

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**ОДЕСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ**  
**ХАРЧОВИХ ТЕХНОЛОГІЙ**



**Шестопапов С.В.**

**Дослідження та проектування  
комп'ютерних систем та мереж**

**Конспект лекцій**

**Одеса - 2017**

**Шестопапов С.В.** Дослідження та проектування комп'ютерних систем та мереж: конспект лекцій/ С.В. Шестопапов // Одеська національна академія харчових технологій, 2017. – 82с.

Навчальний посібник знайомить майбутніх фахівців з питаннями дослідження та проектування комп'ютерних систем та мереж.

Навчальний посібник призначено для студентів спеціальності 8.05010203 «Спеціалізовані комп'ютерні системи».

Рецензент: Котлик С.В., к.т.н., доцент Одеської національної академії харчових технологій

Розглянуто та рекомендовано до видання на засіданні кафедри комп'ютерної інженерії.

Протокол № 5 від 6 березня 2017р.

Розглянуто та рекомендовано до видання на засіданні науково-методичної комісії з напряму підготовки 123 «Комп'ютерна інженерія».

Протокол № 3 від 6 березня 2017 р.

## ЗМІСТ

Дослідження та проектування комп'ютерних систем та мереж.....	1
Шестопалов С.В. Дослідження та проектування комп'ютерних систем та мереж: конспект лекцій/ С.В. Шестопалов // Одеська національна академія харчових технологій, 2017. – 82с.....	2
ВСТУП.....	5
1. Особливості проектування комп'ютерної мережі.....	6
1.1. Характеристики комп'ютерної мережі.....	7
1.2. Вплив на мережні характеристики програмного і апаратного складу комп'ютерних мереж.....	10
1.3. Постановка завдання аналізу характеристик комп'ютерної мережі.....	13
2. МАТЕМАТИЧНІ МОДЕЛІ ДИСКРЕТНИХ СИСТЕМ.....	15
2.1. Основні поняття.....	15
2.1.1. Система масового обслуговування.....	15
2.1.2. Мережа масового обслуговування.....	17
2.1.3. Потік заявок.....	18
2.1.4. Тривалість обслуговування заявок.....	20
2.1.5. Стратегії управління потоками заявок.....	20
2.2. Класифікація моделей масового обслуговування.....	24
2.2.1. Базові моделі.....	24
2.2.2. Мережні моделі.....	25
2.3. Параметри і характеристики СМО.....	27
2.3.1. Параметри СМО.....	27
2.3.2. Позначення СМО (символіка Кендалла).....	28
Для компактного опису систем масового обслуговування часто використовуються позначення, запропоновані Д. Кендаллом у вигляді:.....	28
A/B/N/L ,.....	28
Приклади:.....	28
2.3.3. Режими функціонування СМО.....	29
2.3.4. Характеристики СМО з однорідним потоком заявок.....	30
2.3.5. Характеристики СМО з неоднорідним потоком заявок.....	32
Можна довести, що для характеристик об'єднаного (сумарного) потоку справедливі ті ж фундаментальні співвідношення (3.13) – (3.15), що і для однорідного потоку:.....	33
$U=W+B$ ; $L=W$ ; $M=W$ ,.....	33
2.4. Параметри і характеристики МеМО.....	33
2.4.1. Параметри МеМО.....	33
2.4.2. Режими функціонування МеМО.....	34
2.4.3. Характеристики МеМО.....	34
3. АНАЛІТИЧНЕ МОДЕЛЮВАННЯ.....	37

3.1. Одноканальні СМО з однорідним потоком заявок.....	37
3.1.1. Характеристики експоненційної СМО М/М/1.....	37
3.1.2. Характеристики неекспоненціальної СМО М/Г/1.....	38
3.1.3. Характеристики неекспоненціальної СМО Г/М/1.....	38
3.1.4. Характеристики СМО загального вигляду Г/Г/1.....	39
3.2. Багатоканальні СМО з однорідним потоком заявок.....	39
3.2.1. Характеристики багатоканальної СМО М/М/К.....	39
3.3. Одноканальні СМО з неоднорідним потоком заявок.....	40
3.3.1. Характеристики та властивості ДО БП.....	41
3.3.2. Характеристики та властивості ДО ВП.....	41
3.3.3. Характеристики та властивості ДО АП.....	42
3.4. Розімкнені експоненційні МеМО з однорідним потоком заявок .....	42
3.4.1. Опис розімкнутих МеМО.....	42
3.4.2. Розрахунок коефіцієнтів передач і інтенсивностей потоків заявок у вузлах РМеМО.....	43
3.4.3. Перевірка умови відсутності перевантажень в МеМО.....	44
3.4.4. Розрахунок вузлових характеристик РМеМО.....	45
3.4.5. Розрахунок мережних характеристик РМеМО.....	46
4. Чисельне моделювання (МОДЕЛІ ВИПАДКОВИХ ПРОЦЕСІВ).....	47
4.1. Поняття марковського випадкового процесу.....	47
4.2. Марківські моделі систем масового обслуговування.....	47
4.2.1. Одноканальна СМО без накопичувача (М/М/1/0).....	48
4.2.2. Одноканальна СМО з накопичувачем обмеженою ємкості (М/М/1/г).....	51
4.2.3. Одноканальна СМО з накопичувачем необмеженої ємкості (М/М/1).....	53
4.2.4. Одноканальна СМО з неоднорідним потоком заявок і відносними пріоритетами.....	55
4.3. Марківські моделі мереж масового обслуговування.....	58
4.3.1. Розімкнена експоненційна МеМО з накопичувачами обмеженої ємкості.....	58
5. Основи імітаційного моделювання.....	63
5.1. Поняття імітаційного моделювання.....	63
5.1.1. Принципи організації імітаційного моделювання.....	64
6. Синтез структури комп'ютерної мережі для забезпечення її доступності.....	69
6.1. Постановка задачі синтезу структури технічної системи із забезпеченням доступності.....	71
Література.....	75
ДОСЛІДЖЕННЯ ТА ПРОЕКТУВАННЯ КОМП'ЮТЕРНИХ СИСТЕМ ТА МЕРЕЖ	76

## ВСТУП

Апаратно-програмні засоби обчислювальної техніки, об'єднані в комп'ютерні мережі, оновлюються з великою швидкістю і потребують оновлення проектів мережі в короткі терміни і з меншими витратами.

Найважливішою характеристикою мережі є одностороння затримка пакетів (One-Way Delay Metric, OWD) що входить до числа стандартів IPPM і описана в RFC 2679 ([rfc2679.openrfc.org](http://rfc2679.openrfc.org)). Одиничне значення цієї метрики описується як час передачі пакета певного типу між деякими двома вузлами мережі.

Для оцінки продуктивності проектованої мережі потрібен певний механізм прогнозувань. Стохастичний характер надходження даних і недетермінована обробка в вузлах комутації та каналах зв'язку зумовлює використання моделей теорії масового обслуговування для проектування та аналізу проектних характеристик комп'ютерної мережі. Теорія масового обслуговування забезпечує можливість розрахунку характеристик якості функціонування мереж, включаючи оцінку ймовірно-тимчасових характеристик вузлів комутації.

# 1. ОСОБЛИВОСТІ ПРОЕКТУВАННЯ КОМП'ЮТЕРНОЇ МЕРЕЖІ

Процес проектування об'єкта – комп'ютерної мережі носить ітераційний характер. Ітерації можуть включати в себе більш ніж один рівень проектування. Тобто, в процесі проектування доводиться багато разів виконувати процедуру аналізу об'єкта. Тому очевидне прагнення зменшити трудомісткість кожного варіанту аналізу без шкоди для якості остаточного проекту. У цих умовах доцільно на початкових стадіях проектування, коли високої точності результатів не потрібно, використовувати найбільш прості і економічних моделі.

Створити проект мережі означає вибрати структуру мережі, визначити значення всіх її параметрів і представити результати у встановленій формі. Результати (проектна документація) можуть бути виражені у вигляді схем, пояснювальних записок, програм та інших документів на папері або на машинних носіях інформації.

Розробка (або вибір) структури мережі є проектна процедура структурного синтезу, а розрахунок (або вибір) значень параметрів елементів – процедура параметричного синтезу.

Завдання структурного синтезу полягає у визначенні мети, безлічі можливих рішень і обмежуючих умов. Завдання полягає в синтезі (або корекції) структури, визначенні типів серверів (програмно-апаратних засобів), розподілі функцій по серверам таким чином, щоб досягався екстремум цільової функції при виконанні заданих обмежень.

Наступна після синтезу група проектних процедур – процедури аналізу. Мета аналізу – отримання інформації про характер функціонування та значень вихідних параметрів при заданій структурі об'єкта, відомостях про зовнішні параметри і параметри елементів.

Для аналізу комп'ютерної мережі широко використовуються математичні методи і моделі масового обслуговування. Спрощені моделі масового обслуговування дозволяють знаходити явний вигляд цільової функції, в якості якої використовується характеристика мережі, така як час затримки повідомлень (пакетів).

Моделювання життєвого циклу комп'ютерної мережі проводиться за допомогою ітеративної моделі – це виконання робіт паралельно з безперервним аналізом отриманих результатів і коригуванням попередніх етапів роботи. Комп'ютерна мережа при цьому підході в кожній стадії проходить повторюваний цикл: планування – реалізація – перевірка – оцінка (англ. *plan-do-check-act cycle*).

В автоматизованих проектних процедурах замість проектного не існуючого об'єкта оперують моделлю, яка відображає деякі властивості об'єкта, які цікавлять дослідника.

Математичні функціональні моделі в загальному випадку являють собою алгоритм обчислення вектора вихідних параметрів при заданих векторах параметрів елементів і зовнішніх параметрів.

Закон функціонування мережі може бути представлений в наступному вигляді:

$$H(t) = f_c(S, F, Y, X, t), \quad (1.1)$$

де  $f_c$  – функція, функціонал, логічні умови, алгоритм, методика, таблиця або словесний опис, визначальне правило (закон) перетворення вхідних величин (параметрів) у вихідні величини (характеристики);

$H(t)$  – вектор характеристик, що залежить від поточного моменту часу  $t$  ( $t > 0$ ):

$$H = \{V, T, N, C, Z\} \quad (1.2)$$

Параметри – первинні дані мережі:

$S$  – структурні,

$F$  – функціональні,

$Y$  – навантажувальні,

$X$  – зовнішнього середовища.

Характеристики – вторинні дані мережі:

$V$  – потужнісні,

$T$  – часові,

$N$  – надійнісні,

$C$  – економічні,

$Z$  – інші.

### 1.1. Характеристики комп'ютерної мережі

Поняття характеристики функціонування мережі в роботі включає вторинні властивості комп'ютерної мережі, які визначаються в процесі розв'язання завдань аналізу, як функція параметрів. Параметри комп'ютерної мережі описують первинні властивості мережі і є вихідними даними при вирішенні завдань аналізу.

Характеристики комп'ютерних мереж – це сукупність показників ефективності (якість) мережі. Характеристики комп'ютерних мереж можна розділити на якісні та кількісні.

Кількісні характеристики комп'ютерних мереж можна розділити на дві групи:

– глобальні, що визначають найбільш важливі властивості мережі як цілісного об'єкта;

– локальні, що визначають властивості окремих пристроїв або частин мережі та дозволяють отримати більш детальне уявлення про ефективність мережі.

До глобальних відносяться характеристики продуктивності, оперативності, надійності, вартісні, інші (енергоспоживання, масогабаритні і т.п.).

Продуктивність комп'ютерної мережі – міра потужності мережі, що визначається кількістю роботи, яка виконується мережею в одиницю часу. Поняття продуктивності охоплює широку номенклатуру показників ефективності комп'ютерної мережі, що визначають якість функціонування як мережі в цілому, так і окремих її підсистем і елементів – технічних і програмних засобів.

Продуктивність мережі залежить, в першу чергу, від продуктивності окремих її елементів, що називається швидкістю роботи або швидкодією пристроїв, наприклад, швидкість передачі даних по каналах зв'язку, вимірювана обсягом даних, переданих за одиницю часу, швидкодія ПК або, точніше, процесора, вимірюється числом команд, виконуваних в одиницю часу, і т.п. Для оцінки продуктивності комп'ютерної мережі в цілому використовується наступна сукупність показників:

- продуктивність МПД (мережі передачі даних), яка вимірюється числом повідомлень (пакетів, кадрів, біт) переданих по мережі за одиницю часу;
- продуктивність ЗОД (засобів обробки даних), що представляє собою сумарну продуктивність всіх засобів обчислювальної техніки (ПК і систем), що входять до складу мережі.

Продуктивність МПД (комунікаційна потужність) може бути задана наступними показниками:

- максимальна або гранична продуктивність, що називається пропускною здатністю мережі передачі даних і вимірюється кількістю пакетів (кадрів), переданих в мережі за одиницю часу;
- реальна чи фактична продуктивність мережі передачі даних, яка може бути задана як середнє значення на деякому інтервалі часу або як миттєве значення в конкретний момент часу.

Продуктивність ЗОД (обчислювальна потужність) в цілому складається з продуктивності обчислювальних засобів (ОЗ), що виконують обробку даних в мережі. Найбільш важливим показником продуктивності ОЗ, як сукупності технічних і програмних засобів, є системна продуктивність  $\lambda_0$ , вимірювана числом завдань, виконуваних системою за одиницю часу.

Характеристики оперативності описують затримки, що виникають при передачі та обробці даних в мережі.

Для оцінки оперативності мережі в цілому використовуються наступні показники: час доставки пакетів (повідомлень), час відгуку (відповіді).

Час доставки (час затримки) пакетів характеризує ефективність організації передачі даних в обчислювальній мережі і являє собою інтервал часу, вимірюваний від моменту надходження пакету або повідомлення в мережу до моменту отримання пакета адресатом. У загальному випадку, час затримки – величина випадкова, що обумовлено випадковим характером процесів надходження і передачі даних в мережі.



У комп'ютерних мережах звичайно час доставки задається середнім значенням  $T$ , на котре може накладатися обмеження  $T < T^*$  залежно від типу переданих даних.

При передачі мультимедійних даних крім середнього значення часу доставки пакетів важливою характеристикою є варіація або джиттер затримки, що представляє собою середньоквадратичне відхилення часу затримки різних пакетів.

Час відгуку (відповіді) – інтервал часу від моменту надходження запиту (транзакції) в мережу до моменту завершення його обслуговування, пов'язаного з виконанням деякої прикладної або обслуговуючої програми, із зверненням до бази даних і т.п.

Час відповіді являє собою час перебування запиту в мережі і характеризує ефективність як телекомунікаційних, так і обчислювальних засобів комп'ютерної мережі.

Час відгуку, як і час затримки, – величина випадкова і може задаватися середнім значенням  $U$  або у вигляді ймовірності  $P(t_u < U^*)$  неперевикнення деякого заданого значення  $U^*$ .

У мережах реального часу замість терміна "час відповіді" часто використовують термін "час реакції".

В якості характеристик надійності зазвичай використовуються наступні показники:

- ймовірність безвідмовної роботи мережі  $P(t)$  – ймовірність того, що протягом часу  $t$  не відбудеться відмови;

- інтенсивність відмов  $X_a$  – середнє число відмов за одиницю часу;

- час напрацювання на відмову – проміжок часу між двома суміжними відмовами – величина випадкова, а її середнє значення  $T_0$  називається середнім напрацюванням на відмову  $T_0 = 1/\lambda_0$ ;

- час відновлення – інтервал часу від моменту настання відмови до моменту відновлення працездатності системи – величина випадкова і зазвичай задається середнім значенням так званим середнім часом відновлення;

- коефіцієнт готовності  $K_z$  – частка часу, протягом якого мережа працездатна:  $K_z = T_0 / (T_0 + T_\theta)$ .

Величина  $K_z$  може трактуватися як ймовірність того, що в будь-який момент часу мережа працездатна.

Аналогічно, значення  $(1 - K_z)$  визначає ймовірність того, що мережа знаходиться в стані відновлення (непрацездатна).

В якості вартісних (економічних) характеристик комп'ютерної мережі можуть використовуватися такі показники:

- повна вартість володіння (*Total cost of ownership*) – витрати, що розраховуються на всіх етапах життєвого циклу мережі і включають вартість технічних, інформаційних і програмних засобів (прямі витрати) і витрати на експлуатацію мережі (непрямі затрати);

– вартість (ціна) передачі даних та обробки даних у мережі, обумовлена обсягом і вартістю використовуваних ресурсів мережі відповідно при передачі та обробці даних.

В якості локальних характеристик комп'ютерних мереж можуть використовуватися в залежності від цілей дослідження найрізноманітніші показники ефективності.

Локальні характеристики описують ефективність функціонування:

- вузлів і каналів зв'язку;
- окремих сегментів мережі;
- вузлів обробки даних: обчислювальної системи та її підсистем.

Локальні характеристики можуть бути розбиті на дві групи:

- часові;
- безрозмірні.

До часових характеристик відносяться:

- час доставки (затримки) пакетів при передачі між сусідніми вузлами мережі;
- час очікування передачі даних у вузлах мережі або звільнення ресурсів (сервера);
- час перебування даних в різних вузлах, пристроях або підсистемах.

До безрозмірних характеристик відносяться:

- число пакетів, що знаходяться в буферній пам'яті вузлів (маршрутизаторів, комутаторів);
- коефіцієнти завантажень вузлів, каналів зв'язку й пристроїв і т.д.

Коефіцієнт завантаження або просто завантаження пристрою  $\rho$  це - частка часу, протягом якого пристрій працює:

$$(1.3)$$

,де  $t$  - час, протягом якого пристрій працював;  $T$  - час спостереження.

Завантаження  $\rho$  характеризує ступінь використання пристрою і часто називається коефіцієнтом використання пристрою. Оскільки  $0 < \rho < 1$ , то завантаження може трактуватися як ймовірність того, що в будь-який момент часу пристрій працює. Величина  $\eta = 1 - \rho$  називається коефіцієнтом простою пристрою і характеризує частку часу, протягом якого пристрій не працює (простоює).

## **1.2. Вплив на мержні характеристики програмного і апаратного складу комп'ютерних мереж**

Сучасні досягнення виробників програмного і апаратного складу засобів обчислювальної техніки та мержного обладнання суттєво впливають на можливості зміни структур комп'ютерної мережі (КМ).

Виробники комутаційного обладнання створюють нові технології, що істотно змінюють можливості апаратно-програмних засобів побудови комп'ютерної мережі і тим самим значно збільшують кількість і характер

можливих для реалізації структур комп'ютерної мережі. Можно виділити компоненти, котрі мають найбільший вплив на мережні характеристики:

- розробки фірм-виробників апаратної частини комутаційного обладнання (керовані і некеровані switches L2, L2 +, L3, L3 + рівнів);
- розробки засобів віртуалізації фірм-виробників прикладного програмного забезпечення (*VMWare, HyperVisor*);
- розробки програмного забезпечення фірм-виробників комутаційного обладнання (реалізації технологій *VLAN, Link Aggregation, Spanning Tree, RSTP, MSTP, Loop Guard, ARP Spoofing, DHCP Snooping, VRRP, ...*).

Зміна структури КМ, в свою чергу, неоднозначно впливає на мережні характеристики. Набір апаратних засобів і ресурсів комп'ютерної мережі, відповідних першому, другому, третьому рівням семирівневої моделі міжмережної взаємодії *ISO/OSI*, включає в себе мережні інтерфейси, структуровану кабельну систему, різне комутаційне обладнання, що дозволяє направляти потоки даних між користувачами. З наявного набору засобів є можливість створити певну кількість структур комп'ютерної мережі:  $S_1, S_2, S_3, \dots, S_k$ , де:

$S_k$  – структура з  $k$  кількістю вузлів і конкретною фізичною топологією.

Структури  $S_j$  розрізняються кількістю вузлів і способом взаємодії між ними, утворюючи певну фізичну топологію  $S_i, i=1, k$ .

Фірми-виробники комутаційного обладнання пропонують рішення на базі апаратно-програмних можливостей своїх продуктів, які можна назвати типовими структурами комп'ютерної мережі:

$$S_i, i=(n+1,k); n < k. \quad (1.4)$$

Крім типових фірмових структур, можуть бути запропоновані структури на основі аналітичного та імітаційного моделювання та методики аналізу характеристик мережі цілеспрямованою зміною структури мережі, запропонованої в даній роботі:

$$S_i, i=(n+1,k); n < k. \quad (1.5)$$

Виробники програмного забезпечення комутаційного обладнання пропонують кошти з використанням технології віртуальних локальних мереж *VLAN*, стандартів функцій і протоколів фірм-виробників, які також при налаштуванні помітно змінюють потоки в комп'ютерній мережі і характеристики мережі. Зокрема додавання однієї або декількох *VLAN* кардинально перерозподілять потоки даних у мережі з однією і тією ж фізичною топологією. З'являється необхідність додаткових досліджень конкретного варіанту мережі з раціонального використання програмного забезпечення комутаційного обладнання.

