарантії якості обслуговування (hard QoS), низький рівень масштабування, що призвело до необхідності створення більш гнучких механізмів забезпечення якості обслуговування. Методи, що закладені у технології DiffServ, являють собою сукупність механізмів, які на відміну від IntServ забезпечують відносну або іншими словами «м'яку» (soft QoS) якість обслуговування [6, 10].

Основна ідея механізмів, що реалізуються технологією DiffServ, полягає у наданні диференційованих послуг для набору класів трафіку, які відрізняються вимогами до характеристик QoS. Як і у випадку механізмів технології IntServ, для реалізації диференційованих послуг широко застосовуються механізми, що входять до складу розглянутої вище базової моделі підтримки якості обслуговування в мультисервісних IP-мережах. Одним із ключових понять моделі DiffServ є угода про рівень обслуговування (SLA), що входить до складу механізмів QoS (рисунок 2.1) [10].

Далі проаналізуємо як досягаються необхідні вимоги щодо якості обслуговування на основі застосування розглянутих вище мережних механізмів з використанням технологічних моделей IntServ та DiffServ.

3 АНАЛІЗ АРХІТЕКТУРНИХ І ФУНКЦІОНАЛЬНИХ ПРИНЦИПІВ РЕАЛІЗАЦІЇ МОДЕЛЕЙ INTSERV І DIFFSERV В АСПЕКТІ ЗАБЕЗПЕЧЕННЯ QoS В МУЛЬТИСЕРВІСНИХ IP-МЕРЕЖАХ

3.1 Архітектура і функціонування моделі інтегрованих послуг (IntServ)

Модель інтегрованого обслуговування (IntServ) описується в документі RFC 1633. Вона була запропонована на початку 90-х років і створювалася для обслуговування одиничних потоків, яким надається два види послуг: гарантовані та з керованим рівнем навантаженням. Гарантовані послуги дозволяють забезпечити певному обсягу трафіку мінімальне значення затримки у разі проходження пакетів із кінця в кінець, тобто іншими словами дозволяють забезпечити наскрізну (із кінця в кінець – «End-to-End») якість обслуговування, гарантуючи потрібну пропускну здатність. Послуги з керованим рівнем навантаження надають певному обсягу трафіка обслуговування, що передбачене принципами технології негарантованої доставки даних (Best-Effort) (див. підрозділ 2.3). Таке обслуговування надається без жорстких гарантій у разі наявності низького мережного навантаження [14].

IntServ більше підходить для концентрації трафіка в прикордонній мережі IP і не рекомендована для застосування в транзитних мережах IP (через проблеми з масштабованістю). Проаналізуємо архітектуру моделі інтегрованого обслуговування (рисунок 3.1) [14].

У кожному мережному вузлі, що підтримує технологію IntServ має бути кілька обов'язкових функціональних компонентів [14]:

- класифікатор – направляє пакет, що надходить, в один із класів обслуговування відповідно до інформації, що отримана із заголовків пакета (мережного і транспортного рівнів). Клас обслуговування реалізується у вигляді окремої черги і всі пакети в його межах повинні одержувати однаковий QoS;

- диспетчер пакетів (ДП) – вилучає з кожної черги пакети і направляє їх на канальний рівень. У моделі інтегрованих послуг застосовується двоступінчатий ДП. Всі пакети, які надходять, в рамках ізоляції потоків, що отримують гарантовані послуги, від всіх інших, – обробляються за дисципліною обслуговування черг WFQ. Потоки з керованим навантаженням і потоки з негарантованою доставкою даних розділяються за допомогою пріоритетів;

- блок управління доступом – вирішує про надання можливості отримання трафіком потрібних ресурсів у необхідній кількості, не впливаючи при цьому на раніше надані гарантії. Управління доступом здійснюється на кожному мережному вузлі для забезпечення прийняття або відхилення запиту на виділення потрібних ресурсів по всьому шляху просування потоку;

- протокол резервування ресурсів – інформує всіх учасників з'єднання (зокрема: відправника, одержувача, проміжні маршрутизатори) про необхідні параметри здійснення обслуговування. Для моделі IntServ рекомендується використовувати протокол RSVP.

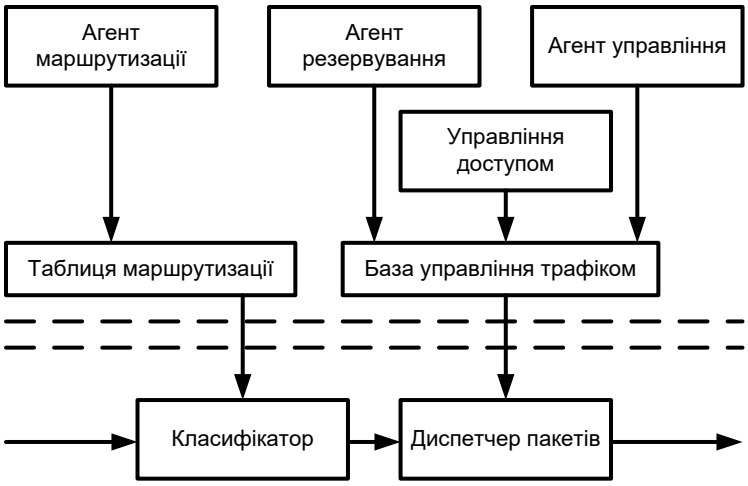

Рисунок 3.1 – Архітектура моделі інтегрованого обслуговування (IntServ)

Технологічна модель IntServ у поєднанні з протоколом RSVP дає можливість організувати гнучке обслуговування трафіка різних типів з максимальним урахуванням потреб у ресурсах кожного додатка, а

використання алгоритму WFQ для обслуговування черг пакетів гарантує мінімізацію значень часових затримок. Ця особливість робить модель інтегрованого обслуговування досить ефективною у разі здійснення обслуговування мультимедійного трафіка. За своєю сутністю, RSVP є протоколом сигналізації, відповідно до якого здійснюється резервування та управління ресурсами з метою гарантії забезпечення QoS, тому саму модель можна стисло охарактеризувати як «резервування ресурсів». Протокол сигналізує про запити на здійснення резервування ресурсів по всьому доступному шляху в мережі. При цьому він не підтримує алгоритми маршрутизації, а використовує інші, призначені для цього протоколи маршрутизації трафіку в IP-мережах [14, 15].

Резервування проводиться для певного конкретного потоку пакетів IP-мережі перед початком передачі цього потоку. Визначення пакетів, що належать одному потоку (ідентифікація потоку), робиться за спеціальною міткою, що розташовується в основному заголовку кожного IP-пакету. Після резервування шляху для подальшого проходження потоку починається його передача. Пакети потоку обслуговуються по всьому з'єднанні «із кінця в кінець» із заданою якістю [10].

Переважне застосування протоколу RSVP – це мультимедійні додатки із груповою розсилкою (наприклад, додатки аудіо- і відеоконференцій). Але цей протокол також можна застосовувати для резервування смуги пропускання для односпрямованого трафіка (наприклад для трафіка мережної файлової системи (Network File System, NFS) або управляючого трафіка віртуальних приватних мереж (Virtual Private Networks, VPN)).

Проаналізуємо більш докладно як працює цей протокол. Для цього розглянемо функціональну модель протоколу RSVP і його основні модулі (рисунок 3.2) [2, 6].

Мережні вузли використовують протокол RSVP для запитів в мережі потрібного рівня QoS від імені потоку даних відповідного додатка. Ці запити передаються по мережі при проходженні кожного вузла, який забезпечує

передачу потоку. Протокол RSVP намагається зарезервувати потрібні ресурси для потоку даних додатку на кожному із цих вузлів. Маршрутизатори, що підтримують роботу з RSVP, допомагають доставити потрібні потоки даних у визначену точку за призначенням [2, 6].

