

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

ІВАНІСЕНКО ІГОР МИКОЛАЙОВИЧ

УДК 004.7:519.2

**МЕТОДИ БАЛАНСУВАННЯ НАВАНТАЖЕННЯ У РОЗПОДІЛЕНИХ
СИСТЕМАХ З УРАХУВАННЯМ САМОПОДІБНИХ ВЛАСТИВОСТЕЙ
ВХІДНИХ ПОТОКІВ**

05.13.05 – комп'ютерні системи та компоненти

Автореферат дисертації на здобуття наукового ступеня
кандидата технічних наук

Харків – 2017

Дисертацією є рукопис.

Робота виконана у Харківському національному університеті радіоелектроніки Міністерства освіти і науки України.

Науковий керівник доктор технічних наук, професор
Кіріченко Людмила Олегівна,
Харківський національний
університет радіоелектроніки,
професор кафедри прикладної математики.

Офіційні опоненти: доктор технічних наук, професор
Мірошник Марина Анатоліївна,
Український державний університет
залізничного транспорту,
професор кафедри інформаційних технологій;

кандидат технічних наук, доцент
Куланов Віталій Олександрович,
Національний аерокосмічний університет імені
М. Є. Жуковського «Харківський авіаційний інститут»,
доцент кафедри комп'ютерних систем та мереж

Захист відбудеться «13» грудня 2017 р. о 15.00 годині на засіданні спеціалізованої вченої ради Д 64.052.01 у Харківському національному університеті радіоелектроніки, за адресою: 61166, м. Харків, просп. Науки, 14.

З дисертацією можна ознайомитися в бібліотеці Харківського національного університету радіоелектроніки, за адресою: 61166, м. Харків, просп. Науки, 14.

Автореферат розісланий «11» листопада 2017 р.

Учений секретар
спеціалізованої вченої ради

Є.І. Литвинова

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Сьогодні разом зі збільшенням швидкостей передачі даних в інфокомунікаціях збільшується частка інтерактивного трафіка, вкрай чутливого до параметрів середовища транспортування. Для надання необхідної кількості ресурсів при передачі різних видів трафіка, що висувають різні вимоги до характеристик телекомунікаційної мережі, використовуються різні механізми забезпечення якості обслуговування QoS (Quality of service). Одним із таких механізмів є балансування навантаження. Система балансування навантаження вирішує завдання забезпечення якості обслуговування і підвищення продуктивності розподілених систем за рахунок оптимального розподілу завдань між вузлами обчислювальної системи.

Завдання балансування завантаження в розподіленій інфокомунікаційній системі полягає в тому, щоб виходячи з комплекту завдань, що передбачають обчислення і передачу даних, і системи серверів різної ресурсомісткості, знайти такий розподіл завдань по серверах, який забезпечує приблизно однакове обчислювальне завантаження кожного сервера і мінімальні витрати на передачу даних. Для виконання цього завдання можуть використовуватися різні методи та алгоритми балансування навантаження, які враховують оцінки завантаження обчислювального вузла.

Найбільш відомими дослідженнями в галузі управління навантаженням у розподілених системах, балансування, теоретичних досліджень і розробки фундаментальних основ розподілу навантаження, у створенні математичного апарату, моделей і методів управління для розподілу навантаження займалися такі вчені, як: Є.І. Ігнатенко, В.Н. Тарасов, F. Wang, V. Cardellini, Xing-Guo Luo, Hisao Kameda, Н. Mehta, P. Kanungo, M. Casalicchio, Y.S. Hong, а також інші дослідники, що працюють над проблемами розподілу навантаження. Розробляли і вдосконалили алгоритми балансування навантаження такі науковці, як: S. Keshav, O. Elzeki, M. Reshad, H. Chen, Y. Hu, Shamsollah Ghanbari, Ratan Mishra, Dhinesh Babu, а також багато інших.

Експериментальні дослідження, проведені в останні десятиліття, свідчать, що трафік у багатьох мультисервісних комп'ютерних мережах має самоподібні (фрактальні) властивості. Причина такого ефекту полягає в особливостях розподілу файлів по серверах, їхніх розмірах, типовій поведінці користувачів і значною мірою пов'язана зі змінами мережних ресурсів і топології мережі. Самоподібний трафік викликає значні затримки і втрати пакетів, навіть якщо сумарна інтенсивність всіх потоків далека від максимально допустимих значень.

Існує велика кількість публікацій, присвячених аналізу фрактальних властивостей трафіка. Самоподібні властивості інформаційних потоків виявлені в локальних і глобальних мережах, зокрема в трафіку Ethernet, АТМ, додатках TCP, IP, VoIP. Значний внесок у розвиток теорії самоподібних процесів, дослідження фрактальних властивостей телетрафіка і побудову моделей фрактального трафіка внесли K. Park, W. Willinger, P. Abry, M. Taqqu, I. Norros, Потапов А.А., Цибаков Б.С., Шелухін А.І.

Наявність у переданих клієнтами інформаційних потоках властивостей самоподібності істотно впливає на ефективність роботи розподілених систем.