Функції, які можуть бути налаштовані на комутаційному обладнанні або в мережі пропонуються фірмами-виробниками для підвищення

продуктивності і надійності комп'ютерної мережі. До них відносяться: технології віртуальних мереж *VLAN*, підтримка протоколу *Spanning Tree*, об'єднання (агрегування) портів, підтримка *SNMP*-управління, *Port Security* або прив'язка *MAC*-адреси до визначеного порту, підтримка 802.1x, протоколи групового мовлення, управління потоком, обмеження смуги пропускання, налаштування *IP*-доменів на комутаторах, використання класифікатора і політики, застосування *OSPF* і т.д. Технології та функції, перераховані вище, збільшують число і характер структур комп'ютерної мережі  $S_i$ :

$$S_i, i=(k,f); k < f. \quad (1.6)$$

Виробники прикладного програмного забезпечення пропонують засоби віртуалізації, такі як додатки *VMWare*, створені на її основі віртуальні (гостьові) робочі станції і сервери, віртуальні мережні сервери і свічі. Віртуальні (гостьові) операційні системи, віртуальні сервери і свічі мають можливість взаємодіяти між собою і з материнськими (хостовими) операційними системами і фізичними комп'ютерними мережами за допомогою вбудованих засобів комутації та маршрутизації.

Використання різних засобів прикладної віртуалізації значно змінює кількість і характер можливих для реалізації структур комп'ютерної мережі  $S_i$ .

Таким чином, при додатковому налаштуванні або адмініструванні однієї і тієї ж структури  $S_k$  комп'ютерної мережі з однаковими постійними вкладеннями можна отримати кілька варіантів  $S_k$  структур, з іншими характеристиками мережі, одна з яких є найкращою за заданим критерієм  $S_k$ .

Зміна структурно-функціональної організації комп'ютерної мережі призводить до поліпшення одних показників ефективності функціонування мережі і до погіршення інших показників мережі, що істотно ускладнює вибір найкращого варіанту структурно-функціональної організації проектованої комп'ютерної мережі, тому що показники ефективності мережі є суперечливими.

При некоректному застосуванні технології *VLAN*, використанні різних засобів прикладної віртуалізації, додаткового налаштування або адмініструванні однієї і тієї ж структури  $S_k$  комп'ютерної мережі з однаковими постійними вкладеннями можна погіршити працездатність мережі. З'являється необхідність додаткових досліджень конкретного варіанту мережі з раціонального використання програмного забезпечення комутаційного обладнання.

Отже, виникає завдання зміни структури мережі, підвищення показників продуктивності та оперативності, не погіршуючи надійності мережі.

Вкладення дають можливість урізноманітнити структури фізично і логічно. Наприклад, функція вкладень створює безліч структур  $S_i$ :

$$f(c) = \{S_1^{\text{Фізична}}, \dots, S_1^{\text{Логічна}} \dots\}. \quad (1.7)$$

Нові структури дадуть деякі характеристики:

$$H = \{S_1^{\text{Фізична}}, \dots, S_1^{\text{Логічна}} \dots\}. \quad (1.8)$$

Потрібно знайти раціональний шлях аналізу характеристик функціонування структури мережі, при обмеженні – забезпеченні надійності не нижче необхідної.

### **1.3. Постановка завдання аналізу характеристик комп'ютерної мережі**

Бурхливе зростання складності комплектуючих комп'ютерних мереж вимагає підвищення ефективності їх застосування та вдосконалення методів управління і планування мереж.

Теорія масового обслуговування (ТМО) забезпечує можливість розрахунку кількісних характеристик функціонування мереж, включаючи оцінку ймовірностно-тимчасових характеристик вузлів комутації, але не дозволяє розрахувати надійність мережі. Спрощений підхід ТМО вимагає подальшого уточнення характеристик мережі за допомогою більш реальних моделей, що призводить до ітераційної процедури проектування комп'ютерних мереж.

В даний час технічні та програмні засоби комп'ютерної мережі та умови їх роботи стають все більш складними. Кількість елементів в окремих видах пристроїв комп'ютерної мережі обчислюється сотнями тисяч. Імовірність виникнення хоча б однієї відмови сучасного складного пристрою стає досить великою, отже, необхідні спеціальні заходи, що забезпечують доступність ресурсів і послуг. Доступність – такий стан комп'ютерної мережі, при якому мережа відповідає всім вимогам, що пред'являються до її функціонування, тобто це властивість зберігати працездатний стан протягом деякого напрацювання. Для оцінки доступності застосовується кількісний показник: коефіцієнт оперативної готовності  $R_s$ .

Методи теорії надійності використовуються для забезпечення заданих вимог до функціонування комп'ютерної мережі та підвищення доступності на етапах проектування, виготовлення, випробування та експлуатації мережі. Методи теорії ймовірності, як математичного апарату теорії надійності, використовуються для розрахунку показника ймовірності безвідмовної роботи комп'ютерної мережі.

При проектуванні мережі, зміні її структури з метою поліпшення характеристик або оновлення, з'являється небезпека зниження готовності, тому необхідно розраховувати коефіцієнт готовності проекту мережі на

кожному кроці ітерації проектування. Використання запропонованої методики аналізу проекту КМ призводить до цілеспрямованої зміни структури мережі і обчисленню, в тому числі і надійнісних характеристик функціонування.

Характеристики функціонування мережі в цьому випадку пропонується представити як функцію структурно-функціональної організації та коефіцієнта оперативної готовності, не нижче заданого.

$$H(t) = f(SF, R_s \geq R_z, C), \quad (1.9)$$

де  $SF$  – структурно-функціональні параметри комп'ютерної мережі;  
 $R_s, R_z$  – коефіцієнти оперативної готовності;  
умова  $R_s \geq R_z$ , обмежує ймовірність безвідмовної роботи мережі;  
 $C$  – вкладення в комп'ютерну мережу.

## 2. МАТЕМАТИЧНІ МОДЕЛІ ДИСКРЕТНИХ СИСТЕМ

Дослідження складних систем передбачає побудову абстрактних математичних моделей, представлених на мові математичних відносин в термінах певної математичної теорії, що дозволяє отримати функціональні залежності характеристик досліджуваної системи від параметрів. Вивчення процесів, що протікають в дискретних системах зі стохастичним характером функціонування, проводиться в рамках *теорії масового обслуговування (ТМО) та теорії випадкових процесів*. При цьому багато моделей реальних систем будуються на основі **моделей масового обслуговування (ММО)**, які діляться на **базові моделі** у вигляді *систем масового обслуговування* і **мережні моделі** у вигляді *мереж масового обслуговування*, що представляють собою математичні об'єкти, описувані в термінах відповідного математичного апарату.

### 2.1. Основні поняття

Для опису одного і того ж поняття численні літературні джерела з моделями і методами теорії масового обслуговування часто використовують різні терміни. Сама «теорія масового обслуговування» часто називається «теорією черг» (в англійській літературі *Queue Theorie*), поряд з терміном «обслуговуючий прилад» використовуються терміни «пристрій», «канал», «лінія» і т.д. Зазвичай це пов'язано з прикладною областю, в якій застосовуються моделі масового обслуговування. Наприклад, терміни «виклик» і «лінія» використовуються в телефонії (звідки власне і пішла теорія масового обслуговування), термін «клієнт» – в моделях магазинів, банків, перукарень і т.д. Розглядаючи моделі масового обслуговування як абстрактні математичні моделі, нижче вводяться і використовуються терміни безвідносно прикладної області застосування цих моделей. Для кожного терміна в круглих дужках перераховані терміни-синоніми, які можуть зустрітися в інших джерелах.

#### 2.1.1. Система масового обслуговування

**Система масового обслуговування (СМО)** – математичний (абстрактний) об'єкт, що містить один або кілька *приладів П (каналів)*, обслуговуючих *заявки З*, які поступають в систему, и *накопичувач Н*, в якому знаходяться заявки, що утворюють *чергу Ч* і які очікують обслуговування (рис.2.1).

**Заявка (вимога, запит, виклик, клієнт)** – об'єкт, що надходить в СМО і потребує обслуговування в обслуговуючому приладі.

Сукупність заявок, розподілених у часі, утворюють **потік заявок**.

**Обслуговуючий прилад** або просто **прилад (пристрій, канал, лінія)** – елемент СМО, функцією якого є обслуговування заявок. У кожен момент часу в приладі на обслуговуванні може знаходитися тільки одна заявка.

**Обслуговування** – затримка заявки на деякий час в обслуговуючому приладі.

**Тривалість обслуговування** – час затримки (обслуговування) заявки в приладі.

**Накопичувач (буфер)** – сукупність місць для очікування заявок перед обслуговуючим приладом. Кількість місць для очікування визначає **ємкість накопичувача**.

Заявка, що надійшла на вхід СМО, може знаходитися в двох станах:

- в стані *обслуговування* (в приладі);
- в стані *очікування* (у накопичувачі), якщо всі прилади зайняті обслуговуванням інших заявок.

Заявки, що знаходяться в накопичувачі і очікують обслуговування, утворюють **чергу** заявок. Кількість заявок, що очікують обслуговування в накопичувачі, визначає **довжину черги**.

**Дисципліна буферизації** – правило занесення в накопичувач (буфер) заявок, що надходять.

**Дисципліна обслуговування** – правило вибору заявок з черги для обслуговування в приладі.

**Пріоритет** – переважне право на занесення (в накопичувач) або вибір з черги (для обслуговування в приладі) заявок одного класу по відношенню до заявок інших класів.

Таким чином, СМО включає в себе:

- *заявки*, що проходять через систему і утворюють *потоки заявок*;
- *черги заявок*, що утворюються в накопичувачах;
- обслуговуючі *прилади*.

Існує велике різноманіття СМО, що розрізняються структурною та функціональною організацією. Водночас, розробка аналітичних методів розрахунку характеристик функціонування СМО в багатьох випадках передбачає наявність ряду припущень, що обмежують безліч досліджуваних СМО.

Нижче при розгляді СМО, якщо не обумовлено інше, будемо використовувати наступні **припущення**:

- заявка, що надійшла в систему, миттєво потрапляє на обслуговування, якщо прилад вільний;
- в приладі на обслуговуванні в кожен момент часу може знаходитися тільки *одна* заявка;
- після завершення обслуговування будь якої заявки в приладі чергова заявка вибирається на обслуговування з черги миттєво, тобто, іншими словами, пристрій не простоює, якщо в черзі є хоча б одна заявка;
- надходження заявок в СМО і тривалості їх обслуговування не залежать від того, скільки заявок вже знаходиться в системі, або від якихось інших факторів;



- тривалість обслуговування заявок не залежить від швидкості (інтенсивності) надходження заявок в систему.

### 2.1.2. Мережа масового обслуговування

**Мережа масового обслуговування (MeMO)** – сукупність взаємопов'язаних СМО, в середовищі яких циркулюють заявки (рис.2.2, а).

Основними елементами MeMO є вузли (В) та джерела заявок (Д).

**Вузол** мережі являє собою систему масового обслуговування.

**Джерело** – генератор заявок, що надходять в мережу і вимагають певних етапів обслуговування у вузлах мережі.

Для спрощеного зображення MeMO використовується граф MeMO.

**Граф MeMO** – орієнтований граф, вершини якого відповідають вузлам MeMO, а дуги відображають переходи заявок між вузлами (рис. 2.2, б).

Переходи заявок між вузлами MeMO, в загальному випадку, можуть бути задані у вигляді ймовірностей передач.

Шлях руху заявок в MeMO називається **маршрутом**.

### 2.1.3. Потік заявок

Сукупність подій розподілених у часі називається **поток**. Якщо подія полягає в появі заявок, маємо *потік заявок*.

Для опису потоку заявок, в загальному випадку, необхідно задати інтервали часу  $\tau_k = t_k - t_{k-1}$  між сусідніми моментами  $t_{k-1}$  і  $t_k$  надходження заявок з порядковими номерами  $(k - 1)$  і  $k$  відповідно ( $k = 1, 2, \dots$ ;  $t_0 = 0$  – початковий момент часу).

Основною характеристикою потоку заявок є його **інтенсивність**  $\lambda$  – середнє число заявок, що проходять через деяку границю за одиницю часу. Величина  $a = 1/\lambda$  визначає *середній інтервал часу між двома послідовними заявками*.

Потік, в якому інтервали часу  $\tau_k$  між сусідніми заявками приймають певні заздалегідь відомі значення, називається детермінованим. Якщо при цьому інтервали однакові ( $\tau_k = \tau$  для всіх  $k = 1, 2, \dots$ ), то потік називається регулярним. Для повного опису регулярного потоку заявок досить задати інтенсивність потоку  $\lambda$  або значення інтервалу  $\tau = 1/\lambda$ .

Потік, в якому інтервали часу  $\tau_k$  між сусідніми заявками являють собою випадкові величини, називається випадковим. Для повного опису випадкового потоку заявок, в загальному випадку, необхідно задати закони розподілів  $A_k(\tau_k)$  всіх інтервалів  $\tau_k$  ( $k = 1, 2, \dots$ ).

Випадковий потік, в якому всі інтервали  $\tau_1, \tau_2, \dots$  між заявками незалежні в сукупності і описуються функціями розподілів  $A_1(\tau_1), A_2(\tau_2), \dots$ , називається потоком з обмеженим післядією.

Випадковий потік, в якому всі інтервали  $\tau_1, \tau_2, \dots$  розподілені по одному і тому ж закону  $A(\tau)$ , називається рекурентним.

Потік заявок називається стаціонарним, якщо інтенсивність  $\lambda$  і закон розподілу  $A(\tau)$  інтервалів між послідовними заявками не змінюються з часом. В іншому випадку потік заявок є нестаціонарним.

Потік заявок називається ординарним, якщо в кожен момент часу  $t_k$  може з'явитися тільки одна заявка. Якщо в будь-який момент часу може з'явитися більше однієї заявки, то маємо неординарний або груповий потік заявок.

Потік заявок називається потоком без післядії, якщо заявки надходять незалежно один від одного, тобто момент надходження чергової заявки не залежить від того, коли і скільки заявок надійшло до цього моменту.

*Стаціонарний ординарний потік без післядії називається найпростішим.*

Інтервали часу  $\tau$  між заявками в найпростішому потоці розподілені за експоненціальним законом з функцією розподілу

$$A(\tau) = 1 - e^{-\lambda\tau}, \quad (2.1)$$

де  $\lambda > 0$  – параметр розподілу, що представляє собою інтенсивність потоку заявок.

Найпростіший потік часто називають пуасонівським, оскільки число заявок  $k$ , що надходять за деякий заданий проміжок часу  $t$ , розподілено за законом Пуасона:

$$P(k, t) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}, \quad (2.2)$$

де  $P(k, t)$  – ймовірність надходження рівно  $k$  заявок за деякий фіксований інтервал часу  $t$ ;  $\lambda$  – інтенсивність потоку заявок. Тут  $k$  – дискретна випадкова величина, що приймає цілочисельні значення:  $k = 0, 1, 2, \dots$ , а  $t > 0$  и  $\lambda > 0$  – параметри закону Пуасона.

Слід зазначити, що пуасонівський потік, на відміну від найпростішого, може бути:

- *стаціонарним, якщо інтенсивність  $\lambda$  не змінюється з часом;*
- *нестационарним, якщо інтенсивність потоку залежить від часу:  $\lambda = \lambda(t)$ .*

Водночас, найпростіший потік, за визначенням, завжди є стаціонарним.

Аналітичні дослідження моделей масового обслуговування часто проводяться з припущенням про найпростіший потоці заявок, що обумовлено рядом властивих йому чудових особливостей.

**1. Сумовування (об'єднання) потоків.** Сума  $N$  незалежних стаціонарних ординарних потоків з інтенсивностями  $\lambda_1, \dots, \lambda_N$  утворює найпростіший потік з інтенсивністю:

$$\Lambda = \lambda_1 + \dots + \lambda_N \quad (2.3)$$

за умови, що потоки, що складаються, роблять більш-менш однаково малий вплив на сумарний потік. На практиці сумарний потік близький до найпростішого при  $N \geq 5$ .

**2. Імовірнісне розрідження потоку.** Імовірнісне (але не детерміноване) розрідження найпростішого потоку заявок, при якому будь-яка заявка випадковим чином з деякою вірогідністю  $p$  виключається з потоку незалежно від того, виключені інші заявки чи ні, призводить до утворення найпростішого потоку з інтенсивністю  $\lambda' = p\lambda$ , де  $\lambda$  – інтенсивність вихідного потоку. Потік виключених заявок – теж найпростіший з інтенсивністю  $\lambda'' = (1 - p)\lambda$ .

**3. Простота.** Припущення про найпростіший потоці заявок дозволяє для багатьох математичних моделей порівняно легко отримати в явному вигляді залежності характеристик від параметрів. Найбільше число аналітичних результатів отримано для найпростішого потоку заявок. Аналіз моделей з потоками заявок, відмінними від найпростіших, зазвичай ускладнює математичні викладки і не завжди дозволяє отримати

аналітичне рішення в явному вигляді. Свою назву "найпростіший" потік отримав саме завдяки цій особливості.

#### **2.1.4. Тривалість обслуговування заявок**

*Тривалість обслуговування заявок* – час перебування заявки в приладі – у загальному випадку величина випадкова і описується функцією  $V(\tau)$  або щільністю  $b(\tau) = V'(\tau)$  розподілу. У разі неоднорідного навантаження тривалості обслуговування заявок різних класів можуть відрізнятися законами розподілів або тільки середніми значеннями. При цьому зазвичай передбачається незалежність тривалостей обслуговування заявок кожного класу.

Часто тривалість обслуговування заявок передбачається розподіленою за експоненційним законом, що істотно спрощує аналітичні викладки. Це обумовлено тим, що процеси, що протікають в системах з експоненційним розподілом інтервалів часу, є марківськими. Величина, зворотна середньої тривалості обслуговування  $b$ , характеризує середнє число заявок, яке може бути обслуговано за одиницю часу, і називається інтенсивністю обслуговування:  $\mu = 1/b$ .

У багатьох випадках аналітичні залежності можуть бути отримані для довільного закону розподілу тривалості обслуговування заявок. При цьому для визначення середніх значень характеристик обслуговування, найчастіше, як буде показано нижче, досить задати, крім математичного очікування  $b$ , другий момент розподілу (дисперсію) або коефіцієнт варіації  $v_b$  тривалості обслуговування.

Час  $T_0$ , який залишився до завершення обслуговування заявки, що знаходиться в приладі, від моменту надходження деякої заявки в систему, і враховує, що на момент надходження в системі може і не виявитися заявок, тобто враховує простої системи, називається часом дообслуговування. Математичне очікування цього часу:

$$M[T_0] = \lambda b^2(1 + v_b) / 2, \quad (2.4)$$

де  $\lambda$  – інтенсивність найпростішого потоку заявок, що надходять в систему.

#### **2.1.5. Стратегії управління потоками заявок**

Стратегія управління потоками заявок в моделях масового обслуговування задається у вигляді:

- *дисципліни буферизації (ДБ);*

- *дисципліни обслуговування (ДО).*

ДБ і ДО можуть бути класифіковані за такими ознаками:

- наявність пріоритетів між заявками різних класів;

- спосіб (режим) витіснення заявок з черги (для ДБ) і призначення заявок на обслуговування (для ДО);

- правило витіснення або вибору заявок на обслуговування;
- можливість зміни пріоритетів.

Одна з можливих класифікацій дисциплін буферизації відповідно з перерахованими ознаками представлена на рис.2.3.

Залежно від наявності або відсутності пріоритетів між заявками різних класів всі ДБ можуть бути розбиті на дві групи:

- безпріоритетні;
- пріоритетні.

За способом витіснення заявок з накопичувача можна виділити наступні класи ДБ:

- без витіснення заявок (БВЗ) – заявки, що надійшли в систему і застали накопичувач заповненим до кінця, губляться;
- з витісненням заявки даного класу (ВЗДК), тобто такого ж класу, що і заявка, що надійшла;
- з витісненням заявки самого низькопріоритетного класу (ВЗНК) ;
- з витісненням заявки, що належить групі низькопріоритетних класів (ВЗГК) .

Два перших класи відносяться до безпріоритетних ДБ, а решта – до пріоритетних.

ДБ можуть використовувати наступні правила витіснення заявок з накопичувача :

- витіснення випадкове (ВВ);
- витіснення останньої заявки (ВОЗ), тобто заявки, що надійшла в систему пізніше за всіх;
- витіснення «довгої» заявки (ВДЗ), тобто що знаходиться в накопичувачі довше за всіх.

Часто ємність накопичувача в моделях передбачається необмеженою, незважаючи на те, що в реальній системі відповідна ємність обмежена. Таке припущення виправдане в тих випадках, коли ймовірність втрати заявки в реальній системі через переповнення обмеженої ємності накопичувача менше  $10^{-3}$ , оскільки в цьому випадку ДБ практично не впливає на характеристики обслуговування заявок.

На рис. 2.4 представлена класифікація дисциплін обслуговування заявок у відповідності з тими ж ознаками, що і для ДБ.

Залежно від наявності або відсутності пріоритетів між заявками різних класів всі ДО, як і ДБ, можуть бути розбиті на дві групи:

- безпріоритетні;
- пріоритетні.

За способом призначення заявок на обслуговування ДО можуть бути розділені на дисципліни:

- одиночного режиму;
- групового режиму;
- комбінованого режиму.

У ДО одиночного режиму всякий раз на обслуговування призначається тільки одна заявка (перегляд черг з метою призначення на обслуговування в приладі чергової заявки виконується після обслуговування кожної заявки).