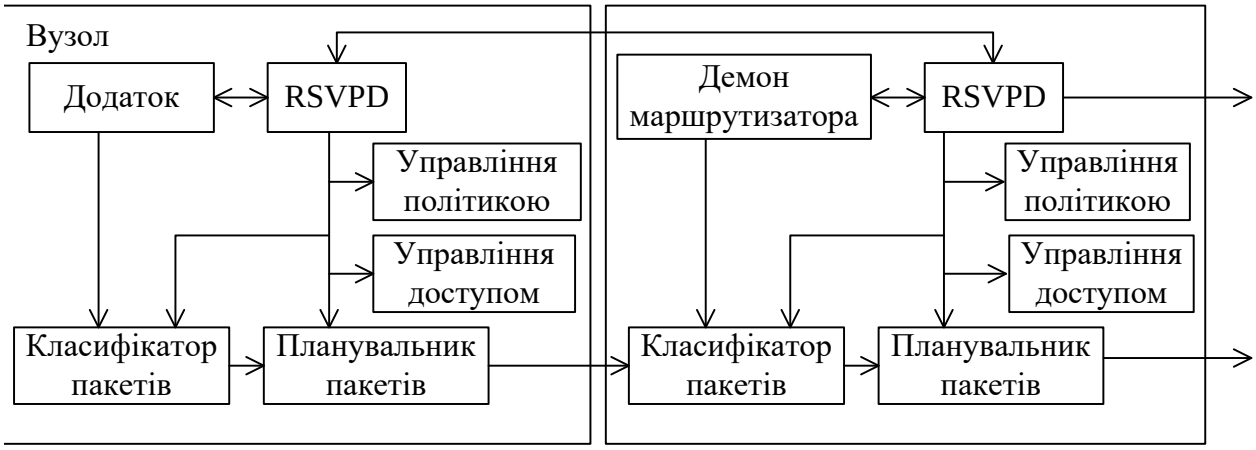

Рисунок 3.2 – Функціональна модель протоколу RSVP і його основні модулі

Перед тим, як зарезервувати ресурси, RSVP-демон маршрутизатора (RSVPD) з'єднується із двома локальними модулями прийняття рішення: модулем управління політикою і модулем управління доступом. Модуль управління політикою визначає, чи має користувач права адміністратора для того, щоб здійснити процедуру резервування. Модуль управління доступом визначає, чи має вузол достатньо вільних ресурсів для забезпечення потрібного рівня якості обслуговування, що був запитаний. У випадку, коли якась з цих перевірок не пройшла, RSVPD генерує повідомлення про помилку для додатка, який створив запит на здійснення резервування. У випадку, коли обидві перевірки були пройдені, RSVPD налаштовує відповідні параметри модулів класифікатора пакетів і планувальника пакетів для отримання потрібного рівня QoS. Задачею класифікатора пакетів є визначення класу якості обслуговування для кожного пакета. Задачею планувальника пакетів є здійснення управління передачею пакетів, опираючись на їх клас. Підтримку потрібної QoS на рівні планувальника

реалізують алгоритми обслуговування черг WFQ і WRED, що вже згадувалися вище [6].

Під час відпрацювання процесу ухвалення рішення модулем управління доступом, резервування смуги пропускання, що була запитана, виконується тільки в тому випадку, якщо для визначеного класу трафіка буде досить тієї її частини, яка залишилася ще не розподіленою. В іншому випадку запит на доступ буде відхилений, але трафік при цьому все одно передається з визначеною для цього класу QoS. Потрібно зазначити, що навіть якщо запит на здійснення доступу буде відхилений на одному або кількох маршрутизаторах, то у багатьох випадках модуль все одно все ще може реалізувати прийнятну QoS, здійснивши резервування на перевантажених маршрутизаторах. Це можливо реалізувати через те, що інші потоки даних пакетів можуть не повністю використовувати смугу, яка була ними замовлена [6].

Резервування має завжди проходити по одному ж і тому самому одноадресному шляху або по багатоадресному дереву. У випадку, якщо канал зв'язку вийшов з ладу – маршрутизатор має обов’язково попередити про це RSVPD, щоб RSVP-повідомлення, що ним генеруються, передавалися по новому шляху [6].

Розглянемо послідовність здійснення процесу встановлення резервування, який складається з п'яти окремих кроків [2, 6]:

1) Відправники даних посилають управляючі повідомлення PATH за тим же шляхом, за яким вони відправляють звичайні дані. У цих повідомленнях описуються дані, які або вже відправляються, або тільки будуть відправлятися;

2) Кожний маршрутизатор, що підтримує RSVP, перехоплює PATH-повідомлення, зберігає IP-адресу попередньої точки призначення, записує замість неї свою власну адресу і відправляє таким чином відредаговане повідомлення далі за тим же шляхом, за яким здійснюється передача даних додатку;

3) Станції-одержувачі вибирають підмножину сеансів, для яких вони одержали інформацію у PATH-повідомленнях і за допомогою RESV-повідомлення запитують резервування потрібних мережних ресурсів у попереднього маршрутизатора. RESV-повідомлення йдуть від одержувача до відправника в протилежному напрямку по маршрутові, який був до цього пройдений PATH-повідомленнями;

4) Маршрутизатори, що підтримують RSVP, визначають можливості щодо задоволення цих RESV-запитів. У випадку неможливості, вони відмовляють у здійсненні резервування. У іншому, позитивному випадку, вони об'єднують отримані запити на резервування і відсилають запит попередньому маршрутизаторові;

5) Відправники, у разі отримання запитів на резервування ресурсів від відповідних маршрутизаторів, вважають, що резервування ресурсів є реалізованим. Тобто фактично резервування ресурсів реалізується RESV-повідомленнями.

Реалізація процесу здійснення резервування мережних ресурсів за протоколом RSVP спрощено наведений на рисунку 3.3 [2].

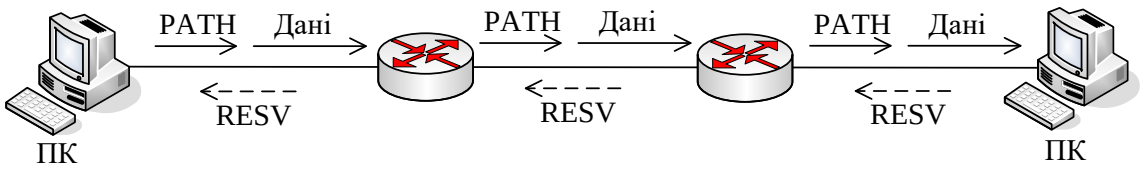

Рисунок 3.3 - Реалізація процесу резервування ресурсів за протоколом RSVP

Слід зазначити, що у протоколі RSVP використовуються сім типів повідомлень: два обов'язкові: PATH і RESV, – і п'ять опціональних: PATH ERROR, PATH TEARDOWN, RESV ERROR, RESV CONFIRM та RESV TEARDOWN. Маршрутизатори, що підтримують RSVP, та клієнти використовують ці повідомлення для створення та підтримки станів резервування.

Зазвичай протокол RSVP працює безпосередньо поверх протоколу IP. Отже, повідомлення RSVP є ненадійними дейтаграмами. Вони допомагають створювати у маршрутизаторах гнучкі стани, які необхідно періодично оновлювати. Нижче, у таблиці 3.1, наведені типи повідомлень RSVP [6].

Таблиця 3.1 – Типи повідомлень протоколу RSVP

PATH SETUP (відкривання шляху) або PATH MESSAGE (повідомлення про шлях)	Відправляється ініціатором з’єднання для формування шляху до кінцевої системи.
RESV SETUP (резервування ресурсу) або RESV MESSAGE (повідомлення про ресурс)	Відправляється ініціатором з’єднання у зворотному напрямку для резервування ресурсів по всій довжині шляху з’єднання.
PATH TEAR (розрив шляху)	Відправляється ініціатором з’єднання для розриву шляху.
RESV TEAR (закриття резервування)	Відправляється ініціатором з’єднання для закриття резервування.
PATH ERROR (похибка шляху)	Відправляється маршрутизатором ініціатору з’єднання, для попередження останнього про похибку шляху.
RESV ERROR (похибка резервування)	Відправляється маршрутизатором ініціатору з’єднання, для попередження останнього про похибку у здійсненні резервування ресурсу.
RESV CONFIRM (підтвердження резервування)	Відправляється кінцевою системою або маршрутизатором у відповідь на запит про підтвердження, що може бути включеним в повідомлення про резервування ресурсу. Це повідомлення не буде відправлене, якщо в повідомленні RESV SETUP був відсутній запит про RESV CONFIRM.

Таким чином можна бачити, що протокол RSVP є основним дієвим інструментом для забезпечення гарантованої QoS в рамках реалізації моделі IntServ, але цей протокол не вирішує всі проблеми, що пов'язані з якістю обслуговування. Зокрема найбільший недолік технології IntServ як раз і пов'язаний із низькою масштабованістю RSVP, що особливо проявляється на високошвидкісних магістралях транспортних мереж. Так обсяг ресурсів, що

необхідні маршрутизатору для обробки та зберігання інформації RSVP, зростає пропорційно до кількості потоків QoS. Вимірювання трафіку у мультисервісних IP-мережах показують, що більшість з'єднань «із кінця в кінець» існує дуже недовго, і в кожен момент часу магістральним маршрутизатором підтримується кілька тисяч активних з'єднань. Отже, багаточисленні потоки моделі IntServ у каналі з великою пропускною здатністю значно збільшують навантаження маршрутизатори. Більш того, щоразу у разі зміни топології, все зарезервовані шляхи необхідно прокладати заново, що у свою чергу призводить до зростання обсягів службової інформації та великих часових витрат на організацію резервування. Останні дві особливості є недоліками використання безпосередньо протоколу RSVP [10, 14].

Таким чином, модель IntServ у поєднанні з протоколом RSVP дозволяє організувати гнучке обслуговування різнотипного трафіку із максимальним врахуванням потреб кожного додатка, а використання алгоритмів WFQ та WRED для обслуговування черг потоків мультисервісного IP-трафіку гарантує мінімальні значення часових затримок. Ця особливість робить модель IntServ ідеальною для здійснення обслуговування мультимедійного трафіку.

Однак слід пам'ятати, що продуктивність технології інтегрованих послуг залежить від кількості потоків, що обробляються, отже, таку сервісну модель практично неможливо реалізувати в мережі з дуже великою кількістю користувачів. Тому модель IntServ із резервуванням за протоколом RSVP не практично реалізувати у великомасштабних середовищах. Для великих мереж потрібна більш проста та масштабована технологія, а сфера застосування техгнології IntServ обмежується внутрішніми та кінцевими мережами. У найкращому разі, магістральний маршрутизатор має можливість резервувати ресурси для кількох тисяч потоків та здійснювати управління чергами для кожного з них [10, 14].

3.2 Архітектура і функціонування моделі диференційованих послуг (DiffServ)

3.2.1 Мережна архітектура моделі DiffServ, її компоненти і функціональні модулі

Модель диференційованих послуг (DiffServ) була вперше описана в 1999 році в документі RFC-2475, який є логічним продовженням робіт IETF над архітектурою моделі IntServ. Наведені вище недоліки технологічної моделі інтегрованих послуг (жорсткі гарантії QoS, низький рівень масштабування) призвели до необхідності створення більш гнучких механізмів забезпечення якості обслуговування. Механізми, що становлять модель DiffServ, на відміну від механізмів моделі IntServ забезпечують «відносну» або «м'яку» якість обслуговування. Їх основна ідея полягає в диференціюванні трафіку шляхом його розбивки на класи з різним пріоритетом. Тут, як і у разі надання інтегрованих послуг, широко застосовуються механізми, що входять до складу розглянутої у попередньому розділі базової моделі підтримки QoS в мультисервісних IP-мережах. Одним із центральних понять моделі DiffServ є угода про рівень обслуговування, що входить до складу механізмів QoS на площині управління менеджменту [10].

Вимоги до необхідного набору показників якості обслуговування задаються у спеціальному однобайтовому полі кожного пакета: у октеті типу обслуговування (Type of Service, ToS) протоколу IPv4 або в октеті класу трафіку (Traffic Class, TC) протоколу IPv6. У моделі DiffServ це поле називається байтом диференційованої послуги (DS-байтом) або полем коду диференційованих послуг (Differentiated Services Code Point, DSCP). Іншими словами, зміст DS-байта визначає вид диференційованих послуг, що надаються, а вибрати необхідний рівень послуг клієнт може шляхом встановлення відповідного його значення для кожного пакета певного додатку. Формат DSCP-байта наведено на рисунку 3.4 [10].

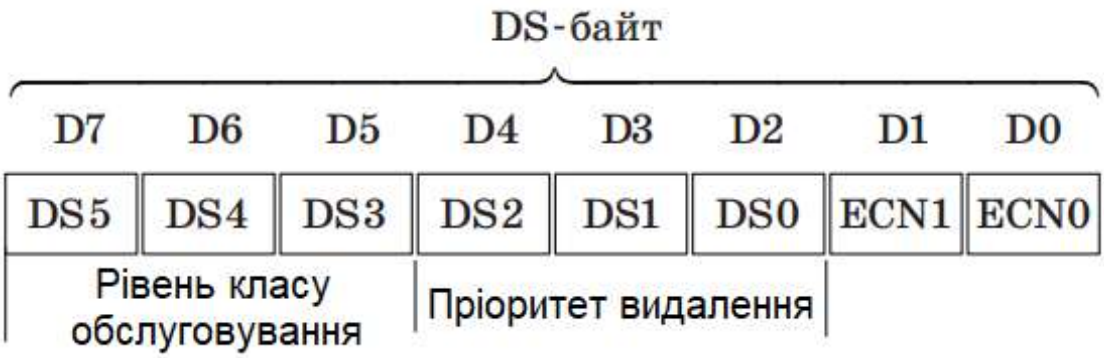

Рисунок 3.4 – Формат поля кода диференційованих послуг (DSCP)

Біти DS5 - DS3 кодують рівень класу обслуговування від 0 (мінімальний пріоритет) до 7 (максимальний пріоритет). Біти DS2 - DS0 кодують пріоритет видалення від 0 (пріоритет видалення максимальний), до 7 (пріоритет видалення мінімальний). У результаті виходить код пріоритету – число від 0 до 63, де чим більше число, тим трафік більш пріоритетний. Наприклад, для VoIP-трафіку застосовується клас сервісу 5 (байт DSCP дорівнює 0xA0 або 10100000b), а для звичайного трафіку – клас сервісу 0 (байт DSCP дорівнює 0x00 або 00000000b). Біти ECN1, ECN0 не визначено. Клас обслуговування тут означає механізм обробки та просування пакету з даного вузла до наступного вузла (так звана PHB-політика або політика покрокового обслуговування – Per-Hop Behavior, PHB) відповідно до необхідної якості обслуговування [10, 16].

Як можна бачити, PHB-політика визначає функціонування мережного вузла відносно пакетів з певним значенням байту DSCP. Всі пакети потоку трафіка із специфічною вимогою до обслуговування несуть у собі одне і те саме значення цього байту. Наприклад, PHB-політика визначає спосіб резервування ресурсів маршрутизатора, що обслуговує потоки трафіка. Правила PHB реалізуються за допомогою декількох механізмів управління буфером і планування обробки пакетів [6, 16].

Мережна архітектура моделі DiffServ представлена на рисунку 3.5 [16].

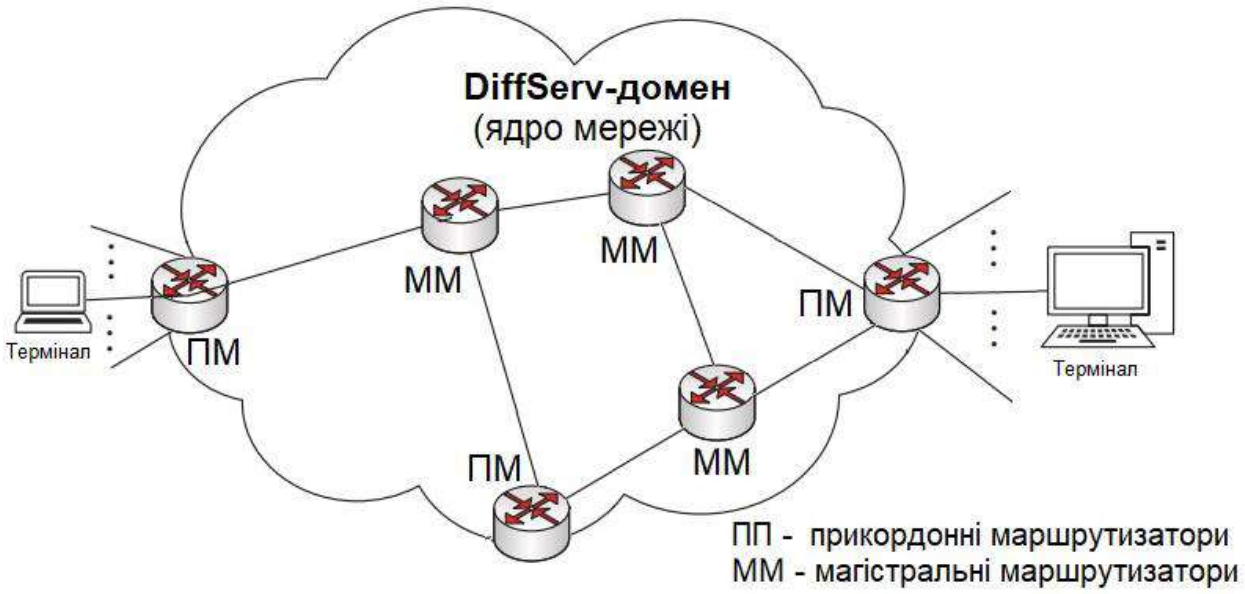

Рисунок 3.5 – Мережна архітектура моделі DiffServ

У наведеній на рисунку 3.5 моделі архітектура мережі представляється у вигляді двох сегментів – прикордонних ділянок і ядра. На вході в мережу у вузлі доступу (прикордонному маршрутизаторі) пакети класифікуються (механізм Traffic classification) для того, щоб була можливість виділити пакети одного потоку, який характеризується загальними вимогами QoS. Потім трафік піддається процедурі нормування (механізм Traffic conditioning), яка передбачає вимірювання характеристик трафіку та порівняння результатів вимірювань з характеристиками, що прописані в угоді SLA. Якщо умови SLA порушуються, частина пакетів може бути відкинута. Магістральні маршрутизатори становлять ядро мережі і забезпечують передачу пакетів відповідно до необхідного рівня QoS [10].