У ДО групового режиму всякий раз на обслуговування призначається група заявок однієї черги (перегляд черг з метою чергового призначення на обслуговування виконується тільки після обслуговування всіх заявок раніше призначеної групи). У граничному випадку призначена на обслуговування група заявок може включати в себе всі заявки даної черги. Заявки призначеної на обслуговування групи послідовно вибираються з черги і обслуговуються приладом, після чого на обслуговування призначається наступна група заявок іншої черги відповідно до заданої ДО.

Комбінований режим – комбінація одиночного і групового режимів, коли частина черг заявок обробляється в одиночному режимі, а інша частина – у груповому.

ДО можуть використовувати наступні правила вибору заявок на обслуговування:

- безпріоритетна:

- обслуговування в порядку надходження (ОПН або *FIFO – First In First Out*), коли на обслуговування вибирається заявка, що надійшла в систему раніше за інших;

- обслуговування в зворотному порядку (ОЗП або *LIFO – Last In First Out*) коли на обслуговування вибирається заявка, що надійшла в систему пізніше інших;

- обслуговування у випадковому порядку (ОВП), коли на обслуговування заявка вибирається випадковим чином;

- обслуговування в циклічному порядку (ОЦП), коли на обслуговування заявки вибираються в процесі циклічного опитування накопичувачів в послідовності 1,2 , ... ,  $N$  ( $N$  – кількість накопичувачів ), після чого зазначена послідовність повторюється;

- пріоритетні:

- з відносними пріоритетами (ВП), які демонструють, що пріоритети враховуються тільки в моменти завершення обслуговування заявок при виборі нової заявки на обслуговування та не впливають на процес обслуговування фонові заявки в приладі; іншими словами, надходження в систему заявки з більш високим пріоритетом порівняно з обслуговуємою заявкою в приладі не призводить до переривання обслуговуємої заявки;

- з абсолютними пріоритетами (АП), які демонструють, що, на відміну від ВП, при надходженні високопріоритетної заявки обслуговування заявки з низьким пріоритетом переривається і на

обслуговування приймається високопріоритетна заявка, що надійшла; при цьому перервана заявка може бути повернута в накопичувач або видалена з системи; якщо заявка повернута в накопичувач, то її подальше обслуговування може бути продовжено з перерваного місця або розпочато заново, тобто з самого початку;

- зі змішаними пріоритетами (ЗП), що представляють собою будь-яку комбінацію безпріоритетного обслуговування, ВП і АП;

- з пріоритетами, що чергуються (ПЧ), які є аналогом ВП і проявляються тільки в моменти завершення обслуговування групи заявок однієї черги і призначення нової групи;

- обслуговування за розкладом (ОР), коли заявки різних класів (що знаходяться в різних накопичувачах) вибираються на обслуговування відповідно за деяким розкладом (планом), що задає послідовність опитування черг заявок, наприклад, у випадку трьох класів заявок (накопичувачів) розклад може мати вигляд:  $\{1, 2, 1, 3, 1, 2\}$ .

Дисципліни ОПН, ОЗП, ВП, АП і ЗП відносяться до дисциплін одиночного режиму. Очевидно, що дисципліни групового режиму ОЦП, ПЧ і ОР, в окремому випадку можуть бути реалізовані як ДО одиночного режиму, якщо розмір призначається на обслуговування групи дорівнює 1, при цьому ДО ПЧ вироджується в ДО ВП.

Серед представлених ДО особливе місце займають дисципліни зі змішаними пріоритетами (ЗП), що володіють спільністю по відношенню до перерахованих ДО одиночного режиму.

Для математичного опису ДО ЗП використовується матриця пріоритетів (МП), що представляє собою квадратну матрицю:  $\mathbf{Q} = [g_{ij} | i, j = 1, \dots, H]$ , де  $H$  – число класів заявок, що надходять в систему.

Елемент  $g_{ij}$  матриці задає пріоритет заявок класу  $i$  може приймати такі значення:

- 0 – немає пріоритету;
- 1 – пріоритет відносний (ВП);
- 2 – пріоритет абсолютний (АП).

Елементи МП повинні відповідати таким вимогам:

- $g_{ij} = 0$ , так як між заявками одного і того ж класу не можуть бути встановлені пріоритети;

- якщо  $g_{ij} = 1$  або 2, то  $g_{ji} = 0$ , тому що якщо заявки класу  $i$  мають пріоритет до заявок класу  $j$ , то останні не можуть мати пріоритет до заявок класу  $i$  ( $i, j =$  ).

Залежно від можливості зміни пріоритетів у процесі функціонування системи пріоритетні дисципліни буферизації та обслуговування діляться на два класи:

- зі статичними пріоритетами, які не змінюються з часом;
- з динамічними пріоритетами, які можуть змінюватися в процесі функціонування системи залежно від різних факторів, наприклад, при досягненні деякого критичного значення довжини черги заявок того класу,

що володіє низьким пріоритетом, йому може бути наданий більш високий пріоритет.

## **2.2. Класифікація моделей масового обслуговування**

### **2.2.1. Базові моделі**

При моделюванні реальних систем з дискретним характером функціонування широке застосування знаходять базові моделі у вигляді СМО, які можуть бути класифіковані (рис.2.5):

- за кількістю місць у накопичувачі;
- за кількістю обслуговуючих приладів;
- за кількістю класів заявок, що надходять у СМО.

1. За кількістю місць у накопичувачі СМО діляться на системи:

- без накопичувача, в яких заявка, що надійшла в систему і застала всі обслуговуючі прилади зайнятими обслуговуванням більш високопріоритетних заявок, отримує відмову і втрачається. Такі системи називаються СМО з відмовами;

- з накопичувачем обмеженої ємності (СМО з втратами), в яких заявка, що надійшла втрачається, якщо вона застає накопичувач заповненим до кінця;

- системи з накопичувачем необмеженої ємності (СМО без втрат), в яких для будь якої заявки, що надійшла, завжди знайдеться місце в накопичувачі для очікування.

Надалі накопичувач необмеженої ємності будемо зображати так, як це показано на рис.2.5, а, і накопичувач обмеженою ємності – як на рис.2.5, б.

Припущення про необмежену ємність накопичувача може використовуватися для моделювання реальних систем, в яких ймовірність втрати заявки через переповнення накопичувача обмеженої ємності менше  $10^{-3}$ .

2. За кількістю обслуговуючих приладів СМО поділяються на:

- одноканальні (рис.2.5, а, б, г), що містять один прилад П;
- багатоканальні (рис.2.5, в), що містять  $K$  обслуговуючих приладів

$\Pi_1, \dots, \Pi_K$  ( $K > 1$ ).

У багатоканальних СМО звичайно передбачається, що всі прилади ідентичні і рівнодоступні для будь-якої заявки, тобто за наявності декількох вільних приладів заявка, що надійшла, з рівною імовірністю може потрапити в будь-який з них на обслуговування.

3. За кількістю класів (типів) заявок, що надходять у СМО, розрізняють системи:

- з однорідним потоком заявок (рис.2.5, а, б, в) ;
- з неоднорідним потоком заявок (рис.2.5, г).

Однорідний потік заявок утворюють заявки одного класу, а неоднорідний потік являє собою потік заявок декількох класів.



У СМО, що представляє собою абстрактну математичну модель, заявки відносяться до різних класів в тому випадку, якщо вони в реальній системі розрізняються хоча б одним з таких чинників:

- тривалістю обслуговування;
- пріоритетами.

Якщо ж заявки не розрізняються тривалістю обслуговування і пріоритетами, то в СМО вони можуть бути представлені як заявки одного класу, незалежно від їх фізичної сутності.

### 2.2.2. Мережні моделі

Залежно від структури і властивостей досліджуваних систем їх моделями можуть служити МеМО різних класів. Одна з можливих класифікацій мережних моделей наведена на рис.3.6.

1. Залежно від характеру процесів надходження та обслуговування заявок в мережі МеМО поділяються на:

- стахастичні, в яких процеси надходження та/або обслуговування заявок носять випадковий характер, тобто інтервали часу між заявками, що надходять, та/або тривалість їх обслуговування у вузлах являють собою випадкові величини, описувані відповідними законами розподілів;
- детерміновані, в яких інтервали часу між заявками, що надходять і тривалості їх обслуговування у вузлах є детермінованими величинами.

2. По виду залежностей, що пов'язують інтенсивності потоків заявок у різних вузлах, МеМО поділяються на:

- лінійні, якщо ці залежності лінійні;
- нелінійні, якщо ці залежності є нелінійними.

У лінійних МеМО, як це впливає з визначення, інтенсивність потоку заявок у вузол  $j$  пов'язана з інтенсивністю потоку заявок у вузол  $i$  лінійною залежністю:

$$\lambda_j = \alpha_{ij} \lambda_i,$$

де  $\alpha_{ij}$  – коефіцієнт пропорційності, що показує, у скільки разів відрізняються інтенсивності потоків заявок у вузол  $j$  і у вузол  $i$  ( $i, j \neq$ ).

Оскільки зазначена залежність справедлива для будь-якої пари вузлів, цей вираз можна записати у іншому вигляді і виразити інтенсивність надходження заявок в усі вузли  $j = \dots$  через одну і ту ж інтенсивність, наприклад, через інтенсивність  $\lambda_0$  потоку заявок, що надходить в МеМО з джерела заявок:

$$\lambda_j = \alpha_j \lambda_0. \quad (2.5)$$

В останньому виразі коефіцієнт пропорційності  $\alpha_j \geq 0$  показує, у скільки разів інтенсивність потоку заявок у вузол  $j$  ( $j \neq 0$ ) відрізняється від

інтенсивності джерела заявок, і називається коефіцієнтом передачі. Коефіцієнт передачі може приймати будь-яке додатне значення.

Коефіцієнт передачі відіграє важливу роль при розробці математичних залежностей і розрахунку характеристик функціонування мережних моделей. Це обумовлено тим фізичним змістом, який несе в собі коефіцієнт передачі.

Коефіцієнт передачі можна трактувати як середнє число надходжень заявки в даний вузол за час її знаходження в мережі. Наприклад, якщо коефіцієнт передачі вузла МеМО дорівнює 3, то це означає, що будь-яка заявка за час знаходження в мережі в середньому 3 рази побуває на обслуговуванні в даному вузлі. Значення коефіцієнта передачі, рівне 0,25, означатиме, що в середньому тільки одна заявка з чотирьох потрапить на обслуговування в даний вузол, а три інші обійдуть даний вузол стороною.

У нелінійних МеМО інтенсивності потоків заявок у вузлах пов'язані складнішими нелінійними залежностями, що значно ускладнює їх дослідження.

Нелінійність МеМО може бути обумовлена:

- втратою заявок в мережі, наприклад через обмежену місткість накопичувачів у вузлах;

- розмноженням заявок в мережі, котре полягає, наприклад, у формуванні кількох нових заявок після завершення обслуговування деякої заявки в одному з вузлів мережі.

Таким чином, МеМО є лінійною, якщо в ній заявки не розмножуються і не втрачаються. Нижче розглядаються лінійні МеМО.

3. За кількістю циркулюючих в мережі заявок розрізняють МеМО:

- розімкнуті;
- замкнуті;
- замкнуто-розімкнуті.

Разімкнута (відкрита) МеМО (РМеМО) містить один або декілька зовнішніх незалежних джерел заявок, які генерують заявки в мережу незалежно від кількості заявок, що знаходяться в мережі (рис.2.7, а). У РМеМО одночасно може перебувати будь-яке число заявок, у тому числі, і як завгодно велике, тобто від 0 до нескінченності. З РМеМО пов'язане зовнішнє середовище, з якого надходять заявки в мережу і в яке вони повертаються після обслуговування в мережі. Зовнішнє середовище в РМеМО позначається зазвичай як нульовий вузол "0", і РМеМО, в цьому випадку, зображується у вигляді рис.2.7, б.

Замкнута (закрита) МеМО (ЗМеМО) не містить незалежних зовнішніх джерел заявок і характеризується тим, що в ній циркулює постійне число заявок  $M$  (рис.2.7, в). На графі ЗМеМО з фізичних міркувань, пов'язаних з конкретним поданням процесу функціонування досліджуваної реальної системи, зазвичай виділяється особлива дуга, що відображає процес завершення обслуговування заявок в мережі і миттєвого формування нової

заявки з такими ж параметрами обслуговування, що і завершила обслуговування. Таке трактування дозволяє розглядати заявку, яка завершила обслуговування, як нову заявку, що надійшла в мережу із залежного джерела заявок.

За аналогією з РМеМО на виділеній дузі ЗМеМО відзначається умовна точка "0", що розглядається як нульовий вузол і трактується іноді як фіктивна СМО з нульовою тривалістю обслуговування або як залежне джерело заявок, що генерує заявки тільки в момент надходження деякої заявки на його вхід. Виділення нульового вузла в ЗМеМО переслідує двояку мету: по-перше, досягається однозначність у поданні та математичному описі РМеМО і ЗМеМО, по-друге забезпечується можливість визначення тимчасових характеристик ЗМеМО щодо виділеного вузла "0". Зокрема, час перебування заявок в ЗМеМО розглядається як проміжок часу між двома сусідніми моментами проходження заявки через нульовий вузол.

Замкнуто-разомкнута МеМО (комбінована) являє собою комбінацію ЗМеМО і РМеМО, в яку, крім заявок, що постійно циркулюють в мережі,  $M^*$ , з зовнішнього незалежного джерела надходять заявки такого ж або іншого класу, при цьому сумарне число заявок в мережі  $M \geq M^*$ .

4. За типом циркулюючих заявок розрізняють МеМО:

- однорідні, в яких циркулює один клас заявок (однорідний потік заявок);

- неоднорідні, в яких циркулює кілька класів заявок (неоднорідний потік заявок), що розрізняються хоча б одним з таких чинників:

- тривалістю обслуговування у вузлах;
- пріоритетами;
- маршрутами.

Маршрути заявок різних класів задаються шляхом зазначення номерів класів заявок на відповідних дугах мережі (рис.2.7, г).

## 2.3. Параметри і характеристики СМО

### 2.3.1. Параметри СМО

Для опису СМО використовуються три групи параметрів:

- структурні;
- навантажувальні;
- функціональні параметри (параметри керування).

До структурних параметром відносяться:

- кількість обслуговуючих приладів  $K$ , рівне 1 для одноканальної СМО і  $K > 1$  для багатоканальної СМО;
- кількість  $k$  і ємкості накопичувачів  $E_j (j = )$ ;
- спосіб взаємозв'язку накопичувачів з приладами (у разі багатоканальних СМО), наприклад у вигляді матриці зв'язків.

Навантажувальні параметри СМО включають в себе:

- кількість класів заявок, що надходять в систему,  $H$ , яка дорівнює 1 для СМО з однорідним потоком заявок і  $H > 1$  для СМО з неоднорідним потоком;

- закон розподілу  $A_i(\tau)$  інтервалів часу між заявками, що надходять в систему, класу  $i =$  або, принаймні, перші два моменти розподілу, що задаються, наприклад, у вигляді інтенсивності  $\lambda_i$  і коефіцієнта варіації  $v_{ai}$  інтервалів;

- закон розподілу  $B_i(\tau)$  тривалості обслуговування заявок класу  $i =$  або, принаймні, перші два моменти розподілу, що задаються, наприклад, у вигляді у вигляді середньої тривалості  $b_i$  чи інтенсивності  $\mu_i$  і коефіцієнта варіації  $v_{bi}$ .

Завдання двох перших моментів навантажувальних параметрів найчастіше виявляється достатнім для оцінки характеристик обслуговування заявок на рівні середніх значень .

*Функціональні параметри* задаються у вигляді конкретних стратегій управління потоками заявок в СМО, що визначають правило занесення заявок різних класів в накопичувачі обмеженої ємкості (дисципліна буферизації) і правило вибору їх з черги на обслуговування (дисципліна обслуговування).

### 2.3.2. Позначення СМО (символіка Кендалла)

Для компактного опису систем масового обслуговування часто використовуються позначення, запропоновані Д. Кендаллом у вигляді:

$A/B/N/L$  ,

де  $A$  і  $B$  – задають закони розподілів відповідно інтервалів часу між моментами надходження заявок в систему і тривалості обслуговування заявок в приладі;  $N$  – число обслуговуючих приладів в системі ( $N = 1, 2, \dots, \infty$ );  $L$  – число місць у накопичувачі, яке може приймати значення 0, 1, 2, ... (відсутність  $L$  означає, що накопичувач має необмежену ємкість).

Для завдання законів розподілів  $A$  і  $B$  використовуються наступні позначення:

$G$  (*General*) – довільний розподіл загального вигляду;

$M$  (*Markovian*) – експоненціальне (показниковий) розподіл;

$D$  (*Deterministik*) – детерміноване розподіл;

$U$  (*Uniform*) – рівномірний розподіл;

$E_k$  (*Erlangian*) – розподіл Ерланга  $k$ -го порядку (з  $k$  послідовними однаковими експоненційними фазами);

$h_k$  (*hipoexponential*) – гіпоекспоненціальний розподіл  $k$ -го порядку (з  $k$  послідовними різними експоненційними фазами);

$H_r$  (*Hiperexponential*) – гіперекспоненціальний розподіл порядку  $r$  (з  $r$  паралельними експоненційними фазами);

$g$  (*gamma*) – гамма-розподіл;

$P$  (*Pareto*) – розподіл Парето і т.д.

**Приклади:**

**M/M/1** – одноканальна СМО з накопичувачем необмеженої ємності, в яку надходить однорідний потік заявок з експоненційним розподілом інтервалів часу між послідовними заявками (найпростіший потік) і експоненційною тривалістю обслуговування заявок в приладі.

**M/G/3/10** – триканальна СМО з накопичувачем обмеженою ємністю, рівної 10, в яку надходить однорідний потік заявок з експоненційним розподілом інтервалів часу між послідовними заявками (найпростіший потік) і тривалістю обслуговування заявок, розподіленою за законом загального вигляду.

**D/E<sub>2</sub>/7/0** – семиканальна СМО без накопичувача (ємність накопичувача дорівнює 0), в яку надходить однорідний потік заявок з детермінованими інтервалами часу між послідовними заявками (детермінований потік) і тривалістю обслуговування заявок в приладі, розподіленою за законом Ерланга 2-го порядку.

Для позначення більш складних СМО додатково можуть використовуватися позначення, що описують неоднорідний потік заявок та пріоритети між заявками різних класів.

### 2.3.3. Режими функціонування СМО

СМО може працювати в наступних режимах:

- **сталому або стаціонарному**, коли імовірнісні характеристики системи не змінюються з часом;
- **несталому**, коли характеристики системи змінюються з часом, що може бути обумовлено:
  - початком роботи системи, коли значення характеристик функціонування, міняючись з часом, наближаються в межі до стаціонарних значень (**перехідний режим**);
  - нестаціонарним характером потоку заявок і обслуговування в приладі (**нестаціонарний режим**).

Крім цього, в деяких системах, наприклад в СМО з накопичувачем необмеженої ємності, несталий режим функціонування може бути обумовлений перевантаженням системи, коли інтенсивність надходження заявок перевищує інтенсивність обслуговування, і система не справляється з покладеним на неї навантаженням (режим перевантаження). При цьому характеристики функціонування СМО з плином часу зростають необмежено. Зокрема, довжина черги перед приладом з плином часу стає все більше і в межі наближається до нескінченності.

Зазвичай дослідження СМО з накопичувачем необмеженої ємності проводиться в припущенні про існування стаціонарного режиму, неодмінною умовою якого є вимога відсутності перевантажень, для чого необхідно, щоб інтенсивність надходження заявок була менше, ніж інтенсивність обслуговування. Ця вимога записується для одноканальних СМО у вигляді умови:

$$\lambda < \mu \quad \text{або} \quad \lambda b < 1.$$

Для багатоканальних СМО аналогічна умова має вигляд:

$$\lambda < K\mu \quad \text{або}$$

де  $K$  – число обслуговуючих приладів, а значення  $K\mu$  являє собою сумарну інтенсивність обслуговування заявок в  $K$ -канальній СМО.

У СМО з накопичувачем обмеженої ємкості перевищення інтенсивності надходження заявок над сумарною інтенсивністю обслуговування не призводить до необмеженого росту довжини черги, що обумовлено втратою заявок. Отже, в СМО з накопичувачем обмеженої ємкості перевантаження не приводять до роботи системи в несталому режимі, а призводить лише до зростання числа втрачених заявок. При цьому втрата частини заявок, що надходять у систему, при наявності накопичувача обмеженої ємкості може розглядатися як один з механізмів боротьби з перевантаженнями.

#### 2.3.4. Характеристики СМО з однорідним потоком заявок

Характеристики систем зі стахостичним характером функціонування є випадковими величинами і повністю описуються відповідними законами розподілів. На практиці при моделюванні часто обмежуються визначенням тільки середніх значень (математичних очікувань), рідше – визначенням двох перших моментів цих характеристик.