Всі вузли, що знаходяться всередині домена Diffserv, визначають PHB-політику, яка має застосовуватися до кожного IP пакета на основі значення байту DSCP, що зберігається в пакеті. Крім того, прикордонні маршрутизатори виконують важливу функцію формування трафіка, що надходить далі в домен. Формування трафіка містить у собі виконання таких функцій, як: класифікація пакетів (встановлення значення поля DS-байту), обмеження трафіка [2].

Формування трафіка зазвичай виноситься на вхідний інтерфейс домена Diffserv, на який надходять пакети. Формування має велике значення для забезпечення управління трафіком, що надходить в домен, оскільки саме в цьому випадку для кожного пакета мережа може визначити відповідну йому PHB-політику.

Таким чином можна бачити, що базовими функціональними модулями архітектурної моделі диференційованих послуг є формувачі трафіка та пристрої, що реалізують PHB-політику. Їх стисла характеристика приведена у таблиці 3.2 [6].

Таблиця 3.2 – Характеристика найважливіших функціональних модулів архітектурної моделі Diffserv

Функціональний модуль	Розташування	Функція	Дія
Формувачі трафіка	Вхідний інтерфейс прикордонного маршрутизатора домена DiffServ	Класифікація пакетів, вирівнювання і обмеження трафіка	Обмеження вхідного трафіка і установка значення DS-байту з урахуванням профілю трафіка
Пристрої, що реалізують PHB-політику	Всі без виключення маршрутизатори домена DiffServ	Розподіл ресурсів і політика відкидання пакетів	PHB-політика обробки пакетів визначається на основі характеристик QoS, що відповідають заданому значенню поля DSCP

Формувач трафіка – це функціональний модуль, що виконує різні функції QoS, які повинні бути реалізовані в прикордонних маршрутизаторах. Його структура наведена на рисунку 3.6 [17].

Як було сказано вище, у разі надходження в домен DiffServ всім пакетам має бути присвоєно значення поля коду DSCP, виходячи з типу інформації, що передається. Код диференційованої послуги є полем, на підставі значення якого визначається спосіб обробки пакету в домені DiffServ. Для цього пакети обробляються класифікатором. Класифікатор

пакетів визначає підмножину трафіку, яка може отримати сервіс певного рівня шляхом обробки згідно з правилами PHB-політики [6, 17].

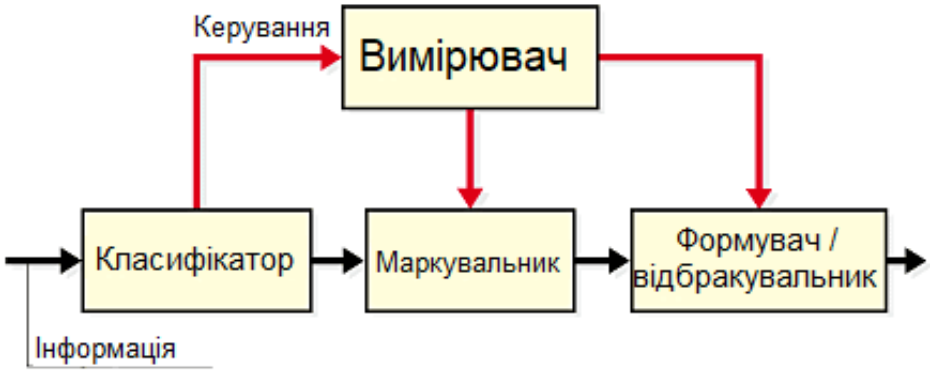

Рисунок 3.6 – Структурна реалізація формувача трафіка

Наступна за класифікатором група елементів називається блоком формування трафіку. Ця група включає в себе: вимірювач, маркувальник, формувач/відбракувальник. Інформація про класифікацію пакета передається у вимірювач. Він визначає відповідність параметрів потоку трафіка характеристикам, які задані у профілі за допомогою угоди про формування трафіку. Ця угода визначає правила класифікатора та відповідних профілів трафіку, а також правила маркування, формування та відкидання пакетів, що застосовуються до тих або інших потоків [17].

Таким чином, в залежності від результату роботи вимірювача налаштовуються маркувальник, відбракувальник та формувач. Маркувальник встановлює код поля DSCP. Для певної групи пакетів маркер може бути налаштований для встановлення єдиного значення коду або маркування пакета одним з декількох кодів відповідно до результатів роботи вимірювача. Формувач використовується для здійснення затримки відправлення пакетів у разі, якщо потоком були перевищені ліміти, які визначені у профілі. Відбракувальник робить видалення пакетів для приведення параметрів потоку до конфігурації профілю. Потрібно зазначити, що блок формування трафіку може містити не всі елементи, які показані на рисунку 3.6, а лише

деякі з них. На внутрішніх вузлах домену DiffServ зазвичай проводяться лише процедури класифікації трафіку та простого пересилання пакетів відповідно до PHB-політики. Проте процедури формування трафіку також дозволяється робити і у внутрішніх вузлах [17].

3.2.2 Особливості застосування PHB-політики, її основні типи і способи формування

Далі більш докладно дослідимо механізми політики покрокового обслуговування або PHB-політики.

Вже неодноразово зверталася увага на те, що вузли мережі, в яких реалізована підтримка диференційованого обслуговування, використовують байт DSCP в заголовку IP-пакета для визначення відповідної цьому пакету PHB-політики. Вона може бути визначена в термінах пріоритету в поданні мережних ресурсів відповідно до інших PHB-політик або ж опираючись на характеристики трафіка, які постійно вимірюються (затримка пакетів, джиттер, рівень втрати пакетів, тощо). Тобто іншими словами PHB-політика визначає поводження вузла мережі, за яким ведеться спостереження як би ззовні відносно пакетів, що надходять, не нав'язуючи при цьому конкретні дії або реалізацію [2].

У якості стандартного покрокового обслуговування в моделі диференційованого обслуговування можна розглядати принцип негарантованої доставки (Best-Effort). Відповідно з особливостями реалізації моделі Diffserv кожній PHB-політиці рекомендується призначити певне значення байту DSCP, однак сервіс провайдер може вільно вибрати ці значення для своєї власної мережі і вони будуть не співпадати з рекомендованими. Рекомендоване значення поля коду DSCP для політики за принципом Best-Effort становить 000000 [2].

PHB-політика відповідає певному класу трафіка і залежить від низки факторів [2]:

- інтенсивності вхідного навантаження або потоку пакетів для заданого класу трафіка. Цей параметр контролюється формувачем трафіка, що реалізується прикордонним маршрутизатором;

- розподілу мережних ресурсів для заданого класу трафіка. Цей параметр контролюється відповідними функціями розподілу мережних ресурсів, що забезпечуються вузлами Diffserv-домена;

- рівня втрати трафіка. Цей параметр залежить від механізмів відкидання пакетів, що підтримуються вузлами Diffserv-домена.

Розглядають два основних типи PHB-політики: негайної передачі і гарантованої доставки.

Політика негайної передачі пакетів (Expedited Forwarding PHB, EF PHB) застосовується для забезпечення обслуговування пакетів за принципом «із кінця в кінець» у мережних вузлах Diffserv-домена. Її відмінними рисами є мала затримка, низький рівень втрати пакетів, незначна величина джиттеру і гарантована смуга пропускання. Ця політика задіється для обслуговування потоків пакетів, що генеруються такими додатками, як додатки VoIP, відеододатки, а також для забезпечення послуг передачі інформації по віртуальних орендованих каналах, тому що ця послуга є двоточковим з'єднанням кінцевих мережних вузлів DiffServ-домена. Тип обслуговування, що реалізується політикою EF PHB, часто відносять до послуг високого класу [2, 6].

Головними факторами, які призводять до великої затримки пакетів та джиттеру, є затримки, що пов'язані з виникненням великих накопичених черг. Подібні черги характерні для перевантажених ділянок мережі. Причиною перевантаження мережі є переважання інтенсивності вхідного потоку трафіку над інтенсивністю його вихідного потоку. Один із підходів мінімізувати часові затримки пакетів, які з’являються за рахунок появи великих черг – це зробити обмеження максимальної інтенсивності вхідного потоку трафіка мінімальною інтенсивністю його вихідного потоку. Політика EF PHB дозволяє встановлювати потрібне значення інтенсивності вихідного потоку трафіка, в той час як за інтенсивністю вхідного потоку здійснюється

контроль формувачами трафіка, які реалізовані в прикордонних маршрутизаторах мережі [6].

Тому що згідно із EF PHB вхідні пакети не повинні утворювати чергу (черга може бути, але дуже малого розміру), інтенсивність вихідного потоку трафіка має дорівнювати інтенсивності вхідного потоку або навіть перевищувати її. Потрібно звернути увагу на те, що інтенсивність вихідного потоку (смуга пропускання) не повинна залежати від інших потоків трафіка. Як правило, інтенсивність як вхідного, так і вихідного потоків, вимірюється з інтервалами, що дорівнюють часу, який потрібний для передачі так званого пакета максимального розміру (MTU-пакета), який може бути переданий через інтерфейс маршрутизатора [6].

Маршрутизатор може виділити ресурси у достатній кількості для забезпечення потрібної інтенсивності вихідного трафіка для відповідного інтерфейсу за рахунок використання різних функціональних реалізацій політики EF PHB. У разі передачі трафіка через перевантажений мережний сегмент (тобто иожлива наявність великих накопичених черг), ця функціональна можливість може бути реалізована за рахунок застосування різних механізмів та алгоритмів обслуговування черг, про які неодноразово вже згадувалося вище [2, 6].

Другий тип PHB-політики, що є прийнятною для більшості TCP-додатків – це політика гарантованої доставки пакетів (Assured Forwarding PHB, AF PHB), яка являє собою механізм, що дозволяє сервіс-провайдеру забезпечити кілька різних рівнів надійності доставки IP-пакетів, що надходять з DiffServ-домена клієнта. Політика AF PHB має у наявності різні рівні обслуговування для кожного із чотирьох класів AF-трафіка, кожному з яких відповідає своя власна черга пакетів, що дозволяє реалізувати ефективне управління смугою пропускання. Також кожний клас AF-трафіка характеризується наявністю 3-х рівнів пріоритету відкидання пакетів (низький, середній і високий). Це дозволяє задіяти механізм управління чергою по типу механізму, що реалізується алгоритмом RED [2, 6].

Існує три способи щодо формування політики PHB в DiffServ-мережі [6]:

- ініціалізація мережі;
- сигналізація про якість обслуговування;
- диспетчер політик.

Перший із зазначених вище способів формування політики PHB в DiffServ-мережі полягає в ініціалізації ресурсів мережі з використанням евристичних методів або техніки систематичного моделювання. Потрібно зазначити, що цей метод може бути застосований тільки в мережах досить невеликого розміру, для яких політики якості обслуговування і профілі трафіка залишаються незмінними протягом достатньо довгого інтервалу часу [6].

Другий спосіб формування PHB-політики полягає в тому, що додатки інформують мережу про вимоги до QoS за допомогою протоколу RSVP, з погляду якого DiffServ-домен виступає у якості ще однієї ланки мережі, яка вимагає забезпечити управління доступом. За допомогою протоколу RSVP можна встановити відповідність між запитами до QoS додатків та класами послуг DiffServ-мережі. Наприклад, гарантованому обслуговуванню RSVP може бути зіставлена у відповідність DiffServ-послуга політики EF PHB [6].

Сигналізація про QoS є досить масштабованим рішенням у великих мережах, оскільки протокол RSVP виконується лише у прикордонних маршрутизаторах DiffServ-домену, як це показано на рисунку 3.7 [6].

Рисунок 3.7 – Передача сигнальної інформації протоколу RSVP через DiffServ-мережу

Встановлення відповідності між резервуванням ресурсів протоколу RSVP та DiffServ-класами проводиться на кордоні DiffServ-мережі. Протокол RSVP має досить широку підтримку, тому прикордонні маршрутизатори, які проінформовані про існуючу політику, можуть використовувати його для висування вимог щодо якості обслуговування, не звертаючи при цьому уваги на можливе збільшення масштабів мережі. Тобто це рішення добре підійде для застосування у великих корпоративних мережах [6].

Третій спосіб формування політики обумовлює вибір рівнів якості обслуговування, що застосовується до потоку трафіка. Політики призначаються за допомогою протоколу розповсюдження політик (Common Open Policy Service, COPS), що розроблений групою IETF. В термінології протоколу COPS централізований сервер політик називається точкою визначення політики (Policy Decision Point, PDP). Вузол мережі, якому нав'язується політика, отримав назву точки застосування політики (Policy Enforcement Point, PEP). PDP-сервер використовує протокол COPS для завантаження політик у PEP-вузли мережі. PEP-пристрій може згенерувати повідомлення, у якому він інформує PDP-сервер про неможливість реалізації запропонованої PDP-сервером політики [6].

Таким чином, підводячи підсумок щодо проведеного аналізу моделі DiffServ, можна сказати, що її перевагами є відносна простота і висока масштабованість. З цієї причини цій моделі відведено місце на магістральних та високошвидкісних ділянках мультисервісних IP-мереж.

3.3 Порівняльний аналіз технологічних моделей IntServ і DiffServ

Порівняльна аналіз параметрів моделей IntServ і DiffServ приведена в таблиці 3.3 [14, 15].

Із наведеного в табл. 3.3 порівняльного аналізу, можна зробити висновок, що на цей час не існує оптимальної універсальної моделі QoS, яка здатна задовольнити одночасно всі вимоги для побудови мультисервісних IP-

мереж. Однак існуючі моделі і механізми забезпечення якості обслуговування в мультисервісних IP-мережах можливо ефективно застосовувати, якщо виділити декілька основних вимог до QoS, а іншими знехтувати.

Таблиця 3.3 - Порівняльний аналіз параметрів моделей IntServ і DiffServ

Параметр	IntServ	DiffServ
Метод забезпечення QoS	Резервування	Пріоритезація
Число класів QoS, що обслуговуються	3	3
Перелік показників якості, що задаються	Смуга пропускання	Швидкість передачі трафіка
	Максимальна мережна затримка	Мережна затримка
	Джитер	Коефіцієнт втрати пакетів
Необхідність використання додаткових протоколів	RSVP	Немає
Вимоги до продуктивності маршрутизаторів	Високі	Низькі
Ефективність масштабування мережі	Невисока	Висока
Сумісність обладнання різних виробників	Середня	Висока
Гарантованість забезпечення якості	Висока	Середня

Простота пріоритезації трафіка в моделі DiffServ, невибагливість до обладнання, гнучка можливість масштабування, набагато менші витрати на реалізацію в порівнянні з IntServ, економія часу і трафіка, підвищена надійність за рахунок того, що класифікація відбувається на кордоні DiffServ-домена без виконання службових запитів, визначає гнучкість і ефективність DiffServ. Однак ця технологічна модель не дає повної гарантії забезпечення якості обслуговування, а лише забезпечує відносне збільшення смуги пропускання для більш пріоритетних потоків. Модель DiffServ підходить для застосування у великих локальних обчислювальних мережах і

територіально розподілених мережах, на стику мереж провайдерів, у каналах з малою пропускною здатністю [14, 15].

Переваги технологічної моделі IntServ полягають у чітко визначеній і гарантованій пропускній здатності, а значить, у більш високому ступені деталізації. За процесами, що створюються цією моделлю, легко здійснювати контроль, оскільки можна стежити за кожним маршрутом і з'єднанням. Однак його складно реалізувати для вже існуючих мереж і мережних додатків, тому що використовується протокол RSVP, підтримку якого мають забезпечувати маршрутизатори і сервісні додатки. Протокол RSVP не використовує механізми, які можуть запобігти втраті його службових повідомлень, за рахунок чого зменшується надійність, а у великих мережах можуть виникнути проблеми сумісності обладнання різних виробників. Механізми підтримки QoS, що реалізуються моделлю IntServ, є прийнятними у корпоративних мережах для вирішення обмеженого кола задач, але вони погано підходять для використання у великих високошвидкісних магістральних мережах і Інтернет [14, 15].