Як основні характеристики СМО з однорідним потоком заявок використовуються такі величини:

- **навантаження системи:**

$$\rho = \lambda / K\mu \quad ; \quad (3.6)$$

- **коефіцієнт завантаження** або просто завантаження системи, визначається як частка часу, протягом якого система (у разі одноканальної СМО – прилад) працює, тобто виконує обслуговування заявок; завантаження може бути розраховане як відношення середнього часу  $T_p$  роботи одного приладу багатоканальної СМО, до загального часу спостереження  $T$ :

$$\rho = T_p / T \quad (3.7)$$

Час  $T_p$  для СМО з  $K$  обслуговуючими приладами визначається шляхом усереднення часу роботи по всіх приладах:

де  $T_i$  - час роботи приладу ;

Підставляючи останній вираз в (3.7) остаточно отримаємо:

очевидно, що  $0 \leq \rho \leq 1$ ;

• **коефіцієнт простою системи:**

$$\eta = 1 - \rho; \quad (3.8)$$

• **ймовірність втрати заявок:**

$$(3.9)$$

де  $T$  – час роботи системи (спостереження за системою);  $N(T)$  – число заявок, що надійшли в систему за час  $T$ ;  $N_n(T)$  – число втрачених заявок за час  $T$ ;

• **вірогідність обслуговування** заявки, тобто верогідність того, що заявка, що надійшла в систему, буде обслугована:

$$(3.10)$$

де  $N_o(T)$  – число обслугованих в системі заявок за час  $T$ , причому  $N_n(T) + N_o(T) = \pi_o + \pi_n = 1$ ;

• **продуктивність системи, що представляє собою інтенсивність потоку обслугованих заявок, що виходять із системи:**

$$\lambda' = \pi_o \lambda = (1 - \pi_n) \lambda; \quad (3.11)$$

Для СМО з накопичувачем необмеженої ємності, за умови відсутності перевантажень, ймовірність втрати заявок  $\pi_n = 0$  і, отже, продуктивність системи збігається з інтенсивністю надходження заявок в систему:  $\lambda' = \lambda$ ;

• **інтенсивність потоку втрачених (не обслугованих) заявок через обмежену місткість накопичувача:**

$$\lambda'' = \pi_n \lambda = (1 - \pi_o) \lambda; \quad (3.12)$$

Очевидно, що сума інтенсивностей потоків обслугованих і втрачених заявок повинна дорівнювати інтенсивності потоку заявок, що входить в систему:  $\lambda' + \lambda'' = \lambda$ ;

• **середній час очікування заявок в черзі:**  $w$ ;

• **середній час перебування заявок в системі, що складається з часу очікування  $w$  і часу обслуговування  $b$ :**

$$u = w + b; \quad (3.13)$$

• **середня довжина черги заявок:**

$$l = \lambda' w; \quad (3.14)$$

• **середня кількість заявок в системі (в черзі і на обслуговуванні в приладі):**

$$m = \lambda' u; \quad (3.15)$$

Навантаження та завантаження є найважливішими характеристиками СМО, що визначають якість функціонування системи.

Навантаження  $y = \lambda b$  являє собою інтегральну оцінку, об'єднуючу два навантажувальних параметра: частоту використання деякого ресурсу (приладу СМО), що задається у вигляді інтенсивності  $\lambda$  надходження

заявок в СМО, і час використання цього ресурсу, що задається у вигляді середньої тривалості обслуговування заявок в СМО. Навантаження показує кількість роботи, яку необхідно виконати в системі. Якщо значення навантаження  $y < 1$ , то задане навантаження може бути виконане одним обслуговуючим приладом, тобто одноканальна СМО працюватиме без перевантаження. Якщо  $y > 1$ , то реалізація заданого навантаження в одноканальній СМО призведе до режиму перевантаження, що означає, що з плином часу все більше число заявок буде залишатися **НЕ обслуженими**, і у випадку накопичувача необмеженої ємкості черга заявок зростатиме необмежено. Для того щоб система працювала без перевантажень необхідно використовувати багатоканальну СМО, кількість приладів якої повинно бути більше, ніж значення навантаження:  $K > y$ .

У загальному випадку для будь-якої СМО (з накопичувачем обмеженої і необмеженої ємкості) завантаження системи може бути розраховане через навантаження наступним чином:

$$(3.16)$$

де  $K$  – число обслуговуючих приладів в СМО;  $\pi_n$  – ймовірність втрати заявок.

Останній вираз можна трактувати наступним чином:

, якщо СМО працює без перевантаження,  $i=1$ , якщо СМО перевантажена.

### 2.3.5. Характеристики СМО з неоднорідним потоком заявок

Для СМО з неоднорідним потоком заявок, до якої надходять  $H$  класів заявок з інтенсивностями  $\lambda_1, \dots, \lambda_H$  і середніми тривалостями обслуговування  $b_1, \dots, b_H$ , визначаються дві групи характеристик обслуговування заявок:

- характеристики по кожному класу (поток) заявок;
- характеристики об'єднаного (сумарного) потоку заявок.

**Характеристики по кожному класу заявок** ідентичні характеристикам СМО з однорідним потоком:

- навантаження, створюване заявками класу  $i$ :  $y_i = \lambda_i / \mu_i = \lambda_i b_i$ ;
- ймовірність втрати заявок: ;
- ймовірність обслуговування заявки:
- *інтенсивність потоку обслужених заявок (продуктивність по  $i$ -му класу заявок):* ;
- *інтенсивність потоку втрачених заявок:*
- *завантаження системи, створювана заявками класу  $i$ :*

де , – ймовірність втрати заявок класу  $i$  через обмежену місткість накопичувача, якщо ємкість накопичувача – необмежена);  $K$  – число обслуговуючих приладів в СМО;

- *час очікування заявок в черзі:*  $w_i$ ;
- *час перебування заявок в системі:*  $u_i = w_i + b_i$ ;
- *довжина черги заявок:*  $l_i = \lambda_i w_i$ ;



- число заявок в системі (в черзі і на обслуговуванні):  $t_i = \lambda_i u_i$ .

**Характеристики об'єднаного (сумарного) потоку заявок** дозволяють визначити усереднені по всіх класах заявок показники ефективності функціонування СМО:

- сумарна інтенсивність надходження заявок в систему (інтенсивність сумарного потоку):

$$\lambda = \sum_{i=1}^n \lambda_i; \quad (3.17)$$

- сумарне навантаження  $U$  і сумарне завантаження  $R$  системи: (3.18)

причому умова відсутності перевантажень в СМО з неоднорідним потоком заявок і накопичувачем необмеженої ємності має вигляд:

$$R < 1; \quad (3.19)$$

- коефіцієнт простоя системи:
- середній час очікування  $W$  і середній час перебування  $U$  заявок сумарного потоку в системі:

де  $\lambda_i$  – коефіцієнт, що враховує частку заявок класу  $i$  в сумарному потоці, який може трактуватися як ймовірність того, що заявка, яка надійшла в систему, належить класу  $i$ ;

- сумарна довжина черги і сумарне число заявок в системі:

Можна довести, що для характеристик об'єднаного (сумарного) потоку справедливі ті ж фундаментальні співвідношення (3.13) – (3.15), що і для однорідного потоку:

$$U = W + B; \quad L = W; \quad M = W,$$

де  $B$  – середній час обслуговування будь-якої заявки сумарного потоку:

## 2.4. Параметри і характеристики МеМО

### 2.4.1. Параметри МеМО

Для опису лінійних розімкнутих і замкнутих однорідних експоненційних МеМО використовується наступна сукупність параметрів:

- число вузлів в мережі:  $n$ ;
- число обслуговуючих приладів у вузлах мережі:  $K_1, \dots, K_n$ ;
- матриця ймовірностей передач:  $P = [p_{ij} \mid i, j = 0, 1, \dots, n]$ , де  $p_{ij}$  – ймовірність передачі заявки з вузла  $i$  у вузол  $j$ ;
- інтенсивність  $\lambda_0$  джерела заявок, що надходять в **розімкнену** МеМО (РМеМО), або число заявок  $M$ , циркулюючих в замкнутій МеМО (ЗМеМО);
- середні тривалості обслуговування заявок у вузлах мережі:  $b_1, \dots, b_n$ .

Для лінійних МеМО елементи матриці ймовірностей передач повинні задовольняти умові:

Ця умова відображає той факт, що будь-яка заявка, що покинула деякий вузол, обов'язково (з ймовірністю 1) перейде в якийсь вузол, включаючи той же самий чи нульовий. Перехід заявки в нульовий вузол означає, що заявка покинула мережу.

У разі неоднорідних МеМО необхідно додатково задати кількість класів заявок  $H$  в мережі і для кожного класу – матриці ймовірностей передач  $P(h)$ , інтенсивності  $\lambda_0(h)$  або число заявок  $M(h)$ , а також середні тривалості обслуговування  $b_i(h)$  заявок класу у вузлі .

#### 2.4.2. Режими функціонування МеМО

МеМО, як і СМО, може працювати в сталому і несталому режимах.

Очевидно, що для МеМО, як і для СМО, при використанні припущення про стаціонарність вхідного потоку заявок і тривалостей обслуговування заявок у вузлах умова існування сталого режиму збігається з умовою відсутності перевантажень.

Перевантаження в розімкнутій МеМО відсутні, якщо кожен вузол мережі працює без перевантажень. Якщо ж хоча б один з вузлів мережі не справляється з навантаженням, то довжина черги в цьому вузлі почне збільшуватися до нескінченності і, отже, сумарне число заявок в РМеМО зростатиме необмежено. Таким чином, для того щоб в розімкнутій МеМО не було перевантажень, необхідна відсутність перевантажень у всіх вузлах РМеМО, тобто завантаження  $\rho_j$  будь-якого вузла повинна бути строго менше одиниці:

для всіх .

для всіх .

Ця умова може бути записана також у наступному вигляді:

#### 2.4.3. Характеристики МеМО

Характеристики МеМО діляться на два класи:

- вузлові, що описують ефективність функціонування окремих вузлів МеМО;

- мережні, що описують функціонування МеМО в цілому.

Склад вузлових характеристик МеМО, що працює в стаціонарному режимі, такий же, як і для СМО, і для вузла включає в себе наступні характеристики:

- навантаження вузла:  $y_j = \lambda_j b_j = \alpha_j \lambda_0 b_j$ ;
- завантаження вузла: , причому
- коефіцієнт простоя вузла:
- час очікування заявок в черзі:  $w_j$ ;

- час перебування заявок в вузлі:  $u_j = w_j + b_j$ ;
- довжина черги заявок:  $l_j = \lambda_j w_j = \alpha_j \lambda_0 w_j$ ;
- число заявок в вузлі (в черзі і на обслуговуванні):  
 $m_j = \lambda_j u_j = \alpha_j \lambda_0 (w_j + b_j) = l_j + u_j$ .

На основі вузлових характеристик розраховуються мережні характеристики МеМО:

- сумарне навантаження в всіх вузлах:

де  $\rho_j$  – навантаження вузла  $j$ , причому ;

- сумарне завантаження в всіх вузлах:

де  $\rho_j$  – завантаження вузла  $j$ , причому ;

- середнє число заявок, що надходять в чергу всіх вузлів і очікують обслуговування:

$$(3.24)$$

де  $l_j$  – середня довжина черги заявок у вузлі  $j$ ;

- середнє число заявок, що знаходяться в мережі:

$$(3.25)$$

де  $l_j$  – середня число заявок у вузлі  $j$ ;

- середній час очікування в мережі:

$$(3.26)$$

де  $l_j$  – середній час очікування заявок у вузлі  $j$ ; – коефіцієн передачі для вузла  $j$ ;

- середній час перебування заявок в мережі:

$$(3.27)$$

де  $l_j$  – середній час перебування заявок у вузлі  $j$ ;

• продуктивність замкнутої МеМО, визначається як інтенсивність потоку заявок, що проходять через виділений нульовий вузол замкнутої мережі, і представляє собою середнє число заявок, обслугованих в ЗМеМО за одиницю часу; продуктивність ЗМеМО може бути розрахована наступним чином:

$$(3.28)$$

Слід зазначити, що для мережних характеристик МеМО виконуються ті ж фундаментальні співвідношення, що і для СМО, а саме:

$$(3.29)$$

$$(3.30)$$

$$(3.31)$$

$$(3.32)$$

де сумарний час обслуговування заявки у всіх вузлах за час її знаходження в мережі.

Вирази (3.29) і (3.30) являють собою формули Літтла для розрахунку мережних характеристик МеМО.

З (3.30) може бути отримана ще одна важлива формула для розрахунку продуктивності ЗМеМО:

(3.33)

Для неоднорідної МеМО перераховані характеристики визначаються як для кожного класу окремо, так і для об'єднаного (сумарного) потоку заявок.

### 3. АНАЛІТИЧНЕ МОДЕЛЮВАННЯ

#### 3.1. Одноканальні СМО з однорідним потоком заявок

Розглянемо одноканальну СМО з однорідним потоком заявок при наступних припущеннях (рис.3.1):

1) СМО містить один обслуговуючий прилад, в якому в кожен момент часу може обслуговуватися тільки одна заявка;

2) перед приладом маєтся накопичувач **N** необмеженої ємкості, що означає присутність відмов заявкам при їх постановці в чергу **Ч**, тобто будь-яка заявка, що надійшла, завжди знайде в накопичувачі місце для очікування не залежно від того, скільки заявок вже знаходиться в черзі;

3) заявки надходять в СМО з інтенсивністю  $\lambda$ ;

4) середня тривалість обслуговування однієї заявки в приладі дорівнює **b**, причому тривалості обслуговування різних заявок не залежать одне від одного;

5) обслуговуючий пристрій не простоє, якщо в системі (накопичувачі) є хоча б одна заявка, причому після завершення обслуговування чергової заявки миттєво з накопичувача вибирається наступна заявка;

6) заявки з накопичувача вибираються відповідно до безпріоритетної дисципліною обслуговування в порядку надходження за правилом «першим прийшов – першим обслужений» (*FIFO – First In First Out*).

7) в системі існує стаціонарний режим, що припускає відсутність перевантажень, тобто навантаження  $\rho$ , отже, завантаження системи менше:  $\rho = \lambda b < 1$ .

В якості розрахункової характеристики обслуговування заявок в СМО будемо використовувати середній час очікування заявок.

##### 3.1.1. Характеристики експоненційної СМО М/М/1

Нехай заявки, що надходять в одноканальну СМО, утворюють найпростіший потік з інтенсивністю  $\lambda$ , а тривалість обслуговування заявок розподілена за експоненційним законом із середнім значенням **b**, причому  $\rho = \lambda b < 1$ , тобто система працює в сталому режимі. Така СМО з однорідним потоком заявок називається експоненційною.

З використанням методу середніх значень або марківської моделі можна отримати наступні вирази для розрахунку середніх значень:

- часу очікування заявок

- часу перебування заявок
- довжини черги заявок
- числа заявок в системі (в черзі і на обслуговуванні)

### 3.1.2. Характеристики неекспоненціальної СМО M/G/1

Нехай заявки, що надходять в одноканальну СМО, утворюють найпростіший потік з інтенсивністю  $\lambda$ , а тривалість  $\tau_b$  обслуговування заявок розподілена за довільним законом  $B(\tau)$  із середнім значенням  $b$  і коефіцієнтом варіації  $v_b$ .

З використанням методу середніх значень можна показати, що середній час очікування заявок визначається за формулою Поллачека-Хинчина:

де  $\rho = \lambda b < 1$  – завантаження системи.

Середній час перебування заявок в системі:

### 3.1.3. Характеристики неекспоненціальної СМО G/M/1

Нехай у одноканальну СМО з інтенсивністю  $\lambda$  надходить випадковий потік заявок довільного виду, що задається функцією розподілу інтервалів між заявками  $A(\tau)$ , а тривалість  $\tau_b$  обслуговування заявок розподілена за експоненціальним законом  $B(\tau)$  із середнім значенням  $b$  (інтенсивністю  $\mu=1/b$ ).

СМО G/M/1 є симетричною по відношенню до СМО M/G/1. Однак отримання кінцевого результату у вигляді аналітичного виразу для розрахунку середнього часу очікування, за аналогією з попередньою моделлю, в загальному випадку, виявляється неможливим. Це обумовлено тим, що середній час очікування, як і інші числові моменти, залежить не тільки від двох перших моментів інтервалів між заявками, що надходять, а й від моментів вищого порядку, тобто від закону розподілу інтервалів між заявками.

Середній час очікування заявок в черзі може бути розрахований наступним чином:

$$w = \zeta b / (1 - \zeta), \quad (3.6)$$

де  $\zeta$  – єдиний в області  $0 < \zeta < 1$  корінь рівняння

$$\zeta = A^*(\mu - \mu\zeta). \quad (3.7)$$

Тут  $A^*(s)$  – перетворення Лапласа щільності розподілу  $\alpha(\tau)$  інтервалів між заявками, що надходять в систему:

### 3.1.4. Характеристики СМО загального вигляду G/G/1

Найбільш загальним випадком одноканальних систем масового обслуговування є СМО типу G/G/1, в яку надходить довільний потік заявок загального виду з функцією розподілу інтервалів між заявками  $A(\tau)$ . Тривалість обслуговування заявок в приладі розподілена за довільним законом  $B(\tau)$ .

Розрахунок таких систем вимагає завдання конкретних законів розподілів, що не дозволяє отримати аналітичне рішення в загальному вигляді. Аналітичне рішення можливе тільки для деяких окремих розподілів, пов'язаних, наприклад, з експоненціальним розподілом. Для більшості законів розподілів інтервалів між заявками, надходять в систему, і тривалостей їх обслуговування в приладі неможливо отримати точне рішення в аналітичній формі.

Як показав аналіз численних опублікованих результатів, одним з найбільш вдалих наближень для розрахунку середнього часу очікування в СМО G/G/1 є наступна формула:

де  $p = \lambda b < 1$  – завантаження системи;  $\lambda$ , – інтенсивність потоку заявок і коефіцієнт варіації інтервалів між заявками, що надходять в систему;  $b$ , – середнє значення і коефіцієнт варіації тривалості обслуговування заявок;  $f()$  – коригувальна функція, що розраховується залежно від значення коефіцієнта варіації :

## 3.2. Багатоканальні СМО з однорідним потоком заявок

Розглянемо багатоканальну СМО з  $K$  ідентичними обслуговуючими приладами і накопичувачем необмеженої ємності, в яку надходять заявки, що утворюють найпростіший потік з інтенсивністю  $\lambda$  (рис.3.2). Тривалість  $\tau_b$  обслуговування заявок розподілена за експоненціальним законом із середнім значенням  $b$ . Вибір заявок з черги на обслуговування здійснюється відповідно до безпріоритетної дисципліни обслуговування в порядку надходження за правилом «першим прийшов – першим обслужений» (*FIFO – First In First Out*).

### 3.2.1. Характеристики багатоканальної СМО M/M/K

В якості основної характеристики функціонування СМО будемо

використовувати середній час очікування  $w$  заявок.

Точний метод розрахунку характеристик обслуговування заявок в багатоканальній СМО розроблений при наступних припущеннях:

- потік заявок – *найпростіший*;
- тривалість обслуговування заявок розподілена за *експоненційним законом* із середнім значенням  $b$ ;
- всі  $K$  приладів – *ідентичні*, і будь-яка заявка може бути обслужена будь-яким прибором;
- ємкість накопичувача – не обмежена;
- в системі *відсутні перевантаження*, тобто завантаження системи менше 1:

При цих припущеннях середній час очікування заявок визначається наступним чином:

$$(3.9)$$

де  $P$  – ймовірність того, що всі  $K$  приладів зайняті обслуговуванням заявок.

Ймовірність  $P$  визначається як:

де  $P_0$  – ймовірність простою багатоканальній СМО, тобто ймовірність того, що в системі немає заявок:

### 3.3. Одноканальні СМО з неоднорідним потоком заявок

Розглянемо одноканальну СМО з неоднорідним потоком заявок, до якої надходять  $H$  класів заявок, що утворюють найпростіші потоки з інтенсивностями  $\lambda_1, \dots, \lambda_H$ . Тривалість обслуговування заявок класу  $k$  розподілена за довільним законом із середнім значенням  $b_k$  і коефіцієнтом варіації  $\sigma_k$ . Вибір заявок з черги на обслуговування здійснюється відповідно до заданої дисципліною обслуговування, в якості якої будемо розглядати:

- дисципліну обслуговування безпріоритетну (ДО БП), при якій заявки вибираються на обслуговування в порядку надходження;
- дисципліну обслуговування заявок з відносними пріоритетами (ДО ВП);
- дисципліну обслуговування заявок з абсолютними пріоритетами (ДО АП).

В якості основної характеристики, що описує ефективність функціонування системи, будемо розглядати середні часи очікування заявок різних класів, на основі яких легко можуть бути розраховані всі інші характеристики з використанням фундаментальних залежностей.

При цьому слід мати на увазі, що представлені нижче формули були отримані при наступних припущеннях:

- 1) СМО містить один обслуговуючий прилад, який в кожен момент часу може обслуговувати тільки одну заявку;



2 ) СМО має накопичувач заявок необмеженої ємкості, що означає відсутність відмов заявкам при їх постановці в чергу, тобто будь-яка заявка, що надходить, завжди знайде в накопичувачі місце для очікування незалежно від того, скільки заявок вже знаходиться в черзі;

3 ) заявки різних класів, що надходять в СМО незалежно одна від одної, утворюють найпростіші потоки;

4 ) тривалості обслуговування заявок кожного класу в приладі розподілені за довільним законом і не залежать одна від одної;

5) обслуговуючий пристрій не простоює, якщо в системі ( накопичувачі) є хоча б одна заявка будь-якого класу, причому після завершення обслуговування чергової заявки миттєво з накопичувача вибирається наступна заявка відповідно до заданої дисципліною обслуговування;

б) при використанні ДО БП заявки різних класів вибираються на обслуговування тільки залежно від часу надходження в систему за правилом «раніше прийшов – раніше обслужений», незалежно від номера класу, до якого належить заявка;

7) при використанні пріоритетних дисциплін (ДО ВП та ДО АП) пріоритети класам заявок призначені за принципом «клас з меншим номером має більш високий пріоритет », тобто найвищий пріоритетом мають заявки класу 1;

8) у разі ДО АП заявка, обслуговування якої перервано більш високопріоритетною заявкою, повертається в накопичувач, де очікує подальшого обслуговування, причому її обслуговування продовжується з перерваного місця.

### **3.3.1. Характеристики та властивості ДО БП**

При безпріоритетній ДО середні часи очікування однакові для всіх класів заявок і визначаються за такою формулою:

$$\bar{t} = \frac{1}{\mu} \sum_{k=1}^n \rho_k \quad (3.10)$$

де – сумарне завантаження системи.

### **3.3.2. Характеристики та властивості ДО ВП**

Пріоритети називаються відносними, якщо вони враховуються тільки в момент вибору заявки на обслуговування і не позначаються на роботі системи в період обслуговування заявки будь-якого класу (пріоритету) . Відносність пріоритету пов'язана з наступним. Після завершення обслуговування будь-якої заявки з черги на обслуговування вибирається заявка класу з найбільш високим пріоритетом, яка надійшла раніше інших заявок цього класу (такого ж пріоритету) . Якщо в процесі її обслуговування в систему надійдуть заявки з більш високим пріоритетом, то обслуговування розглянутої заявки не буде припинена. Таким чином, пріоритет відносний в тому сенсі, що він має місце лише в момент вибору заявок на обслуговування і відсутній, якщо прилад зайнятий

обслуговуванням заявки.

Введення відносних пріоритетів (ВП) дозволяє зменшити в порівнянні з ДО БП час очікування високопріоритетних заявок.

Для ДО ВП середній час очікування заявок класу  $k$  визначається за наступною формулою:

де  $\sum_{k=1}^n$  – сумарні завантаження, створювані заявками, які мають пріоритет не нижче  $(k-1)$  і  $k$  відповідно:

### 3.3.3. Характеристики та властивості ДО АП

Пріоритет, перериваючий обслуговування фонові заявки, називається абсолютним, а відповідна дисципліна – дисципліною обслуговування з абсолютними пріоритетами (ДО АП).

Перерване заявка може бути втрачена або повернута в накопичувач, де вона буде чекати подальшого обслуговування. В останньому випадку можливі два варіанти продовження обслуговування перерваної заявки:

- обслуговування спочатку, тобто перервана заявка буде обслуговуватися заново з самого початку;
- дообслуговування, коли обслуговування перерваної заявки в приладі буде виконуватися з перерваного місця.

Надалі, якщо не обумовлено інше, будемо припускати дообслуговування перерваної заявки.

Для ДО АП середній час очікування заявок класу  $k$  визначається за наступною формулою :

де  $\sum_{k=1}^n$  – сумарні завантаження, створювані заявками, які мають пріоритет не нижче  $(k-1)$  та  $k$  відповідно.

## 3.4. Розімкнені експоненційні МеМО з однорідним потоком заявок

### 3.4.1. Опис розімкнутих МеМО

Розглянемо розімкнену експоненційну мережу масового обслуговування (МеМО) з однорідним потоком заявок при наступних припущеннях:

- 1) розімкнена МеМО (РМеМО) довільної топології містить  $n$  вузлів;
- 2) після завершення обслуговування в якому-небудь вузлі передача заявки в інший вузол відбувається миттєво;
- 3) в якості вузлів можуть бути як *одноканальні*, так і *багатоканальні* СМО;
- 4) всі прилади багатоканального вузла є ідентичними, і будь-яка заявка може обслуговуватися будь-яким приладом;
- 5) заявка, що надійшла в багатоканальний вузол, коли всі або декілька приладів вільні, направляється випадковим чином в будь-який вільний прилад;

6) в кожному вузлі РМеМО мається накопичувач заявок необмеженої ємкості, що означає відсутність відмов заявкам, що надходять, при їх постановці в чергу, тобто будь-яка заявка, що надходить у вузол, завжди знайде в накопичувачі місце для очікування незалежно від того, скільки заявок вже знаходиться в черзі;

7) заявки надходять в РСеМО із зовнішнього незалежного джерела і утворюють найпростіший потік заявок;

8) тривалості обслуговування заявок у всіх вузлах мережі являють собою випадкові величини, розподілені за експоненційним законом;

9) обслуговуючий прилад будь-якого вузла не простоє, якщо в його накопичувачі є хоча б одна заявка, причому після завершення обслуговування чергової заявки миттєво з накопичувача вибирається наступна заявка;

10) в кожному вузлі мережі заявки з накопичувача вибираються в відповідності до безпріоритетної дисципліни обслуговування в порядку надходження (ОПН) за правилом «першим прийшов – першим обслужений» (FIFO – First In First Out).

Для опису лінійних розімкнутих однорідних експоненційних МеМО необхідно задати таку сукупність параметрів:

- число вузлів в мережі:  $n$ ;
- число обслуговуючих приладів у вузлах мережі:  $K_1, \dots, K_n$ ;
- матрицю ймовірностей передач:  $P = [p_{ij} \mid i, j = 0, 1, \dots, n]$ , де ймовірності передач  $p_{ij}$  повинні задовольняти умові: сума елементів кожного рядка повинна дорівнювати 1;
- інтенсивність  $\lambda_0$  джерела заявок, що надходять в РМеМО;
- середні тривалості обслуговування заявок у вузлах мережі:  $b_1, \dots, b_n$ .

На основі перерахованих параметрів можуть бути розраховані вузлові і мережні характеристики, що описують ефективність функціонування відповідно вузлів і РМеМО в цілому.

Розрахунок характеристик функціонування лінійних розімкнутих однорідних експоненційних МеМО базується на еквівалентному перетворенні мережі і проводиться в чотири етапи:

- розрахунок коефіцієнтів передач  $\alpha_j$  і інтенсивностей потоків заявок  $\lambda_j$  в вузлах МеМО;
- перевірка умови відсутності перевантажень в МеМО;
- розрахунок вузлових характеристик;
- розрахунок мережних характеристик.

### **3.4.2. Розрахунок коефіцієнтів передач і інтенсивностей потоків заявок у вузлах РМеМО**

Покажемо що, інтенсивності  $\lambda_0, \dots, \lambda_n$  потоків заявок, що надходять у вузли  $0, \dots, n$  мережі, однозначно визначаються ймовірностями передач  $p_{ij}$  ( $i, j = 0, 1, \dots, n$ ) і задають маршрути заявок в МеМО.

Будемо розглядати тільки сталий режим.

Так як в лінійній МеМО заявки не розмножуються і не губляться, то

інтенсивності вхідних і вихідних потоків для будь-якого вузла будуть рівні між собою.

Інтенсивність потоку заявок, що входять в будь-який вузол  $j$  мережі, дорівнює сумі інтенсивностей потоків заявок, що надходять в нього з інших вузлів (рис.3.3). Оскільки заявки з вузла  $i$  надходять у вузол  $j$  з імовірністю  $p_{ij}$ , то інтенсивність потоку заявок, що надходять з  $i$  в  $j$ , дорівнює  $p_{ij}\lambda_i$ , де  $\lambda_i$  – інтенсивність що виходить з  $i$ , отже, вхідного потоку заявок вузла  $i$ . З урахуванням цього, на вході вузла  $j$  мається потік з інтенсивністю:

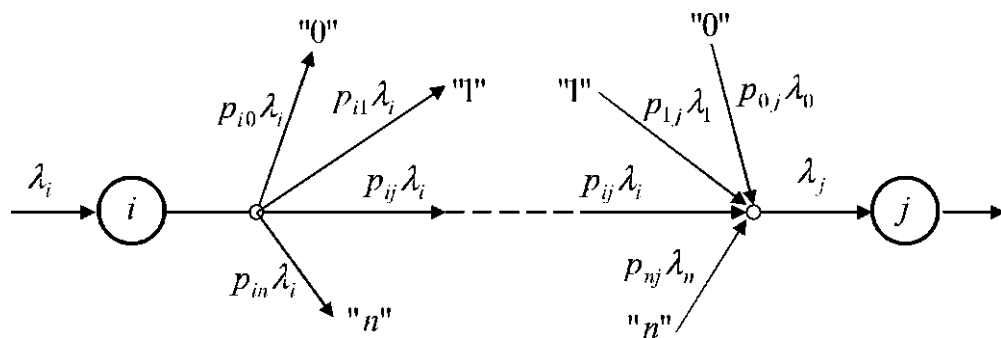


Рис.3.3. – До розрахунку інтенсивностей потоків заявок в вузлах РМеМО

Вираз (3.14) являє собою систему лінійних алгебраїчних рівнянь  $(n + 1)$ -го порядку, з якої можуть бути знайдені інтенсивності потоків заявок у вигляді співвідношення). коефіцієнт – називається коефіцієнтом передачі і визначає середнє число входжень заявки у вузол у за час її знаходження в мережі, причому .

Для розімкнутої МеМО відома інтенсивність джерела заявок. Можна показати, що система рівнянь для розрахунку інтенсивностей має єдине рішення виду , де – задана величина.

### 3.4.3. Перевірка умови відсутності перевантажень в МеМО

В розімкнутій МеМО відсутні перевантаження, якщо виконується умова:

Якщо зазначена умова не виконується, то стаціонарний режим в розімкнутій МеМО може бути реалізований одним із таких способів:

- зменшенням інтенсивності зовнішнього джерела заявок до значення, при якому ця умова буде виконуватися;
- збільшенням кількості обслуговуючих приладів у перевантажених вузлах;
- зменшенням тривалостей  $b$  обслуговування заявок в перевантажених вузлах;
- зменшенням коефіцієнтів передачі в перевантажувальних вузлах.

### 3.4.4. Розрахунок вузлових характеристик РММО

Один і той же об'єкт, розглянутий на різних рівнях деталізації, можна представити різними моделями масового обслуговування, характеристики яких однакові або відрізняються на величину, що не перевищує заданої похибки. При виконанні певних умов такі моделі легко перетворюються одна в одну.

Для мережних моделей у вигляді розімкнутих і замкнених ММО можуть використовуватися два види перетворень:

- еквівалентне перетворення;
- толерантне перетворення.

Дві мережні моделі еквівалентні, якщо порівнювані характеристики цих моделей не відрізняються одна від одні.

Дві мережні моделі толерантні (подібні), якщо значення певних характеристик відрізняються одна від одні на величину, що не перевищує задану.

Використання властивостей еквівалентних і толерантних моделей дозволяє спростити розрахунок характеристик моделей шляхом заміни складних мережних моделей простішими. Еквівалентними можуть бути мережні моделі одного типу (наприклад, дві замкнуті мережі), толерантними – моделі як одного, так і різних типів.

Розрахунок характеристик функціонування лінійних розімкнутих однорідних експоненційних ММО базується на еквівалентному перетворенні мережі, що полягає в поданні розімкнутої ММО з  $n$  вузлами у вигляді  $n$  незалежних експоненційних СМО типу М/М/ N (найпростіший потік заявок, тривалість обслуговування розподілена за експоненційним законом,  $N$  обслуговуючих приладів). При цьому інтенсивність вхідного потоку заявок в СМО, що відображає вузол  $j$  ( $j = \dots$ ) мережі, визначається з системи алгебраїчних рівнянь через інтенсивність вхідного в мережу потоку і коефіцієнту передачі вузла:  $\lambda_j = \alpha_j \lambda_0$ , а середня тривалість обслуговування заявок в СМО дорівнює тривалості обслуговування  $b_j$  заявок у відповідному вузлі ММО.

Характеристики всіх  $n$  СМО (час очікування заявок в черзі і перебування в системі, довжина черги і число заявок в системі, середня кількість зайнятих приладів і т.д.) являють собою вузлові характеристики ММО.

Середній час очікування заявок в черзі може бути розрахований з використанням виразу для багатоканальних СМО типу М / М / N або виразу для одноканальних СМО типу М/М/1, інші характеристики вузла  $j$  ( $j = \dots$ ) – з використанням фундаментальних співвідношень:

- навантаження у вузлі  $j$ , що показує середнє число зайнятих приладів:  
 $y_j = \lambda_j b_j$ ,
- завантаження вузла  $j$ :  $\rho_j = \min(y_j / K_j; 1)$ , де  $K_j$  – число обслуговуючих приладів у вузлі  $j$ ;
- коефіцієнт простою вузла:  $\pi_j = 1 - \rho_j$ ;

- час перебування заявок у вузлі:  $u_j = w_j + b_j$  ;
- довжина черги заявок:  $l_j = \lambda_j w_j$  ;
- число заявок у вузлі (в черзі і на обслуговуванні в приладі):  
 $m_j = \lambda_j u_j$ .

Розраховані таким чином характеристики окремих СМО в точності відповідають вузловим характеристикам вихідної МеМО.

### 3.4.5. Розрахунок мережних характеристик РМеМО

Мережні характеристики, що описують ефективність функціонування МеМО в цілому, розраховуються на основі отриманих значень вузлових характеристик.

До складу мережних характеристик входять:

- середнє число заявок, що очікують обслуговування в мережі, і середнє число заявок, що знаходяться в мережі:

$$L = ; \quad M = , \quad (3.14)$$

де  $l_j$  – середня довжина черги і  $m_j$  – середнє число заявок у вузлі  $j$ ;

- середній час очікування і середній час перебування заявок в мережі:

$$W = ; \quad U = , \quad (3.15)$$

де  $w_j$  і  $u_j$  – відповідно середній час очікування і середній час перебування заявок у вузлі  $j$ ;  $\alpha_j$  – коефіцієнт передачі для вузла  $j$ , що показує середнє число проходжень заявки у вузол  $j$  за час її знаходження в мережі.

## 4. ЧИСЕЛЬНЕ МОДЕЛЮВАННЯ (МОДЕЛІ ВИПАДКОВИХ ПРОЦЕСІВ)

При вивченні складних систем зі стохастичним характером функціонування корисною математичною моделлю є випадковий процес, який розвивається в залежності від ряду випадкових факторів. Прикладами випадкових процесів можуть служити процеси надходження і передачі даних в телекомунікаційній мережі, процеси виконання завдань і обміну даними з зовнішніми пристроями в обчислювальній системі і т.п.

Більшість моделей дискретних систем зі стохастичним характером функціонування будується на основі моделей масового обслуговування, процеси в яких є випадковими і, в багатьох випадках, марківськими або деяким чином пов'язані з марківськими процесами. Тому для вирішення таких завдань **теорії масового обслуговування** може використовуватися математичний апарат теорії марківських процесів. *Застосування марківських процесів виявляється особливо ефективним і результативним при дослідженні систем і мереж масового обслуговування з накопичувачами обмеженою ємкості.*

Математичний опис марківських процесів зазвичай представляється у вигляді систем диференційних (у разі нестационарного режиму ) або алгебраїчних (для стаціонарного режиму) рівнянь, рішення яких, в загальному випадку, отримати в явному вигляді не вдається. Це обумовлює необхідність застосування чисельних методів розв'язання систем диференційних або алгебраїчних рівнянь.

### 4.1. Поняття марковського випадкового процесу

Якщо безліч станів, в яких може перебувати процес злічені, тобто всі можливі стани можуть бути пронумеровані, то відповідний процес називається випадковим процесом з дискретними станами або просто дискретним випадковим процесом.

Якщо безліч станів не може бути пронумеровано, то маємо випадковий процес з безперервними станами або просто безперервний випадковий процес.

Випадковий процес називається марківським, якщо ймовірність будь-якого стану в майбутньому залежить тільки від його стану в сьогоднішній і не залежить від того, коли і яким чином процес виявився в цьому стані.

Процес  $Z(t)$ , що описує поведінку системи, називається ланцюгом Маркова.

### 4.2. Марківські моделі систем масового обслуговування

Розробка марківської моделі досліджуваної системи в термінах випадкових процесів передбачає виконання таких етапів:

- кодування станів випадкового процесу;
- побудова розміченого графа переходів;
- формування матриці інтенсивностей переходів;

- складання системи лінійних алгебраїчних рівнянь для розрахунку стаціонарних ймовірностей станів марківського процесу.

Матриця інтенсивностей переходів може використовуватися для завдання системи лінійних алгебраїчних рівнянь в матричному вигляді при комп'ютерному розрахунку стаціонарних ймовірностей.

При дослідженні різного роду реальних систем, моделями яких служать СМО, навряд чи когось цікавлять ймовірності станів. Набагато більший інтерес представляють такі характеристики СМО, як довжина черги заявок перед обслуговуючим приладом, час очікування і час перебування заявок в системі, завантаження і коефіцієнт простою системи, частка втрачених заявок і т.д. , значення яких можуть бути розраховані по знайденим значенням стаціонарних ймовірностей станів.

#### **4.2.1. Одноканальна СМО без накопичувача (М/М/1/0)**

Розглянемо найпростішу одноканальну систему масового обслуговування (СМО) з відмовами, в яку надходить випадковий потік заявок, затриманих в приладі на випадковий час (рис.4.1, а). Оскільки перед обслуговуючим приладом немає накопичувача, то заявка, що надійшла в систему і застала прилад зайнятим, отримує відмову в обслуговуванні і втрачає-ся. Таким чином, в системі, крім вхідного потоку заявок з інтенсивністю  $\lambda$ , утворюються ще два потоки: виходить потік обслужених в приладі заявок з інтенсивністю  $\lambda'$  і потік не обслужених заявок (що отримали відмову в обслуговуванні) з інтенсивністю  $\lambda''$ . Очевидно, що  $\lambda' + \lambda'' = \lambda$ .

### **1. Опис системи.**

1.1. Система містить один обслуговуючий прилад (П), тобто є одноканальною.

1.2. У систему надходить один клас заявок, тобто потік заявок однорідний.

1.3. У приладі відбувається затримка (обслуговування) заявок, що надходять в систему, на деякий випадковий час.

1.4. Перед приладом не передбачені місця для очікування заявок, тобто в системі відсутній накопичувач.

### **2. Припущення і допущення.**

2.1. Заявки, що надходять в систему, утворюють найпростіший потік з інтенсивністю  $\lambda$ .

2.2. Тривалість обслуговування заявок в приладі розподілена за експоненційним законом з інтенсивністю  $\mu = 1/b$ , де  $b$  – середня тривалість обслуговування.

2.3 . Дисципліна буферизації – з відмовами: заявка, що надійшла в систему і застала прилад зайнятим обслуговуванням іншої заявки, втрачається.



2.4 . Дисципліна обслуговування – в природному порядку: заявка, що надійшла в систему і застала прилад вільним, приймається на обслуговування.

Очевидно, що в СМО з відмовами завжди існуватиме сталий режим, оскільки навіть при великих значеннях навантаження () число заявок в системі не може вирости до незкінченності. Це обумовлено тим, що з ростом навантаження збільшується частка заявок, які отримують відмову в обслуговуванні.

### 3. Кодування станів випадкового процесу.

Як параметр, що описує стан випадкового процесу, будемо розглядати кількість заявок  $k$ , що знаходяться в СМО. Очевидно, що система в будь-який момент часу може знаходитися в одному з двох станів:

$E_0$ :  $k = 0$  – в системі немає заявок;

$E_1$ :  $k = 1$  – в системі (на обслуговуванні в приладі) знаходиться 1 заявка (прибор працює).

### 4. Розмічений граф переходів випадкового процесу (рис.4.1, б).

У процесі функціонування даної системи в один і той же момент часу може наступити тільки одна з двох можливих подій, які призводять до зміни стану випадкового процесу, що протікає в системі.

1. Надходження заявки в систему з інтенсивністю  $\lambda$ . При цьому:

- якщо випадковий процес знаходиться в стані  $E_0$  (прилад простоює), то відбудеться перехід в стан  $E_1$ , (почнеться обслуговування заявки, що надійшла), причому інтенсивність переходу збігається з інтенсивністю надходження заявок в систему  $\lambda$ ;

- якщо ж випадковий процес знаходиться в стані  $E_1$  (прилад працює), то стан  $E_1$  випадкового процесу не зміниться, що буде відповідати відмови в обслуговуванні заявці, що надійшла.

Таким чином, перехід зі стану  $E_0$  в стан  $E_1$  відбувається з інтенсивністю  $\lambda$ .

2. Завершення обслуговування заявки, що знаходиться в приладі. Ця подія може наступити тільки в тому випадку, якщо в приладі на обслуговуванні знаходиться заявка, тобто випадковий процес знаходиться в стані  $E_1$ . При цьому відбувається перехід в стан  $E_0$ , причому інтенсивність переходу збігається з інтенсивністю обслуговування заявки в приладі  $\mu$ .

### 5. Матриця інтенсивностей переходів.

Графу переходів відповідає матриця інтенсивностей переходів:

$E_i$	0	1
$G = 0$	$-\lambda$	$\lambda$
1	$\mu$	$-\mu$

Дійсно, перехід зі стану  $E_0$  в стан  $E_1$  відповідає надходженню заявки в систему з інтенсивністю  $\lambda$ , а перехід зі стану  $E_1$  в стан  $E_0$  відповідає завершенню обслуговування заявки в приладі з інтенсивністю  $\mu$ .