3.4 Аналіз принципів взаємодії технологічних моделей IntServ і DiffServ для забезпечення QoS

Принципи організації взаємодії моделей IntServ (протокол RSVP) та DiffServ для надання потрібної якості обслуговування «з кінця в кінець» регламентуються стандартом RFC 2998, що був створений і опублікований наприкінці 2000 р. Вище у цьому розділі було показано, що слабкі місця однієї моделі є можливість компенсувати відповідними рішеннями іншої. Зокрема, наприклад, погана масштабованість моделі IntServ на протяжних магістральних ділянках мережі може бути замінена на більш гнучку у цьому питанні модель DiffServ, але за допомогою протоколу RSVP можна практично повністю вирішити проблему з невизначеністю сервісу, що одержується в «чистій» DiffServ-мережі [14].

Основною проблемою у разі організації такої взаємодії є відповідність мережних ресурсів, що запитуються за допомогою протоколу RSVP та надаються в DiffServ-регіоні. Під DiffServ-регіоном мається на увазі безперервна послідовність DiffServ-доменів, у межах яких можуть надаватися DiffServ-послуги. Можлива організація двох варіантів взаємодії [14]:

- DiffServ-регіон не підтримує сигналізацію за протоколом RSVP, а мережні ресурси виділяються на статичній основі;

- обробка повідомлень RSVP проводиться безпосередньо у DiffServ-регіоні.

У першому випадку спільна робота ґрунтується на статичній угоді клієнта з оператором SLS (Service Level Specification – специфікація рівня сервісу). У найпростішій ситуації ця угода описує значення пропускної здатності, що отримується трафіком користувача, в DiffServ-мережі. У цьому випадку відправник (Tx) створює повідомлення PATH, які передаються до вузла-одержувача (Rx) через DiffServ-регіон, як це показано на рисунку 3.8 [14].

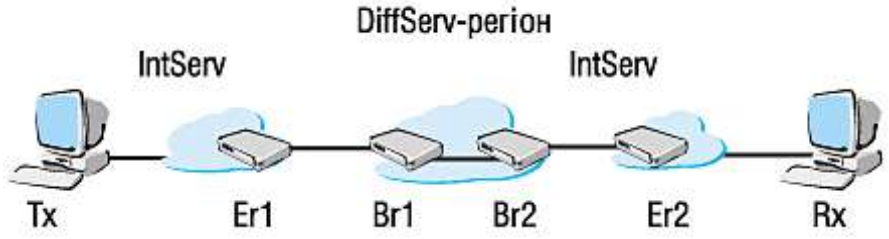

Рисунок 3.8 – Схема взаємодії моделей IntServ и DiffServ

При проходженні через DiffServ-регіон вміст RSVP-повідомлень ігнорується і вони передаються як звичайні пакети з даними. У разі, якщо вузол Rx отримав повідомлення PATH, – генерується запит на резервування ресурсів (RESV), який потім передається назад до вузла Tx. У разі успішної обробки запиту кожним маршрутизатором, що підтримує роботу з повідомленнями протоколу RSVP, та проходження через DiffServ-регіон повідомлення RESV досягає маршрутизатора Er1. Цей маршрутизатор на

підставі угоди SLS здійснює порівняння мережних ресурсів, що запитуються в повідомленні RESV, та ресурсів, які є доступними у DiffServ-регіоні. Якщо Er1 підтверджує запит, тоді повідомлення RESV надсилається далі до вузла Tx. У іншому випадку повідомлення відкидається, а вузлу Rx буде надіслане повідомлення про помилку. В повідомленні, що отримується вузлом Tx, може міститися інформація про маркування певним кодом пакетів, що адресуються вузлу Rx. Значення коду визначається за замовчуванням або безпосередньо із повідомлення RESV [14].

У другому випадку передбачається, що прикордонні маршрутизатори в регіоні DiffServ (наприклад, маршрутизатор Br1) підтримують протокол RSVP. Зазначимо, що, незважаючи на підтримку сигналізації за протоколом RSVP, обробляються не поодинокі потоки пакетів, а лише агреговані потоки, як у мережах, що реалізують модель IntServ/RSVP. Порядок обміну повідомленнями RSVP буде таким же самим, як і в попередньому випадку. Однак завдяки підтримці протоколу RSVP у DiffServ-регіоні блок управління доступом є частиною DiffServ-мережі. У кінцевому підсумку, маршрутизатор Br1 має можливість безпосередньо обробити RSVP-запит, виходячи з доступності мережних ресурсів [14].

Таким чином, спільне функціонування моделей забезпечення якості обслуговування IntServ і DiffServ є оптимальним варіантом надання необхідного QoS «із кінця в кінець». Реалізація такої моделі дозволить значно підвищити якість надання мультимедійних послуг в мультисервісних IP-мережах, а також підвищити продуктивність вже давно існуючих традиційних сервісів.

4 ЗАГАЛЬНА МЕТОДИКА ОЦІНКИ МЕРЕЖНИХ ХАРАКТЕРИСТИК ЯКОСТІ ОБСЛУГОВУВАННЯ В МУЛЬТИСЕРВІСНІЙ IP-МЕРЕЖІ

Як зазначалося, сучасна тенденція до здійснення конвергенції різних типів мереж призвела до появи мультисервісних мереж NGN, основу яких становить універсальне транспортне середовище на базі IP технологій. Універсальність такої мультисервісної IP мережі передбачає необхідність забезпечити передачу трафіку різного виду і тут особливого значення набувають механізми підтримки потрібних характеристик QoS, особливо, коли цей трафік передається по мережі одночасно, наприклад, голосовий трафік і трафік Web-додатків. Це пов'язано з тим, що різні типи трафіку пред'являють різні вимоги до характеристик якості обслуговування. У зв'язку з тим, що постійно збільшується обсяг мультимедійної інформації, що передається по каналам сучасних мультисервісних IP-мереж, зростають також і вимоги до QoS трафіку, що генерується користувачами, абонентськими системами та безпосередньо вузлами мережі [18].

Найважливішими характеристиками в тестах QoS серед тих, що ми розглядали у попередніх розділах, є такі параметри [18]:

- кругова затримка (Round-Trip Delay time, RTD);
- варіація затримки пакетів (джиттер);
- втрата пакетів.

Для здійснення розрахунку часової затримки у мережі для подальшої оцінки її значення потрібно брати до уваги такі параметри як довжина середовища передачі та відповідні втрати, що воно вносить. На цей час у разі організації сучасних мереж будь-якого рівня найбільш оптимальним середовищем є використання волоконно-оптичного кабелю (ВОК). Тому в нашому випадку вироблення загальної методики оцінки мережних характеристик якості обслуговування будемо враховувати довжину ВОК, його тип, втрати у волокні, дисперсійні значення, відношення оптичного

сигнал/шум, частота та рівень каналу, тип та параметри обладнання, у тому числі транспондерів (приймачів), кількість мережних елементів та точок регенерації [18].

Основне значення часової затримки вноситься довжиною оптичного кабелю. Затримка, що вноситься обладнанням, незначно позначається на загальній величині часової затримки каналу.

Вимірювання зазначених параметрів може проводитись для різних класів сервісу [18]:

- сервісу що функціонує у реальному часі (real-time);

- сервісу, що є критичним для бізнесу (business critical);

- сервісу, що надається у режимі негарантованої доставки (best effort).

Визначимо теоретичні значення мережних показників якості обслуговування. Вони можуть бути використані під час укладання угоди SLA із сторонніми операторами по забезпеченню потрібної QoS в з'єднанні типу «із кінця в кінець» (end-to-end) для різних видів трафіку.

Необхідно визначити такі параметри [18]:

- кругова затримка;

- затримка поширення;

- час очікування пакета у черзі в маршрутизаторі;

- затримка, що вноситься активним обладнанням;

- втрата пакетів;

- джиттер.

1) Почнемо з проведення оцінки кругової затримки (RTD), яка представляє собою сумарний час, який необхідний для передачі пакета від джерела до одержувача і назад. У загальному випадку вона складається з таких видів затримок [18]:

- затримка, що утворюється у разі поширення сигналу, D_p;

- затримка, що утворюється за час очікування пакета у черзі в маршрутизаторі, D_Q;

- затримка, яка вноситься активним обладнанням, $D_{a.e.}$.

Якщо маршрутизація є симетричною (тобто використовується один і той же маршрут від джерела до одержувача і назад), і проходить найкоротшими шляхами, тоді RTD розраховується за формулою [18]:

$$RTD = 2 \cdot (D_p + \sum D_{Qi} + \sum D_{a.e.i}). \quad (4.1)$$

2) Затримка, що утворюється при поширенні сигналу залежить від довжини маршруту та швидкості поширення світлового потоку в оптичному волокні. Виходячи з цього, затримку D_p можна оцінити за формулою [19]:

$$D_p = {R \cdot n_1}/{C}, \quad (4.2)$$

де C – швидкість світла у вакуумі, $3 \cdot 10^8$ м/с;

R – довжина маршруту;

n_1 – коефіцієнт заломлення матеріалу серцевини оптичного волокна, значення якого лежить у межах від 1,45 до 1,55 (ближче до 1,5).