Діагональні елементи матриці визначаються з умови – сума елементів кожного рядка повинна дорівнювати нулю.

### 6. Система рівнянь.

Система рівнянь для визначення стаціонарних ймовірностей, складена по графу переходів з застосуванням правила 1, має вигляд:

$$(4.1)$$

де останнє рівняння є нормувальною умовою.

З огляду на те, що перше і друге рівняння однакові (або, як кажуть математики, лінійно залежні) і видаляючи одне з них, остаточно отримаємо:

$$(4.2)$$

Вирішуючи цю систему, отримаємо наступні значення стаціонарних ймовірностей станів марківського процесу:

$$p_0 = \dots; \quad p_1 = \dots, \quad (3.18)$$

де  $y = \lambda/\mu$  – навантаження системи.

### 7. Розрахунок характеристик СМО.

Для розрахунку характеристик СМО можна скористатися наступними математичними залежностями:

1) навантаження:  $y = \lambda/\mu = \lambda b$  (за визначенням);

2) завантаження визначається як ймовірність роботи приладу  $\rho = p_1$  і не збігається з навантаженням навіть в разі  $y < 1$ , що характерно для систем з відмовами і втратами заявок, причому завжди  $\rho < y$ ;

3) коефіцієнт простою системи визначається як ймовірність відсутності заявок в системі або, за визначенням, через завантаження системи:

$$\eta = p_0 = 1 - \rho;$$

4) середнє число заявок в системі:  $m = p_1/\rho$ , яке визначається як математичне сподівання випадкової величини: в системі може знаходитися або нуль заявок з ймовірністю  $p_0$ , або одна заявка з ймовірністю  $p_1$ , тоді середнє число заявок  $m = 0 \cdot p_0 + 1 \cdot p_1 = p_1$ ;

5) ймовірність втрати заявок в результаті відмови в обслуговуванні через зайнятість приладу відповідно збігається з ймовірністю того, що система зайнята обслуговуванням заявок:

$$\pi = 1 - K = 1 - \dots = \dots$$

де враховано, що для даної СМО без накопичувача  $p_1 =$

6) продуктивність системи:  $\lambda' = (1 - \pi) \lambda$ , яка визначається як інтенсивність потоку обслужених заявок на виході системи;

7) інтенсивність потоку не обслужених заявок, тобто тих, що отримали відмову:  $\lambda'' = \pi \lambda$ ;

8) середній час перебування заявок в системі:  $u = m/\lambda' = b$  (визначається за формулою Літтла і, як слід було очікувати, дорівнює середній тривалості обслуговування заявок).

#### **4.2.2. Одноканальна СМО з накопичувачем обмеженою ємкості (М/М/1/г)**

##### **1. Опис системи.**

1.1. Система (рис.4.2) містить один обслуговуючий прилад (П), тобто є одноканальною.

1.2. Потік заявок, що надходять у систему, однорідний.

1.3. Тривалість обслуговування заявок в приладі – величина випадкова.

1.4. Перед приладом мається  $r$  місць для заявок, які очікують обслуговування і утворюють чергу, тобто в системі є накопичувач обмеженої ємкості:  $r < \infty$ .

##### **2. Припущення і допущення.**

2.1. Заявки, що надходять в систему, утворюють найпростіший потік з інтенсивністю  $\lambda$ .

2.2. Тривалість обслуговування заявок в приладі розподілена за експоненційним законом з інтенсивністю  $\mu = 1/b$ , де  $b$  – середня тривалість обслуговування заявок в приладі.

2.3. Дисципліна буферизації – з втратами: заявка, що надійшла в систему і застала накопичувач заповненим, втрачається.

2.4. Дисципліна обслуговування – в порядку надходження за правилом «першим прийшов – першим обслужений» (FIFO).

У СМО з накопичувачем обмеженою ємкості завжди існує стаціонарний режим, оскільки довжина черги не буде рости до нескінченності навіть при великих значеннях навантаження.

##### **3. Кодування станів марківського процесу.**

Як параметр, що описує стан марківського процесу, будемо

розглядати кількість заявок  $k$ , що знаходяться в СМО (в приладі і в накопичувачі). Тоді марківський процес в будь-який момент часу може знаходитися в одному з наступних  $(r + 2)$ -х станах:

$E_0: k = 0$  – в системі немає жодної заявки;

$E_1: k = 2$  – в системі знаходиться 1 заявка на обслуговуванні в приладі;

$E_2: k = 2$  – в системі знаходяться 2 заявки: одна – на обслуговуванні в приладі і друга очікує в накопичувачі;

...

$E_{r+1}: k = r + 1$  – в системі знаходяться  $(r + 1)$  заявка: одна – на обслуговуванні в приладі і  $r$  – в накопичувачі.

**4. Розмічений граф переходів випадкового процесу представлений на рис.4.3.**

В один і той же момент часу в системі може відбутися тільки одна подія:

- надходження заявки з інтенсивністю  $\lambda$ , що відповідає збільшенню на одиницю числа заявок в системі і переходу випадкового процесу в стан з номером на одиницю більше;

- завершення обслуговування заявки в приладі з інтенсивністю  $\mu$ , що відповідає зменшенню числа заявок в системі і переходу випадкового процесу в стан з номером на одиницю менше.

#### **5. Система рівнянь.**

Складемо по графу переходів систему рівнянь для визначення стаціонарних ймовірностей:

$$(4.3)$$

Використовуючи метод математичної індукції можна показати, що

$$p_k = y^k p_0 \quad (k = ), \quad (4.4)$$

де  $y = \lambda b$  – навантаження системи.

Підставляючи отриманий вираз в останнє рівняння системи лінійних алгебраїчних рівнянь, знайдемо ймовірність простою системи залежно від навантаження:

$$p_0 = = . \quad (4.5)$$

Тоді стаціонарні ймовірності станів:  $p_k$  ( $k =$ ) :

$$p_k = \dots \quad (4.6)$$

## 5. Розрахунок характеристик СМО.

Характеристики СМО при знайдених значеннях стаціонарних ймовірностей станів випадкового процесу можуть бути розраховані за наступними формулами:

- 1) навантаження  $y = \lambda / \mu = \lambda b$ ;
- 2) завантаження  $\rho = 1 - p_0$ ;
- 3) коефіцієнт простою системи  $\eta = p_0 = 1 - \rho$ ;
- 4) середнє число заявок в черзі  $l = \dots$ ;
- 5) середнє число заявок в системі  $m = l + \rho$ ;
- 6) ймовірність втрати заявок  $\pi = p_{r+1}$ ;
- 7) продуктивність системи (інтенсивність потоку обслужених заявок)  $\lambda' = \lambda (1 - \pi)$ ;
- 8) інтенсивність потоку втрачених заявок  $\lambda'' = \lambda \pi$ ;
- 9) середній час очікування заявок  $w = l / \lambda'$ ;
- 10) середній час перебування заявок  $u = m / \lambda' = w + b$ .

### 4.2.3. Одноканальна СМО з накопичувачем необмеженої ємкості (М/М/1)

#### 1. Опис системи (рис.4.4).

- 1.1. Система – одноканальна – з одним обслуговуючим приладом.
- 1.2. Потік заявок однорідний.
- 1.3. У приладі відбувається затримка заявок, що надходять у систему, на деякий випадковий час.
- 1.4. У системі є накопичувач необмеженої ємкості:  $r = \infty$ , тобто будь-яка заявка, що надійшла в систему, знайде місце для очікування в черзі і не буде втрачена.

#### 2. Припущення і допущення.

- 2.1. Заявки, що надійшли в систему, утворюють найпростіший потік з інтенсивністю  $\lambda$ .
- 2.2. Тривалість обслуговування заявок в приладі розподілена за експоненційним законом з інтенсивністю  $\mu = 1/b$ , де  $b$  – середня тривалість обслуговування заявок в приладі.
- 2.3. Дисципліна буферизації відсутня, оскільки накопичувач має необмежену ємкість.
- 2.4. Дисципліна обслуговування – в порядку надходження за правилом «першим прийшов – першим обслужений» (FIFO).
- 2.5. Навантаження системи співпадає з завантаженням, причому виконується умова:  $y = \rho < 1$ , тобто система працює в сталому режимі без перевантажень. При  $y > 1$ , на відміну від попередніх моделей, в СМО

встановлюється режим перевантажень.

### 3. Кодування станів марківського процесу.

Як параметр, що описує стан марківського процесу, як і в попередньому прикладі, будемо розглядати кількість заявок  $k$ , що знаходяться в СМО (в приладі і в накопичувачі). Оскільки в системі в довільний момент часу може знаходитися будь-яке як завгодно велике число заявок, то кількість станів марковського процесу дорівнює нескінченності:

$E_0$ :  $k = 0$  – в системі немає жодної заявки;

$E_1$ :  $k = 1$  – в системі знаходиться 1 заявка (на обслуговуванні в приладі);

$E_2$ :  $k = 2$  – в системі знаходяться 2 заявки (одна – на обслуговуванні в приладі і друга очікує в накопичувачі);

...

$E_k$ :  $k$  – в системі знаходяться  $k$  заявок (одна – на обслуговуванні в приладі і  $(k - 1)$  – у накопичувачі)...

### 4. Розмічений граф переходів випадкового процесу представлений на рис.4.5.

В один і той же момент часу може відбуватися тільки одна подія: надходження заявки в систему з інтенсивністю  $\lambda$  або завершення обслуговування заявки з інтенсивністю  $\mu$ . Розмічений граф переходів містить нескінченне число станів.

### 5. Система рівнянь.

Не формуючи матрицю інтенсивностей переходів, складемо по графу переходів систему рівнянь для визначення стаціонарних ймовірностей:

(4.7)

Незважаючи на те, що система містить нескінченне число рівнянь і, відповідно, нескінченне число невідомих, неважко методом математичної індукції отримати аналітичні рішення в явному вигляді для розрахунку ймовірностей станів одноканальної експоненційної СМО з однорідним потоком заявок і накопичувачем необмеженої ємності при умові, що навантаження системи  $\rho < 1$ :

$$p_k = y^k (1 - y) = \rho^k (1 - \rho) \quad (k = 0, 1, 2, \dots), \quad (4.8)$$

де  $\rho = \lambda / \mu$  – завантаження системи, що збігається з навантаженням, причому  $\rho < 1$ , що гарантує існування сталого режиму в системі.

Таким чином, ймовірність знаходження марківського процесу в стані

$E_k$  або, що те ж саме, ймовірність того, що в довільний момент часу в системі знаходиться  $k$  заявок, розподілена по геометричному закону з параметром, рівним завантаженню (навантаженню) системи.

### 6. Розрахунок характеристик СМО.

Для розрахунку характеристик СМО можна скористатися наступними математичними залежностями:

- 1) навантаження  $y = \lambda / \mu = \lambda b$ ;
- 2) завантаження  $\rho = 1 - p_0 = \lambda b$  і збігається з навантаженням;
- 3) коефіцієнт простою системи  $\eta = p_0 = 1 - \rho$ ;
- 4) середнє число заявок в черзі  $l =$  ;
- 5) середнє число заявок в системі  $m =$  ;
- 6) ймовірність втрати заявок  $\pi = 0$ ;
- 7) продуктивність системи за відсутності втрат збігається з інтенсивністю надходження заявок в систему:  $\lambda' = \lambda$ ;
- 8) інтенсивність втрачених заявок  $\lambda'' = 0$ ;
- 9) середній час очікування заявок  $w =$  ;
- 10) середній час перебування заявок  $u = w + b$  або  $u =$  ; .

Отримані вирази збігаються з формулами для розрахунку характеристик експоненційою СМО.

#### 4.2.4 . Одноканальна СМО з неоднорідним потоком заявок і відносними пріоритетами

Розглянемо одноканальну СМО, в яку надходить неоднорідний потік заявок. Заявки, що очікують обслуговування, рознесені по різних накопичувачах обмеженої ємності. Між заявками різних класів встановлені відносні пріоритети (ВП) , що означає, що всякий раз з накопичувачів на обслуговування вибирається заявка з найвищим пріоритетом. При цьому при надходженні в систему високопріоритетні заявки обслуговування фонові не переривається. Заявка, що надійшла при заповнених накопичувачах, втрачається.

1. Опис системи ( рис. 4.6) .

1.1 . Система одноканальна.

1.2 . Вхідний потік заявок – неоднорідний: у систему надходить два класи заявок.

1.3 . Накопичувачі для заявок кожного класу – обмеженої місткості:

$$r_1 = r_2 = 1.$$

1.4. Дисципліна буферизації – без витіснення заявок: якщо при вступі до системи заявки будь-якого класу відповідний накопичувач заповнений до кінця, то заявка втрачається.

1.4. Дисципліна обслуговування – з відносними пріоритетами: заявки першого класу мають пріоритет по відношенню до заявок другого класу.

### 2. Припущення і допущення.

2.1. Заявки двох класів, що надходять в систему, утворюють найпростіші потоки з інтенсивностями  $\lambda_1$ , і  $\lambda_2$  відповідно.

2.2. Тривалості обслуговування заявок кожного класу розподілені за експоненційним законом з інтенсивностями  $\mu_1 = 1 / b_1$  и  $\mu_2 = 1 / b_2$ , де  $b_1$  і  $b_2$  – середні тривалості обслуговування заявок класу 1 і 2 відповідно.

У СМО завжди існує стаціонарний режим, так як не може бути нескінченних черг.

### **3 . Кодування станів випадкового процесу .**

Для опису станів марківського процесу будемо використовувати розподіл заявок між приладом і накопичувачами. Закодуємо стани наступним чином:  $(\Pi/\mathcal{C}_1, \mathcal{C}_2)$ , де  $\Pi = \{0, 1, 2\}$  – стан обслуговуючого приладу, що задається класом заявки, яка знаходиться на обслуговуванні («0» – прилад вільний; «1» або «2» – на обслуговуванні в приладі знаходиться заявка класу 1 або 2 відповідно);  $\mathcal{C}_1, \mathcal{C}_2 = \{0, 1\}$  – стан накопичувачів 1 і 2 відповідно («0» – означає відсутність заявки в накопичувачі, «1» – означає наявність однієї заявки в накопичувачі відповідного класу).

При обраному способі кодування система може перебувати в наступних станах:

$E_0$ : (0/0,0) – в системі немає жодної заявки;

$E_1$ : (1/0,0) – на обслуговуванні в приладі знаходиться заявка класу 1;

$E_2$ : (2/0,0) – на обслуговуванні в приладі знаходиться заявка класу 2;

$E_3$ : (1/1,0) – на обслуговуванні знаходиться заявка класу 1 і одна заявка класу 1 очікує обслуговування в першому накопичувачі;

$E_4$ : (1/0,1) – на обслуговуванні знаходиться заявка класу 1 і одна заявка класу 2 очікує обслуговування відповідно в другому накопичувачі;

$E_5$ : (2/1,0) – на обслуговуванні знаходиться заявка класу 2 і одна заявка класу 1 очікує обслуговування в першому накопичувачі;

$E_6$ : (2/0,1) – на обслуговуванні знаходиться заявка класу 2 і одна заявка класу 2 очікує обслуговування в другому накопичувачі;

$E_7$ : (1/1,1) – на обслуговуванні знаходиться заявка класу 1, і по одній заявці кожного класу очікують обслуговування у відповідних накопичувачах;

$E_8$ : (2/1,1) – на обслуговуванні знаходиться заявка класу 2, і по одній заявці кожного класу очікують обслуговування у відповідних накопичувачах.

При кодуванні випадкових процесів можуть бути застосовані різні способи кодування.

У розглянутому прикладі стани випадкового процесу замість представленого вище способу можна закодувати, наприклад, таким чином:  $(\Pi, \mathcal{C})$ , де  $\Pi = \{0, 1, 2\}$  – стан обслуговуючого приладу, що задається класом заявки, що знаходиться на обслуговуванні («0» – прилад вільний; «1» або «2» – на обслуговуванні в приладі знаходиться заявка класу 1 або 2 відповідно);  $\mathcal{C} = \{0, 1, 2, 3\}$  – стан накопичувачів 1 і 2 відповідно («0» – означає відсутність заявок в обох накопичувачах; «1» – наявність в



першому накопичувачі заявки класу 1; «2» – наявність в другому накопичувачі заявки класу 2; «3» – наявність в першому і другому накопичувачах по одній заявці відповідно класу 1 і 2).

Представлені способи кодування не застосовні, якщо для заявок обох класів використовується загальний накопичувач ємкістю  $r = 2$ . У цьому випадку кількість станів випадкового процесу збільшиться, оскільки в накопичувачі можуть перебувати 2 заявки одного і того ж класу і стан накопичувача може бути представлено наступним чином:  $Ч = \{0, 1, 2, 11, 12, 22\}$ , де «0» – означає відсутність заявок в накопичувачі; «1» – наявність у накопичувачі тільки однієї заявки класу 1; «2» – наявність у накопичувачі заявки класу 2; «11» – наявність у накопичувачі двох заявок класу 1; «22» – наявність у накопичувачі двох заявок класу 2 і «12» – наявність в накопичувачі однієї заявки класу 1 і однієї заявки класу 2. Стан «12» не розрізняє, в якій послідовності ці заявки надійшли в систему, що обумовлено наявністю відносного пріоритету між ними – незалежно від моменту надходження на обслуговування першою завжди буде обрана заявка класу 1. У разі безпріоритетного обслуговування, коли заявки різних класів вибираються на обслуговування в порядку надходження, слід ввести ще один стан накопичувача – «21», що означає, що заявка класу 2 поступила в систему раніше заявки класу 1, в той час як стан «12» означає, що в систему раніше надійшла заявка класу 1.

**4. Розмічений граф переходів випадкового процесу представлений на рис. 4.7.**

У кожен момент часу може відбутися тільки одна подія (або надходження заявки якогось класу, або завершення обслуговування заявки, що знаходиться в приладі), оскільки ймовірність появи двох і більше подій в один і той же момент часу дорівнює нулю.

При наявності в накопичувачах заявок першого і другого класів (стану  $E_7$  и  $E_8$ ) після завершення обслуговування деякої заявки в приладі випадковий процес переходить у стан  $E_4$ , що означає, що на обслуговування завжди вибирається високопріоритетні заявка класу 1.

За графом переходів складемо систему рівнянь для визначення стаціонарних ймовірностей::

(4.9)

## **5. Розрахунок характеристик СМО.**

Характеристики обслуговування заявок в СМО з неоднорідним потоком заявок діляться на дві групи:

- характеристики обслуговування заявок кожного класу;
- характеристики обслуговування заявок сумарного потоку.

Розрахунок характеристик обслуговування заявок кожного класу виконується за такими формулами:

- 1) навантаження:  $y_1 = \lambda_1 / \mu_1 = \lambda_1 b_1$ ;  $y_2 = \lambda_2 / \mu_2 = \lambda_2 b_2$ ;
- 2) завантаження, створюване заявками, яке може трактуватися як ймовірність того, що на обслуговуванні в приладі знаходиться заявка класу 1 і 2 відповідно:  $\rho_2 = p_2 + p_5 + p_6 + p_8$ ;
- 3) середнє число заявок в черзі:  
;  $l_2 = p_4 + p_6 + p_7 + p_8$ ;
- 4) середнє число заявок в системі:  
;  
;
- 5) ймовірність втрати заявок:  
;  $\pi_2 = p_4 + p_6 + p_7 + p_8$ ;
- 6) продуктивність по кожному класу заявок (інтенсивність невтрачених заявок):  
 $= \lambda_1 (1 - \pi_1)$ ;  $\lambda_2 = \lambda_2 (1 - \pi_2)$ ;
- 7) середній час очікування заявок:  
 $w_1 = l_1 / \lambda_1$ ;  $w_2 = l_2 / \lambda_2$ ;
- 8) середній час перебування заявок:  
 $u_1 = m_1 / \lambda_1 = w_1 + b$ ;  $u_2 = m_2 / \lambda_2 = w_2 + b$ ;

Розрахунок характеристик обслуговування заявок сумарного потоку виконується за такими формулами:

- 1) сумарне навантаження системи:  $Y = y_1 + y_2$ ;
- 2) завантаження системи:  $R = \rho_1 + \rho_2$ ;
- 3) коефіцієнт простою системи:  $\eta = 1 - R$ ;
- 4) сумарне число заявок у всіх чергах:  $l = l_1 + l_2$ ;
- 5) сумарне число заявок в системі:  $m = m_1 + m_2 = l + R$ ;
- 6) ймовірність втрати заявок:  $\pi = \pi_1 + \pi_2$ ;
- 7) продуктивність системи (інтенсивність сумарного потоку обслугованих заявок):  $\lambda' = \lambda (1 - \pi)$ ;
- 8) середній час очікування заявок сумарного потоку:  
1.  $w = (w_1 + w_2) / \lambda' = l / \lambda'$ ;
- 9) середній час перебування заявок сумарного потоку:  
2.  $u = (u_1 + u_2) / \lambda' = m / \lambda' = w + b$ .

### 4.3. Марківські моделі мереж масового обслуговування

Розглянемо докладно деякі марківські моделі мереж масового обслуговування (MeMO).

#### 4.3.1 Розімкнена експоненційна MeMO з накопичувачами обмеженої ємкості

Розглянемо розімкнену експоненційну MeMO з двома одноканальними вузлами, в яку із зовнішнього середовища надходить найпростіший потік заявок з інтенсивністю  $\lambda_0$  (рис. 4.8). Накопичувачі в обох вузлах мають обмежену ємкість, рівну одиниці:  $r_1 = r_2 = 1$ . Заявка, що надійшла у вузол і застала накопичувач заповненим, втрачається.

Тривалості обслуговування у вузлах розподілені за експоненційним законом з середніми значеннями  $b_1$  и  $b_2$  відповідно. Заявки після обслуговування у вузлі 2 з ймовірністю  $q$  направляються у вузол 1 і ймовірністю  $(1 - q)$  – покидають МеМО.

Оскільки заявки в мережі можуть губитися, розглянута розімкнена МеМо є нелінійною, тобто інтенсивності потоків заявок, що надходять у вузли МеМО, не пов'язані між собою лінійною залежністю і, отже, не можуть бути розраховані шляхом вирішення системи лінійних алгебраїчних рівнянь.

### 1. Опис МеМО (рис. 4.8).

1.1. Мережа масового обслуговування – двовузлова розімкнена.

1.2. Вузли 1 і 2 – одноканальні:  $K_1 = K_2 = 1$ .

1.3. Накопичувачі у вузлах обмеженою ємкості:  $r_1 = r_2 = 1$ .

1.4. Дисципліни буферизації у вузлах – з втратами заявок, якщо накопичувачі заповнені.

1.5. Потік заявок з інтенсивністю  $\lambda_0$ .

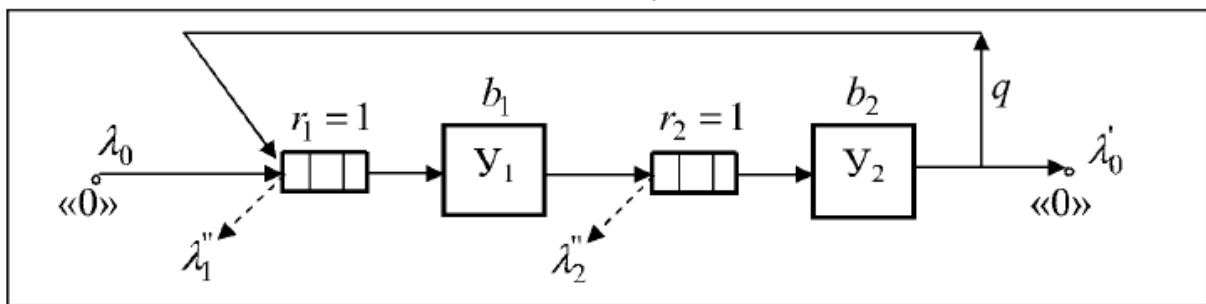


Рис. 4.8. – Двовузлова розімкнена МеМО з втратами

### 2. Припущення і допущення.

2.1. Заявки, що надходять в розімкнену МеМО утворюють найпростіший потік з інтенсивністю  $\lambda_0$ .

2.2. Тривалості обслуговування заявок у вузлах МеМО розподілені за експоненційним законом з параметрами, котрі представляють собою інтенсивності обслуговування:  $\mu_1 = 1/b_1$  і  $\mu_2 = 1/b_2$ .

У розімкнутій МеМО при будь-якому навантаженні існує стаціонарний режим, так як у вузлах мережі не може бути нескінченних черг.

### 3. Кодування станів випадкового процесу.

Для опису станів марківського випадкового процесу будемо використовувати розподіл заявок між вузлами. Закодуємо стану наступним чином:  $(M_1, M_2)$ , де  $M_i = \{0, 1, 2\}$  – кількість заявок у вузлі  $i$  («0» – вузол вільний; «1» – на обслуговуванні у вузлі знаходиться одна заявка; «2» – у вузлі знаходяться дві заявки – одна на обслуговування і друга в накопичувачі).

При обраному способі кодування система може перебувати в наступних станах:

$E_0(0,0)$  – в МеМО немає жодної заявки;

$E_1(1,0)$  – у вузлі 1 знаходиться одна заявка;

- $E_2(2,0)$  – у вузлі 1 знаходяться дві заявки;
- $E_3(0,1)$  – у вузлі 2 знаходиться одна заявка;
- $E_4(1,1)$  – у вузлі 1 і 2 знаходиться по одній заявці;
- $E_5(2,1)$  – дві заявки знаходяться у вузлі 1 і одна – у вузлі 2;
- $E_6(0,2)$  – у вузлі 2 знаходяться дві заявки;
- $E_7(1,2)$  – дві заявки знаходяться у вузлі 2 і одна – у вузлі 1;
- $E_8(2,2)$  – у вузлі 1 і 2 знаходяться по дві заявки.

**4. Розмічений граф переходів випадкового процесу (рис.5.17).**

Побудуємо граф переходів, вважаючи, що в кожен момент часу може відбутися тільки одна подія (надходження заявки в МеМО або завершення обслуговування заявки в одному з вузлів), оскільки ймовірність появи двох і більше подій в один і той же момент часу дорівнює нулю.

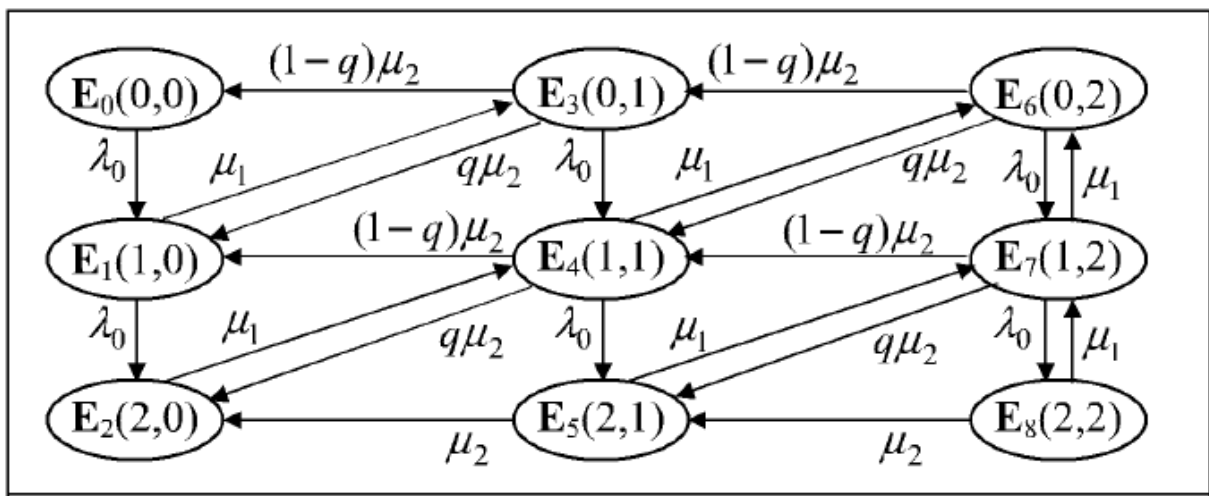


Рис. 4.9. – Граф переходів розімкненої МеМО з втратами

Слід звернути увагу на переходи із станів  $E_3(0,1)$ ,  $E_4(1,1)$ ,  $E_6(0,2)$  і  $E_7(1,2)$ , зумовлені завершенням обслуговування заявки у вузлі 2 з інтенсивністю  $\mu_2$ . У цих випадках з імовірністю  $q$  заявка може повернутися у вузол 1 і з імовірністю  $(1-q)$  – покинути МеМО, тоді інтенсивності відповідних переходів будуть рівні  $q\mu_2$  і  $(1-q)\mu_2$ . Якщо ж випадковий процес знаходиться в станах  $E_5(2,1)$  і  $E_8(2,2)$ , то завершення обслуговування заявки у вузлі 2 призводить до переходу відповідно в стани  $E_2(2,0)$  і  $E_5(2,1)$  з інтенсивністю  $\mu_2$ , що відповідає виходу заявки з МеМО з імовірністю  $(1-q)$  і втрати заявки, яка з імовірністю  $q$  буде спрямована у вузол 1, оскільки в останньому немає місця в накопичувачі. Аналогічно, якщо випадковий процес знаходиться в станах  $E_7(1,2)$  і  $E_8(2,2)$ , то завершення обслуговування заявки у вузлі 1 призводить до переходу відповідно в стани  $E_6(0,2)$  і  $E_7(1,2)$  з інтенсивністю  $\mu_1$  що відповідає втраті заявки, оскільки накопичувач вузла 1 заповнений.

**5. Розрахунок характеристик МеМО.**

Не складаючи матрицю інтенсивностей переходів і не виписуючи систему рівнянь для визначення стаціонарних ймовірностей, отримаємо

математичні вирази для визначення вузлових і мережових характеристик розімкнутої МеМО з втратами при відомих значеннях стаціонарних ймовірностей станів  $p_i$  ( $i = 0, 1, \dots, 8$ ).

Зауважимо, що МеМО з втратами відноситься до класу нелінійних мережних моделей, розрахунок характеристик яких пов'язаний з певними проблемами, зокрема, з необхідністю детального аналізу потоків заявок і з неможливістю застосування в ряді випадків фундаментальних співвідношень для розрахунку мережних характеристик. Крім того, процес формування математичних залежностей для кожної конкретної нелінійної МеМО може істотно відрізнятись.

У зв'язку з цим, нижче досить докладно розглядається процес отримання математичних виразів для розрахунку вузлових і мережних характеристик нелінійної розімкнутої МеМО, представленої на рис.4.8. Вузлові характеристики МеМО розраховуються в такій послідовності:

1) завантаження вузлів визначаються як суми ймовірностей станів, в яких відповідний вузол зайнятий обслуговуванням заявок:

$$\rho_1 = p_1 + p_2 + p_4 + p_5 + p_7 + p_8; \quad \rho_2 = p_3 + p_4 + p_5 + p_6 + p_7 + p_8;$$

2) коефіцієнти простою вузлів:  $\eta_1 = 1 - \rho_1; \quad \eta_2 = 1 - \rho_2;$

3) середнє число заявок в чергах:

$$l_1 = p_2 + p_5 + p_8; \quad l_2 = p_6 + p_7 + p_8;$$

4) середнє число заявок у вузлах:

$$m_1 = p_1 + p_4 + p_7 + 2(p_2 + p_5 + p_8) = l_1 + \rho_1;$$

$$m_2 = p_3 + p_4 + p_5 + 2(p_6 + p_7 + p_8) = l_2 + \rho_2;$$

5) продуктивності вузлів (інтенсивність обслугованих заявок на виході вузлів):

6) ймовірності втрати заявок у вузлах МеМО можуть бути розраховані з урахуванням того, що  $K_1 = K_2 = 1$ :

;

В цих виразах:  $y_1 = \lambda_1 b_1$  и  $y_2 = \lambda_2 b_2$  – створювані у вузлах навантаження, де  $\lambda_1$  і  $\lambda_2$ , – інтенсивності надходження заявок у вузли 1 і 2 МеМО, для розрахунку яких необхідно виконати аналіз потоків у розглянутій МеМО; Інтенсивність  $\lambda_1$  складається (див. рис.4.8) з інтенсивності  $\lambda_0$  надходження заявок із зовнішнього джерела і інтенсивності потоку заявок, які повертаються з ймовірністю  $q$  у вузол 1 після обслуговування у вузлі 2:  $\lambda_1 = \lambda_0 +$  , де – розрахована раніше інтенсивність потоку заявок, що виходять з вузла 2 (продуктивність вузла 2);

Аналогічно, з рис. 4.8 можна бачити, що інтенсивність  $\lambda_2$  надходячих у вузол 2 заявок являє собою інтенсивність потоку виходячих з вузла 1 заявок (продуктивність вузла 1):  $\lambda_2 =$ ;

Остаточнo, після деяких перетворень вирази для розрахунку ймовірностей втрати заявок у вузлах МеМО приймуть вигляд:

7) середній час очікування заявок у вузлах розраховується за

формулами Літтла з урахуванням тільки обслужених заявок:

$$w_1 = l_1/\lambda; \quad w_2 = l_2/\lambda;$$

8) аналогічно, середній час перебування заявок у вузлах:

$$u_1 = m_1/w_1 + b; \quad u_2 = m_2/w_2 + b;$$

Для розрахунку мережних характеристик МеМО можуть використовуватися такі формули:

1) сумарне завантаження вузлів МеМО, що характеризує середнє число одночасно працюючих вузлів в мережі:  $\rho = \rho_1 + \rho_2$ ;

2) сумарне число заявок в чергах:  $L = l_1 + l_2$ ;

3) сумарне число заявок у вузлах:  $M = m_1 + m_2 = L + \rho$ ;

4) продуктивність МеМО (інтенсивність обслужених заявок на виході мережі):

5) ймовірність втрати заявок в мережі:  $\pi$ ; слід звернути увагу, що ймовірність втрати заявок в мережі визначається як частка втрачених заявок по відношенню до заявок, що надійшли в МеМО, в той час як ймовірності втрати  $\pi_1$  і  $\pi_2$  заявок у вузлах МеМО визначається як частка втрачених заявок по відношенню до всіх заявок, що надійшли в конкретний вузол, число яких враховує і ті, що надійшли в МеМО і за час знаходження в мережі можуть потрапити в даний вузол кілька разів.

Математичні залежності для розрахунку сумарного часу очікування заявок і часу перебування заявок в МеМО не можуть бути отримані в загальному вигляді у випадку нелінійних МеМО з втратами.

## 5. ОСНОВИ ІМІТАЦІЙНОГО МОДЕЛЮВАННЯ

### 5.1. Поняття імітаційного моделювання

**Статистичне моделювання** – метод дослідження складних систем, заснований на описі процесів функціонування окремих елементів в їх взаємозв'язку з метою отримання безлічі окремих результатів, що підлягають обробці методами математичної статистики для отримання кінцевих результатів. В основі статистичного моделювання лежить метод статистичних випробувань – метод Монте-Карло.

**Імітаційна модель** – універсальний засіб дослідження складних систем, що представляє собою логіко-алгоритмічний опис поведінки окремих елементів системи та правил їх взаємодії, що відображають послідовність подій, що виникають в системі, яка моделюється.

Якщо статистичне моделювання виконується з використанням імітаційної моделі, то таке моделювання називається імітаційним.

Поняття «статистичне та імітаційне моделювання» часто розглядають як синоніми. Проте слід мати на увазі, що статистичне моделювання не обов'язково є імітаційним. Наприклад, обчислення певного інтеграла методом Монте-Карло шляхом визначення підінтегральної площі на основі безлічі статистичних випробувань, відноситься до статистичного моделювання, але не може називатися імітаційним.

Найбільш широке застосування імітаційне моделювання отримало при дослідженні складних систем з дискретним характером функціонування, в тому числі моделей масового обслуговування. Для опису процесів функціонування таких систем зазвичай використовуються часові діаграми.

**Часова діаграма** – графічне представлення послідовності подій, що відбуваються в системі. Для побудови часових діаграм необхідно досить чітко уявляти взаємозв'язок подій усередині системи. Ступінь деталізації при складанні діаграм залежить від властивостей модельованої системи і від цілей моделювання.

Оскільки функціонування будь-якої системи досить повно відображається у вигляді часової діаграми, імітаційне моделювання можна розглядати як процес реалізації діаграми функціонування досліджуваної системи на основі відомостей про характер функціонування окремих елементів і їх взаємозв'язку.

Імітаційне моделювання зазвичай проводиться на ПК відповідно до програми, що реалізує заданий конкретний логіко-алгоритмічний опис. При цьому кілька годин, тижнів або років роботи досліджуваної системи можуть бути промодельовані на ПК за кілька хвилин. У більшості випадків модель є не точним аналогом системи, а швидше її символічним відображенням. Однак така модель дозволяє проводити вимірювання, які неможливо провести будь-яким іншим способом.

Імітаційне моделювання забезпечує можливість випробування, оцінки та проведення експериментів з досліджуваною системою без будь-яких безпосередніх впливів на неї.

Першим кроком при аналізі будь-якої конкретної системи є виділення

елементів, і формулювання логічних правил, керуючих взаємодією цих елементів. Отриманий в результаті цього опис називається моделлю системи. Модель зазвичай включає в себе ті аспекти системи, які становлять інтерес або потребують дослідження.

Оскільки метою побудови будь-якої моделі є дослідження характеристик модельованої системи, в імітаційну модель повинні бути включені засоби збору та обробки статистичної інформації по всім характеристикам, що цікавлять, засновані на методах математичної статистики.

### 5.1.1. Принципи організації імітаційного моделювання

Розглянемо принципи імітаційного моделювання на прикладі найпростішої базової моделі у вигляді одноканальної системи масового обслуговування з однорідним потоком заявок (рис.5.1), в яку надходить випадковий потік заявок з інтервалами між сусідніми заявками, розподіленими за законом  $A(\tau)$ , а тривалість обслуговування заявок в приладі розподілена за законом  $B(\tau)$ .

Процес функціонування такої системи може бути представлений у вигляді часових діаграм, на основі яких можуть бути виміряні і розраховані характеристики обслуговування заявок. Оскільки процеси надходження та обслуговування заявок в системі носять випадковий характер, то для побудови діаграм необхідно мати генератори випадкових чисел.

Покладемо, що в нашому розпорядженні є генератори випадкових чисел, що формують значення відповідних випадкових величин із заданими законами розподілів  $A(\tau)$  і  $B(\tau)$ . Тоді можна побудувати часові діаграми, що відображають процес функціонування даної системи.

На рис.5.2 представлені чотири діаграми, що відображають:

1) «**процес надходження заявок**» у вигляді моментів надходження заявок в систему, формованих за правилом:  $t_j = t_{j-1} + \tau_{ai} (t_0=0)$ , де  $\tau_{ai} (i=1,2,...)$  – інтервали між заявками, що надходять в систему, значення яких генеруються за допомогою генератора випадкових величин  $A(\tau)$ ;

2) «**процес обслуговування в приладі**», представлений у вигляді тривалостей обслуговування  $\tau_{bi}$ , які генеруються за допомогою генератора випадкових величин  $B(\tau)$ , і моментів завершення обслуговування  $t'_i$  заявок в приладі, що визначаються за наступним правилом:

$t'_i = t_i + \tau_{bi}$ , якщо на момент надходження  $i$ -ї заявки обслуговуючий прилад був вільний;

$t'_i = t_{i-1} + \tau_{bi}$ , якщо на момент надходження  $i$ -ї заявки обслуговуючий прилад був зайнятий обслуговуванням попередньої заявки ( $i=1,2,...; t'_0=0$ );

3) «**модельний або реальний час**», що показує дискретну



(стрибкоподібну) зміну часу в реальній системі, кожен момент якого відповідає одній з таких подій: надходження заявки в систему або завершення обслуговування заявки в приладі; в ці моменти часу відбувається зміна стану системи, що описується числом заявок, що знаходяться в системі;

4) «число заявок в системі», що описує стан дискретної системи і змінюється за правилом: збільшення на 1 в момент надходження заявки в систему і зменшення на 1 в момент завершення обслуговування.

При дотриманні обраного часового масштабу представлені діаграми дозволяють шляхом вимірювання визначити значення ймовірностно-часових характеристик функціонування модельованої системи, зокрема, як показано на другій діаграмі, час знаходження (перебування) кожної заявки в системі:  $\tau_{ii}$  ( $i = 1, 2, \dots$ ).

Очевидно, що час перебування заявок в системі – величина випадкова. У простому випадку, застосовуючи методи математичної статистики, можна розрахувати два перших моменти розподілу часу перебування:

- математичне очікування:

(5.1)

- другий початковий момент:

(5.2)

де  $N$  – кількість значень часу перебування заявок, отриманих на діаграмі, тобто кількість заявок, відображених на діаграмі які пройшли через систему і покинули її.

Звідси легко можуть бути отримані значення дисперсії, середньоквадратичного відхилення і коефіцієнта варіації часу перебування заявок в системі.