У разі, якщо R = 3000 км та n_1=1,5, отримаємо:

$$D_p = {3 \cdot 10^6 \cdot 1{,}5}/{3 \cdot 10^8} = 15 \text{ мс}.$$

Зазначимо, що можлива ситуація, коли довжина маршруту R є невідомою або незаданою. Тоді її значення можна розрахувати з використанням коефіцієнтів, що оцінюються з урахуванням параметра D, який характеризує пряму відстань між мережними вузлами, відповідно із специфікаціями, що надані у рекомендації ITU G.826 (дивись таблицю 4.1) [19]:

Таблиця 4.1 – Оцінка значення *R* у відповідності із специфікаціями G.826

D	*R*
D < 1000 км	*R* = 1,5·*D*
1000 км ≤ *D* ≤1200 км	*R* = 1500 км
D > 1200 км	*R* =1,25·*D*

4) Оцінку величини затримки, що утворюється за час очікування пакета у черзі в маршрутизаторі, зробимо за такою формулою [18]:

$$D_Q = (b/r)\cdot(1/1-u), \tag{4.3}$$

де *b* – середня довжина пакета (біт);

r – швидкість передачі, що забезпечується каналом (біт/с);

u – середній коефіцієнт використання каналу.

Припустимо, що швидкість у каналі становить *r* = 100 Мбіт/с, довжина пакета – *b* = 2000 байт чи 16000 біт, середній коефіцієнт використання каналу – *u* = 0,9. Звідси отримаємо величину затримки D_Q:

$$D_Q = \left(16\cdot 10^3/10^8\right)\cdot(1/1-0{,}9) = 1{,}6 \text{ мс}.$$

При тій же швидкості у каналі (тобто 100 Мбіт/с), але при довжині пакета – *b* = 160 байт (1280 біт) та середньому коефіцієнті використання каналу – u = 0,9, величина затримки D_Q становитиме:

$$D_Q = \left(1280/10^8\right)\cdot(1/1-0{,}9) = 0{,}128 \text{ мс}.$$

Тобто можна бачити, що при зменшенні довжини пакету, відповідно зменшується і час очікування пакета в черзі маршрутизатора. Це стосується одного маршрутизатору на шляху від джерела до одержувача. У цілому ж затримки на маршрутизаторах завжди будуть мати величину меншу за1 мс, якщо канали не будуть перевантажені [18].

5) Зробимо оцінку затримки активного обладнання. У нашому випадку вона являє собою сумарну величину затримок, які внесені, наприклад, таким обладнанням оптичної магістралі : компенсаторами дисперсії, транспондерами, 3R регенераторами, іншим обладнанням. Значення затримок, які вносяться активними елементами мережі та використовуються для оцінки величини $D_{a.e.}$, приводяться у технічній документації до того чи іншого обладнання його постачальниками [19]. Зокрема загальна величина затримки $D_{a.e.}$ обраховується за формулою [19]:

$$D_{a.e.заг.} = D_{a.e.1} + D_{a.e.2} + ... + D_{a.e.n} = \sum D_{a.e.i} \quad (4.4)$$

Наприклад, для обладнання Cisco дані щодо затримки, яка вноситься модулями компенсації дисперсії, будуть мати наступні значення (таблиця 4.2) [19]:

Таблиця 4.2 – Значення затримок у оптичному волокні, що вносяться модулями компенсації дисперсії

DCM із зазначенням довжини волокна, що компенсується, км	Затримка, мс
1	2
DCM-2,5	1
DCM-5	3
DCM-7,5	5
DCM-10	7
DCM-20	15
DCM-30	22
DCM-40	30
DCM-50	38
DCM-60	45
DCM-70	53

Продовження таблиці 4.2

1	2
DCM-80	61
DCM-90	68
DCM-100	76

Потрібно звернути особливу увагу на те, що величина затримки, яка вноситься транспондером (Optical Transponder Unit, OTU), буде залежати від того, чи є транспондер одночасно і концентратором (так званим мукспондером - TRBC), а також від того, чи знаходиться сигнал користувача в OTU чи ні). Відповідно до рекомендації ITU G.709, де визначені стандартні інтерфейси та швидкості транспортних каналів, на сьогоднішній день в оптичних транспортних мережах використовуються переважно дві швидкості передачі даних: OTU1 (2,7 Гбіт/с) служить для прозорої передачі потоків технології SDH рівня STM-16 і OTU2 (10,7 Гбіт/с) забезпечує передачу трафіку рівня STM-64 і трафіку фізичного рівня глобальних мереж зі швидкістю 10 Гбіт/с. Також потрібно зазначити, що згідно G.709, OTU1 і OTU2 підключаються до каналів по інтерфейсам «користувач-мережа» (User Network Interface, UNI) та «мережа-мережа» (Network to Network Interface, NNI). Відповідно значення затримки, що вноситься, спирається на наступні дані [20]:

- TRBD UNI = 150 мс (приклад OTU2 лінійний інтерфейс, STM64/10GE клієнтський інтерфейс);

- TRBC UNI = 150 мс (приклад OTU2 лінійний інтерфейс, STM16 клієнтський інтерфейс);

- TRBD NNI = 160 мс (приклад OTU2 лінійний інтерфейс, OTU-2 клієнтський інтерфейс);

- TRBC NNI = 175 мс (приклад OTU2 лінійний інтерфейс, OTU-1 клієнтський інтерфейс).

Затримку, що виникає на транспондері, слід враховувати сумарно на всіх OTU мережі, а також від початкового інтерфейсу користувача до кінцевого інтерфейсу (тобто на прийомі, на передачі, на проміжних регенераторах 3R типу, що забезпечують ресинхронізацію сигналу і усунення фазових тремтінь) [20].

Далі проаналізуємо методику оцінки рівня втрат пакетів. Зокрема рівень втрати пакетів визначається кількістю пакетів, що скидаються мережею у процесі передачі. Найважливішими причинами втрати пакетів є перевантаження мережі та пошкодження пакетів під час передачі по лінії зв'язку. Також скидування пакетів може бути спричинене недостатнім розміром вхідного буфера. Коефіцієнт втрати пакетів визначається за такою формулою [18]:

$$K_{втрат} = \frac{Nvtr}{Nvtr + N_{otr}} \cdot 100\%. \quad (4.5)$$

де N_{vtr} - кількість втрачених пакетів;

N_{otr} – кількість пакетів, які були отримані успішно.

Зробимо припущення, що в мультисервісній IP-мережі кількість переданих пакетів становить 115288, пакетів, що були втрачені (ті, що були скинуті, або пошкоджені) – 488, а, відповідно, кількість пакетів, що були доставлені, становить 114800. Звідси за формулою (4.5) розрахуємо коефіцієнт втрати пакетів [18]:

$$K_{vtr} = 488/(488+114800) \cdot 100\% = 0{,}4\%$$

І на останок зробимо оцінку величини джиттера або варіації затримки пакетів. Цей параметр описаний у специфікації стандарту RFC 3393 і визначається як різниця наскрізних затримок проходження двох пакетів. Значення джиттера для *i*-ого та *j*-го пакетів визначимо за формулою [18]:

$$D_{i,j} = (R_j - R_i) - (S_j - S_i) = (R_j - S_j) - (R_i - S_i), \quad (4.6)$$

де R – час відправлення пакета;

S – час доставки пакета.

Таким чином, в цьому розділі надана і обґрунтована загальна методика оцінки мережних характеристик якості обслуговування в мультисервісних IP-мережах, які працюють по оптичним лініям зв’язку та є базовою основою для реалізації транспортної платформи мереж NGN.

ВИСНОВКИ

У процесі виконання кваліфікаційної роботи магістра наведені та обґрунтовані основні поняття, визначення та характеристики якості обслуговування, проаналізовані стандарти, механізми і відповідні їм моделі забезпечення потрібної QoS у мультисервісних IP-мережах.

В першому розділі роботи обґрунтовано необхідність отримання від мультисервісних IP-мереж необхідних гарантій якісної доставки чутливої до затримок інформації, такої як мова, відео та мультимедіа в реальному часі та з мінімально можливою затримкою. Для забезпечення цих цілей ITU-T розроблено стандарти, що регламентують забезпечення в таких мережах потрібної QoS. У роботі розглянуті ті з них, які описують термінологію та поняття якості обслуговування і характеристики роботи мережі (рекомендації ITU-T E.800 та I.350); стандартні мережні характеристики передачі пакетів у IP мережах (рекомендація ITU-T Y.1540); нормовані значення цих мережевих характеристик відповідно до різних класів QoS (рекомендація ITU-T Y.1541).