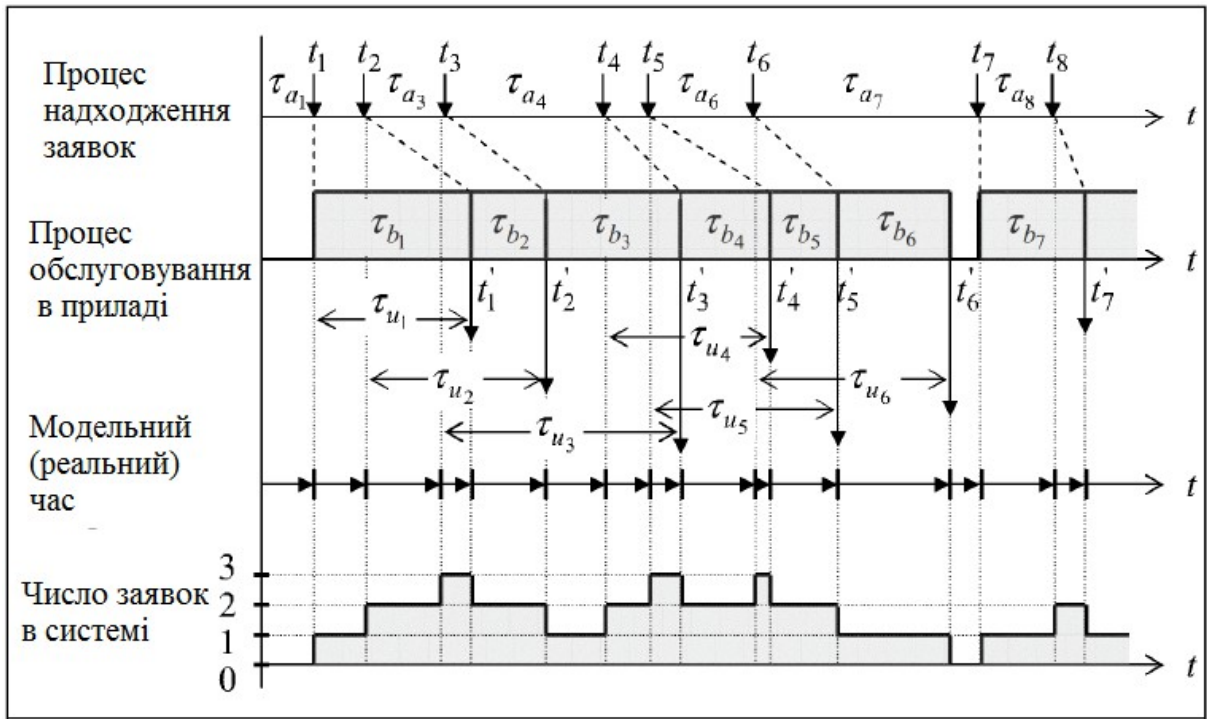


Рис. 5.2. – Діаграми функціонування одноканальної СМО

На основі отриманих за допомогою часових діаграм значень часу перебування заявок в системі можна побудувати гістограму функції або щільності розподілу часу перебування.

Точність отриманих числових моментів розподілу і якість гістограм істотно залежить від кількості значень  $N$  часу перебування заявок, на основі яких вони розраховуються: чим більше  $N$ , тим точніше результати розрахунку. Значення  $N$  може становити від декількох тисяч до десятків мільйонів. Конкретне значення  $N$  залежить від багатьох факторів, що впливають на швидкість збіжності результатів до істинного значення, основними серед яких при моделюванні систем і мереж масового обслуговування є закони розподілів інтервалів між заявками, що надходять, і тривалостей обслуговування, завантаження системи, складність моделі, кількість класів заявок і т.д.

Ясно, що побудова вручну таких часових діаграм з тисячами і більше заявками, що проходять через систему, нереальна. У той же час, використання ПК для реалізації часових діаграм дозволяє істотно прискорити процеси моделювання та отримання кінцевого результату. Тому імітаційне моделювання можна розглядати як процес реалізації діаграми функціонування досліджуваної системи.

Таким чином, імітаційна модель являє собою алгоритм реалізації часової діаграми функціонування досліджуваної системи. Наявність вбудованих в більшість алгоритмічних мов генераторів випадкових чисел значно спрощує процес реалізації імітаційної моделі на ПК. Однак при цьому залишається ряд проблем, що вимагають свого рішення. Одна з них полягає в принципі реалізації часової діаграми і, пов'язаною з нею,

проблемою організації служби часу в імітаційній моделі.

У простому випадку часова діаграма може бути реалізована наступним чином: спочатку формуються моменти надходження всіх заявок в систему, а потім для кожної заявки визначаються тривалості обслуговування в приладі і формуються моменти завершення обслуговування (виходу заявок з системи). Очевидно, що такий підхід неприйнятний, оскільки навіть для нашої дуже простої системи доведеться зберігати в пам'яті ПК одночасно мільйони значень моментів надходження і завершення обслуговування заявок, а також інших змінних, причому із збільшенням кількості класів заявок та кількості обслуговуючих приладів це число збільшиться багаторазово.

Другий підхід, який може бути запропонований для реалізації часової діаграми, – покрокова побудова діаграми. Для цього слід сформувати змінну для модельного часу і вибрати крок її зміни. У кожен момент часу необхідно перевіряти, яка подія (надходження в систему або завершення обслуговування заявки) відбулася в системі за попередній інтервал  $\Delta t$ .

Цей підхід значно скорочує потребу в пам'яті, оскільки в цьому випадку в кожен момент часу необхідно зберігати в пам'яті ПК значення параметрів (моментів надходження і завершення обслуговування) тільки тих заявок, які знаходяться в системі на даний момент часу.

Недоліки такого підходу очевидні. По-перше, проблематичним є вибір довжини інтервалу  $\Delta t$ . З одного боку, інтервал  $\Delta t$  повинен бути якомога менше для зменшення методичної похибки моделювання, з іншого боку, інтервал  $\Delta t$  повинен бути якомога більше для зменшення часу моделювання.

Найбільш ефективним підходом визнаний підхід із змінним кроком прорахунку модельного часу, який реалізується відповідно до принципу «до найближчого події». **Принцип «просування модельного часу до найближчої події» полягає в наступному.** По всіх процесах, що паралельно протікають в досліджуваній системі, в кожен момент часу формуються моменти настання «найближчого події в майбутньому». Потім модельний час просувається до моменту настання найближчої з усіх можливих подій. Залежно від того, яка подія виявилася найближчою, виконуються ті чи інші дії. Якщо найближчою подією є надходження заявки в систему, то виконуються дії, пов'язані із заняттям приладу за умови, що він вільний, і занесення заявки в чергу, якщо прилад зайнятий. Якщо ж найближчою подією є завершення обслуговування заявки в приладі, то виконуються дії, пов'язані зі звільненням приладу і вибором на обслуговування нової заявки з черги, якщо остання не порожня. Потім формується новий момент настання цієї ж події. На третій діаграмі «Модельний (реальний) час» (рис.5.2) просування часу відповідно до цього принципу показано у вигляді стрілок.

Для того щоб забезпечити правильну часову послідовність подій в імітаційній моделі, використовуються системні годинники, що зберігають значення поточного модельного часу. Зміна значення модельного часу

здійснюється відповідно до принципу «перерахунку часу до найближчої події». Наприклад, якщо поточне значення модельного часу дорівнює 25, а чергові події повинні настати в моменти часу 31, 44 і 56, то значення модельного часу збільшується відразу на 6 одиниць і «просувається» до значення 31. Одиниці часу в моделі не обов'язково повинні бути конкретними одиницями часу, такими як секунда або година. Основною одиницею часу в моделі можна вибрати будь-яку одиницю, яка дозволить отримати необхідну точність моделювання. Важливо пам'ятати, одиниці часу вибираються виходячи з вимог користувача до точності моделювання. Яка б одиниця не була обрана, наприклад мілісекунда або одна десята години, вона повинна незмінно використовуватися у всій моделі.

Крім розглянутої служби часу в імітаційній моделі необхідно реалізувати процедури, пов'язані з формуванням потоків заявок та імітацією обслуговування, з організацією черг заявок, з організацією збору та статистичної обробки результатів моделювання.

Таким чином, імітаційне моделювання дискретних систем зі стахостичним характером функціонування, таких як системи та мережі масового обслуговування, припускає використання ряду типових процедур, що забезпечують реалізацію відповідних імітаційних моделей. До таких процедур, в першу чергу, відносяться наступні процедури:

- 1) генерування випадкових величин:
  - рівномірно розподілених;
  - із заданим законом розподілу;
- 2) формування потоків заявок і імітація обслуговування;
- 3) організація черг заявок;
- 4) організація служби часу;
- 5) збір та статистична обробка результатів моделювання.

## **6. СИНТЕЗ СТРУКТУРИ КОМП'ЮТЕРНОЇ МЕРЕЖІ ДЛЯ ЗАБЕЗПЕЧЕННЯ ЇЇ ДОСТУПНОСТІ**

При синтезі структури мережі резервування і ремонтпридатність використані як науково обгрунтовані методи підвищення доступності ресурсів і послуг проектованої комп'ютерної мережі.

Доступність є якісною характеристикою функціонування складного технічного об'єкта, такого як комп'ютерна мережа. Оцінювати доступність пропонується кількісними показниками надійності.

Надійність – це властивість об'єкта зберігати в часі в установлених межах значення всіх параметрів, що характеризують здатність виконувати необхідні функції в заданих режимах і умовах застосування, технічного обслуговування, ремонтів, зберігання і транспортування.

Надійність об'єктів закладається при проектуванні, реалізується при виготовленні і витрачається при експлуатації.

Тому потрібно розглядати схемні методи підвищення доступності за критерієм показників надійності на першому етапі життєвого циклу комп'ютерної мережі – проектуванні.

Засобами підвищення працездатності і продуктивності каналів зв'язку є альтернативні шляхи прямування трафіку, надлишкові по відношенню до єдиного основному шляху.

Працездатність кожного фізичного каналу локальної мережі цілком визначається характеристиками працездатності його компонентів (порту комунікаційного пристрою, роз'єму, середовища передачі, точок з'єднання роз'ємів з середовищем, кросових з'єднань).

Працездатність каналу підвищується за рахунок застосування надлишкових фізичних з'єднань, коли канал дублюється одним або декількома іншими.

Підвищення рівня зв'язності мережі, тобто створення альтернативних маршрутів, – це засіб підвищення експлуатаційних характеристик мережі.

Неминучі додаткові витрати створення альтернативних маршрутів можуть виявитися менше, ніж при підвищенні продуктивності за рахунок застосування більш швидкісної технології каналу, а для підвищення працездатності введення надлишкових з'єднань найчастіше стає єдиним практично здійсненним способом досягнення мети.

Альтернативні з'єднання в мережах використовуються двома способами:

1. В режимі резервування, коли одне з них функціонує, а інші знаходяться в "гарячому" резерві для заміни з'єднання, що відмовило.

2. В режимі балансу навантаження; при цьому дані передаються паралельно по всіх альтернативних з'єднаннях.

Альтернативні з'єднання в режимі резервування покращують одну характеристику мережі – працездатність.

Альтернативні з'єднання в режимі балансу навантаження дозволяють підвищити продуктивність і працездатність.

У локальних мережах, де технології та обладнання реалізують тільки функції першого і другого рівня моделі OSI/ISO базові протоколи підтримують тільки деревовидні топології зв'язків.

Для організації альтернативних каналів потрібні протоколи і технології, що виходять за рамки базових, до яких відноситься Ethernet або алгоритм прозорого моста (*Transparent Bridge*), описаний в стандарті 802.1D і вживаний в сучасних комутаторах другого рівня.

Об'єднання декількох фізичних каналів можливо з допомогою механізмів агрегування каналів *Link Aggregation* за специфікацією IEEE 802.3ad.

Підтримка надлишкових зв'язків комутаторами другого рівня знижує вартість мережі, так як багато ділянок мережі будуються без залучення відносно дорогих маршрутизаторів (комутаторів третього рівня) і підвищується швидкість реакції на обрив зв'язку.

При агрегуванні фізичних каналів всі надлишкові зв'язки залишаються в робочому стані, а наявний трафік розподіляється між ними для розділення навантаження.

При відмові однієї з складових логічного каналу, трафік розподіляється між рештою каналів.

Відповідно до ГОСТ 27.002-89 резервуванням називається застосування додаткових засобів та ( або) можливостей з метою збереження працездатного стану об'єкта при відмові одного або декількох його елементів.

Таким чином, *резервування* – це метод підвищення надійності технічних пристроїв або підтримання їх на необхідному рівні за допомогою введення апаратної надмірності за рахунок включення запасних (резервних) елементів і зв'язків, додаткових порівняно з мінімально необхідним, для виконання заданих функцій в даних умовах роботи.

При резервуванні з дробовою кратністю працездатний стан резервованого з'єднання можливий за умови, якщо число справних елементів не менше необхідного для виконання заданих функцій в даних умовах роботи.

*Кратність резерву* – це відношення числа резервних елементів об'єкта до числа основних елементів, що резервуються ними, виражених нескороченим дробом.

Кратність резерву обчислюється за формулою:

$$k = m/(n-m), \quad (6.1)$$

де  $n$  – загальне число елементів розрахунку резервованої системи;  $(n-m)$  – число елементів, необхідних для нормальної роботи системи;  $m$  – число резервних елементів.

Відмова системи настає при відмові  $(m + 1)$  елемента.

Система буде працездатною протягом часу  $t$  при відмові не більше

ніж  $m$  елементів.

З вимог додатків визначається число елементів ( $n-m$ ), необхідне для виконання заданих функцій в даних умовах роботи системи, наприклад кількість осіб в групі слухачів таке, щоб на двох слухачів була одна робоча станція.

Знаходження кількості резервних  $m$  елементів проводиться при допущенні, що основні елементи системи мають однакову надійність, відмови елементів описуються найпростішим потоком відмов, відновлення є незалежним.

### **6.1. Постановка задачі синтезу структури технічної системи із забезпеченням доступності**

Для забезпечення необхідної доступності використовується структурна надмірність.

Постановка задачі синтезу структури технічної системи по забезпеченню доступності:

1. Мається структурна схема системи.
2. Відомо число елементів  $n$  системи.
3. Відома інтенсивність відмови елементів вихідної системи.
4. Середній час відновлення вихідної системи.
5. Коефіцієнт оперативної готовності системи.

Необхідно визначити показники доступності системи:

1. Структурну схему системи, що задовольняє вимогам доступності.
2. Показники доступності системи:
  - a. ймовірність безвідмовної роботи системи  $P_s(t)$ .
  - b. щільність  $f_s(t)$  розподілу часу безвідмовної роботи (частота відмов) системи залежно від кратності резервування.
  - c. вираз надійності.
  - d. коефіцієнт готовності системи  $K_{rs}$ .
  - e. коефіцієнт оперативної готовності системи  $R_s$ .

Методика синтезу структури системи із забезпечення доступності наступна:

1. Аналіз доступності вихідної систем.
2. Визначення кратності роздільного резервування, що забезпечує вимоги надійності.
3. Визначення показників надійності системи: імовірності безвідмовної роботи системи  $P_s(t)$ , коефіцієнта готовності системи  $K_{rs}$ , коефіцієнта оперативної готовності системи  $R_s$  синтезованої структури.

Аналіз доступності вихідної системи проводиться за критерієм коефіцієнта оперативної готовності системи не менше заданого  $R_s > R_z$  і включає визначення:

- ймовірності безвідмовної роботи вихідної системи  $P_s(t)$  протягом часу  $t$ ,
- коефіцієнта готовності елементів системи  $K_{ri}$ ,
- коефіцієнта оперативної готовності вихідної системи для

експоненційного розподілу  $R_s = K_r P_s(t)$ .

Імовірність безвідмовної роботи для систем з дробовою кратністю резервування за умови, що всі елементи мають однакову надійність, розраховується з виразу:

$$, \quad (6.2)$$

де  $C$  – число поєднань з  $(n)$  по  $(i)$  елементів системи;  $p(t)$  – ймовірність безвідмовної роботи елемента системи;  $m$  – число резервних елементів системи;  $n$  – загальне число елементів системи з числом основних елементів  $(n-m)$ .

За допомогою формули (6.2) можна визначити необхідну кількість резервних елементів  $m$  системи для заданого числа основних  $(n-m)$ , що забезпечують необхідну ймовірність безвідмовної роботи системи  $P_s(t)$ .

Щільність розподілу часу безвідмовної роботи (частота відмов) розраховується з виразу:

$$f_s(t) = (n - m(t)) f(t) \quad (6.3)$$

де  $f(t)$  – щільність розподілу часу безвідмовної роботи елемента;

$Q(t) = 1 - p(t)$  – ймовірність відмови елемента.

Коефіцієнт готовності елементів розраховується з виразу:

$$K_n(t) = 1 / (1 + \rho_i), \quad i = 1, 2, \dots, n. \quad (6.4)$$

З літератури відомо, що інтенсивність відмов однієї мікросхеми дорівнює  $\lambda = [10^{-6}, 10^{-7}]$  час<sup>-1</sup>.

В одному керованому комутаторі другого або третього рівня, робочої станції є приблизно 100 мікросхем, отже, інтенсивність відмов одного комутатора або робочої станції дорівнює  $\lambda = [10^{-4}, 10^{-5}]$  час<sup>-1</sup>.

Доступність послуг мережі в основному закладається при проектуванні, конструюванні та виготовленні.

При резервуванні в проекті мережі заздалегідь передбачається заміна несправних елементів справними елементами. Це первинний метод підвищення надійності мережі.

Оцінка резервування проводиться за критерієм якості для відновлюваних систем – коефіцієнту оперативної готовності.

Оцінка ефективності резервування проводиться при наступних спрощуючих припущеннях: всі елементи мають однакову надійність, потік відмов є найпростішим.

Для постійної інтенсивності відмов елементів системи:

$$p(t) = e^{-\lambda t}, \quad (6.5)$$



формула ймовірності безвідмовної роботи системи приймає вигляд:

(6.6)

Формула безвідмовної роботи системи  $P_s(t)$  являє собою накопичені суми біноміального розподілу ймовірностей з параметрами  $n$  і  $p = 1 - e^{-\lambda t}$ .

Щільність з точки зору надійності системи показує закон розподілу часу до відмови або напрацювання до відмови.

Середнє напрацювання до відмови дорівнює площі під кривою ймовірності безвідмовної роботи об'єкта.

(6.7)

Щільність  $f_s(t)$  розподілу часу безвідмовної роботи (частота відмов) системи є диференціальною функцією розподілу і показує елементарну швидкість, з якою в будь-який момент часу відбувається прирощення числа відмов.

Коефіцієнт готовності – це ймовірність того, що об'єкт виявиться в працездатному стані в довільний момент часу, крім планованих періодів, протягом яких застосування об'єкта за призначенням не передбачається.

Коефіцієнт готовності  $K_r$  оцінює час  $T_{vi}$  відновлення елементів і розраховується з виразу:

(6.8)

де (6.9)

Коефіцієнт оперативної готовності визначається як ймовірність того, що об'єкт виявиться в працездатному стані в довільний момент часу, крім планованих періодів, протягом яких застосування об'єкта за призначенням не передбачається і, починаючи з цього моменту, працюватиме безвідмовно протягом заданого періоду часу.

Коефіцієнт оперативної готовності  $R_s$  характеризує надійність об'єктів, необхідність застосування яких виникає в довільний момент часу, після якого потрібна певна безвідмовна робота.

Коефіцієнт оперативної готовності системи для експоненційного розподілу відмов і відновлення елементів, що відмовили, дорівнює:

$$R_s = K_r P_s(t) \quad (6.10)$$

При аналізі доступності резервованих пристроїв на етапі проектування потрібно порівнювати схемні рішення за критерієм якості резервування – виграшу надійності.

Аналіз виграшу надійності розраховується за основними кількісними характеристиками.

Оцінка ефективності резервування здійснюється за критеріями якості – ймовірності та середньому наробітку на відмову при спрощуючих припущеннях: всі елементи мають однакову надійність, потік відмов елементів найпростіший.

Виграш резервованої системи в порівнянні з нерезерованою обчислюється:

$$G_p = P_s(t)P_0(t), \quad (6.11)$$

$$G_t = T/T_0 \quad (6.12)$$

## ЛІТЕРАТУРА

1. Алиев Т.И. Основы моделирования дискретных систем / Алиев Т.И. – СПб: СПбГУ ИТМО, 2009. – 363 с.
2. Кудрявцев Е.М. GPSS World. Основы имитационного моделирования различных систем / Е.М. Кудрявцев. – М: ДМК Пресс, 2004. – 320 с.
3. Томашевский В. Имитационное моделирование в среде GPSS / В. Томашевский, Е. Жданова. – Бестселлер, 2003. – 416 с.
4. Боев В.Д. Моделирование систем. Инструментальные средства GPSS World / В. Д. Боев. – БХВ-Петербург, 2004. – 368 с.
5. Жерновий Ю.В. Імітаційне моделювання систем масового обслуговування: [Навчальний посібник] / Ю.В. Жерновий. – Львів: Видавничий центр ЛНУ ім. Івана Франка, 2006. – 312 с.
6. Платунова С.М. Методы проектирования фрагментов компьютерной сети – СПб: НИУ ИТМО, 2012. – 51 с.
7. Вишнеvский В. М. Теоретические основы проектирования компьютерных сетей / В. М. Вишнеvский. – Москва: Техносфера, 2003. – 512 с.
8. [Jean Walrand](#). Network Performance Modeling and Simulation. / [Jean Walrand](#), [Kallol Bagchi](#), [George Zobrist](#). – CRC Press, April 1, 1998. – 332p.
9. Steven Elliot. [Modeling and Simulation of Computer Networks and Systems](#)/ Steven Elliot, Benjamin Rearick, Punithavathy Govindaradjane. – Elsevier Inc, 2015. – 924p.
10. <http://narfu.ru/university/library/books/1163.pdf>
11. <http://math.immf.ru/lectures/206.html>
12. <http://window.edu.ru/resource/124/47124/files/sssu068.pdf>
13. <http://www.resolventa.ru/data/metodstud/servtheory.pdf>
14. <http://portal.tpu.ru/SHARED/1/LASUKOV/ms/Tab1/g5.pdf>

**С.В. Шестопапов**

**ДОСЛІДЖЕННЯ ТА ПРОЕКТУВАННЯ  
КОМП'ЮТЕРНИХ СИСТЕМ ТА  
МЕРЕЖ**

Підписано до друку \_\_\_\_\_ 2017 р. Формат 60×84 1/16.

Умовн. друк. арк. **x,x**. Наклад **x** прим.  
Надруковано видавничий центр ОНАХТ «Технолог».  
65039, Одеса, вул. Канатна, 112