Крім визначення мережних характеристик та специфікації норм для них, ITU-T провів також роботи з ідентифікації та стандартизації мережних механізмів, що мають забезпечувати потрібну QoS в мультисервісних IP-мережах. Зокрема у другому розділі кваліфікаційної роботи аналізується базова модель підтримки якості обслуговування в мультисервісних IP-мережах, що описується в рекомендації Y.1291. Показано, що вона визначає набір мережних механізмів, які називаються конструктивними блоками. Ці блоки відповідають трьом логічним площинам: площині контролю, площині даних (інформаційній площині) та площині адміністративного управління.

Також було зазначено, що здатність IP-мережі забезпечувати реалізацію різних рівнів обслуговування, що потребують ті чи інші мережні додатки, поряд із здійсненням контролю за використанням мережних ресурсів та параметрами продуктивності, таких як смуга пропускання, затримка, джиттер

та втрата пакетів, – може бути реалізована із застосуванням трьох технологій обслуговування: негарантованої доставки даних, гарантованого і диференційованого обслуговування.

Останні дві технології отримали практичну реалізацію у двох найбільш широко застосовуваних моделях забезпечення QoS: моделі диференційованого обслуговування (DiffServ) та моделі інтегрованого обслуговування (IntServ). Тому у третьому розділі проаналізовані можливості і принципи досягнення необхідних вимог щодо якості обслуговування на основі застосування розглянутих мережних механізмів з використанням технологічних моделей IntServ та DiffServ.

Технологічна модель IntServ у поєднанні з сигнальним протоколом RSVP дозволяє організувати гнучке обслуговування різнотипного трафіку, максимально враховуючи потреби кожної програми, а використання алгоритму WFQ для обслуговування пакетів гарантує мінімізацію затримки. Ця особливість робить IntServ ідеальною моделлю для обслуговування мультимедійного трафіку, що є чутливим до часових затримок. Однак висока гнучкість і орієнтація перш за все на задоволення потреб одиничних потоків є недоліками IntServ. Зокрема основний недолік моделі IntServ як раз і пов'язаний із низькою масштабованістю RSVP, що особливо проявляється на високошвидкісних магістралях транспортних IP-мереж. Продуктивність IntServ залежить від кількості потоків, що обробляються, отже, таку сервісну модель практично неможливо реалізувати в мережі з дуже великою кількістю користувачів. Тому для великих мереж потрібна більш проста та масштабована технологія, а сфера застосування технології IntServ обмежується внутрішніми та кінцевими мережами.

Відповідний аналіз технологічної моделі DiffServ показав, що її перевагами є відносна простота та висока масштабованість. З цієї причини для моделі DiffServ відведено місце на магістральних та високошвидкісних ділянках мережі.

Також зроблено загальний порівняльний аналіз технологічних моделей IntServ та DiffServ, а також обґрунтовано і доведено, що оптимальним варіантом надання необхідного QoS «з кінця в кінець» є забезпечення спільної роботи цих моделей. Створення такої моделі дозволить значно підвищити якість надання мультимедійних послуг в мультисервісних IP-мережах, а також дозволить підвищити продуктивність вже давно існуючих традиційних послуг.

У практичній частині кваліфікаційної роботи було представлено та обґрунтовано загальну методику оцінки мережних характеристик якості обслуговування в мультисервісних IP-мережах. Для запропонованої методики найбільш значущими мережними характеристиками, що вимагають здійснення оцінки, були зазначені: кругова затримка (RTD), варіація затримки пакетів (джиттер) і втрата пакетів.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Маколина М. А. Разработка и исследование моделей оценки качества передачи видео в IP-сетях : дис. канд. техн. наук : дис. канд. техн. наук : 05.12.13 / Маколина Мария Александровна – Санкт-Петербург, 2014. – 187 с.

2. Гулевич Д.С. Сети связи следующего поколения: учеб. курс [Електронний ресурс] / Д. С. Гулевич // НОУ «ИНТУИТ» – 2007. – Режим доступу до ресурсу: http://www.intuit.ru/studies/courses/1150/157/info.

3. Телекоммуникационные системы и сети: Мультисервисные сети, Том 3 / В.В. Величко, Е.А. Субботин, В.П. Шувалов, А.Ф. Ярославцев. – М.: Горячая линия – Телеком, 2005. – 592 с.

4. Б.С. Гольдштейн, Пинчук А.В. Суховицкий АЛ. IP-телефония. - М.: Радио и связь, 2001. - 336 с.

5. Кудзиновская И.П. Анализ методов обеспечения качества обслуживания в высокоскоростных компьютерных сетях/ И.П. Кудзиновская // Проблеми інформатизації та управління. – №1 (23). – 2008. – С. 182 – 187.

6. Вегешна Ш. Качество обслуживания в сетях IP / Шринивас Вегешна – М.: Издательский дом "Вильяме", 2003. – 368 с.

7. Башарин Г.П. Модели для анализа качества обслуживания в сетях связи следующего поколения: учеб. пособ. / Г.П. Башарин, Ю.В. Гайдамака, К.Е. Самуйлов, Н.В. Яркина – М.: РУДН, 2008. – 137 с.

8. Нетес В.А. Качество обслуживания на сетях связи. Обзор рекомендаций ITU-T/ В.А. Нетес // Сети и системы связи. – 1999. – №3. – С. 66 – 77.

9. Битнер В.И. Мультисервисные сети связи: консп. лекций [Електронний ресурс] / В.И. Битнер // бакалаврская программа «Информационные технологии в телекоммуникациях», 2009. – Режим доступу ло ресурсу: http://gendocs.ru/v3927/лекции-мультисервисные_сети_связи?page=6.

10. Яновский Г.Г. Конвергенция в инфокоммуникациях: учеб. пособие / Г.Г. Яновский. – СПб: 2010. – 172 с.

11. Пример расчета «коэффициента готовности» для IT-системы. – 2018. – Режим доступу ло ресурсу: https://habr.com/ru/post/418769/.

12. Олифер В.Г. Компьютерные сети: 3-е изд. / В.Г. Олифер, Н.А. Олифер – СПб.: Питер, 2006. – 957 с.

13. Семенов Ю.В. Проектирование сетей связи следующего поколения / Ю.В. Семенов. – СПб.: Наука и Техника, 2005. – 240 с.

14. Баскаков И.В. IP-телефония в компьютерных сетях [Електронний ресурс] /И.В. Баскаков, А.В. Пролетарский, Р.А. Федотов, С.А. Мельников // НОУ «ИНТУИТ» – 2008. – Режим доступу до ресурсу: https://intuit.ru/studies/courses/8/8/info.

15. Грошев А.С. Аналіз механізмів забезпечення гарантованої якості обслуговування в мультисервісних IP мережах / А.С. Грошев, Ю.М. Колтун // матеріали 10-ої міжнародної науково-технічної конференції «Проблеми інформатизації». Том 2. – Черкаси –Баку – Бельсько-Бяла – Харків. – 24 - 25 листопада, 2022 р. – С. 102.

16. Татарникова Т.М., Оценка вероятностно-временных характеристик сетевых узлов с дифференциацией трафика / Т.М. Татарникова, А.В. Вольский // Информационно-управляющие системы. – №3. – 2018. – С. 54 - 60.

17. Качество обслуживания в IP-сетях [Електронний ресурс] / ITC.UA. – 2003. – Режим доступу ло ресурсу: https://itc.ua/articles/kachestvo_obsluzhivaniya_v_ip-setyah_15116/.

18. Кадацкая О.И. Методы метрологического обеспечения параметров качества NGN-сетей / О.И. Кадацкая, С.А. Сабурова // Системи обробки інформації. – 2016. – випуск 6 (143). – С. 52 – 54.

19. Комплекс навчально-методичного забезпечення з навчальної дисципліни «Напрямні системи електричного та оптичного зв'язку» (частина 2)» підготовки бакалавра напряму 6.050903 «Телекомунікації» / Розробник: Колтун Ю.М., доц. каф. ІМІ, к.т.н. – Харків: ХНУРЕ, каф. ІМІ, 2017 р. – 237 с.

20. Стив А. Оптические транспортные сети [Електронний ресурс] / Стив Александер // Computerworld. – № 21. – 2006. – Режим доступу ло ресурсу: https://www.osp.ru/cw/2006/21/2046391.