Особливо важливу роль це відіграє для роботи сервісів, що забезпечують передачу мультимедійного трафіка і трафіка реального часу. Таким чином, актуальною є задача розробки та аналізу методів балансування навантаження, які враховують самоподібність трафіка і завантаження кожного вузла та всієї розподіленої системи.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконувалася відповідно до плану науково-технічних робіт Харківського національного університету радіоелектроніки в рамках держбюджетної теми: «Розробка та дослідження застосування GRID-порталу Харківського ресурсно-операційного GRID-центру», договору №08-22 (08.04.07-27.06.08) і №08-22/9 (01.07.08-30.09.08) між ХНУРЕ і «ІПСА» НТУУ «КПІ», що виконувалася на підставі Договору «ІПСА» НТУУ «КПІ» з Міністерством освіти і науки України №ІТ/506-2013, Державної програми «Інформаційні та телекомунікаційні технології в освіті і науці» на 2006 - 2013 роки (№ДР 0108U008261).

В рамках теми здобувачем досліджено властивості інформаційних потоків даних у розподілених системах.

Мета і завдання досліджень. Метою роботи є розробка модифікованих методів балансування навантаження у розподілених системах, які враховують фрактальні властивості трафіка і дозволяють забезпечити високий рівень якості обслуговування.

Відповідно до поставленої мети необхідно вирішити такі наукові завдання:

- провести огляд і аналіз методів балансування навантаження в розподілених системах і досліджень впливу фрактальних властивостей трафіка на якість обслуговування в мережах;
- удосконалити метод розрахунку дисбалансу завантаження ресурсів розподіленої системи для різних класів обслуговування вхідного трафіка, який передбачає комплексне вимірювання загального рівня дисбалансу системи;
- за допомогою імітаційного моделювання дослідити вплив самоподібних і мультифрактальних властивостей трафіка на показники якості обслуговування в мережі для найбільш затребуваних алгоритмів балансування навантаження;
- удосконалити модель балансування навантаження розподіленої системи, з урахуванням мультифрактальних характеристик вхідних потоків;
- на основі удосконаленої моделі розробити метод балансування, який враховує мультифрактальні властивості трафіка і метод розрахунку дисбалансу ресурсів розподіленої системи;
- провести імітаційне моделювання запропонованих методів балансування навантаження і вирішити практичні завдання.

Об'єкт дослідження – процеси управління трафіком у розподілених інфокомунікаційних системах.

Предмет дослідження – моделі та методи балансування самоподібного навантаження в розподілених системах.

Методи дослідження. Для вирішення поставлених завдань у роботі використовувалися методи фрактального і статистичного аналізів даних (для

дослідження властивостей самоподібності та розрахунку характеристик мережного трафіка); методи теорії ймовірностей і випадкових процесів (для побудови математичних моделей фрактального трафіка); методи лінійного програмування, теорії графів і оптимізації (для побудови математичних моделей мультисервісних мереж; імітаційне моделювання (для розробки методів управління потоками даних, перевірки запропонованих моделей і вироблення практичних рекомендацій).

Наукова новизна отриманих результатів. У дисертації отримано такі нові наукові і практичні результати:

1. Вперше запропоновано метод балансування навантаження, який характеризується, урахуванням мультифрактальних властивостей адитивного трафіка і можливістю розрахунку дисбалансу ресурсів розподіленої системи, що дозволяє підвищити ступінь використання ресурсів системи за рахунок спрямування неоднорідних інформаційних потоків на менш завантажені ресурси.

2. Удосконалено модель балансування навантаження розподіленої системи, яка, на відміну від існуючих, враховує обмеження на заданий набір характеристик мережі для різних класів обслуговування та зміну в часі фрактальних параметрів трафіка, що дозволяє розробити методи балансування з урахуванням певних вимог якості обслуговування.

3. Набув подальшого розвитку метод розрахунку дисбалансу ресурсів розподіленої системи, який, на відміну від існуючих передбачає комплексне вимірювання загального рівня дисбалансу системи та враховує вагові коефіцієнти ресурсів серверів, що дозволяє рівномірно розподіляти навантаження в неоднорідній системі.

Практична значущість отриманих результатів.

1. Розроблено програмні засоби для реалізації моделей і методів балансування навантаження розподілених систем, які враховують фрактальні властивості трафіка та дозволяють забезпечити задану якість послуг і підвищити ефективність використання обладнання та каналів передачі інфокомунікаційної мережі.

2. Проведено тестування і верифікацію програмного продукту для балансування навантаження використовуючи запропоновані моделі і методи балансування навантаження шляхом багаторазового імітаційного моделювання розподілених інфокомунікаційних мережах.

Результати дисертаційної роботи впроваджені для балансування навантаження розподіленої комп'ютерної мережі ТОВ «Ипра-софт» (акт від 11.01.2016 р.) та в навчальний процес кафедри електронних обчислювальних машин ХНУРЕ, м. Харків (акт від 16.01.2017 р.).

Основні результати дисертаційної роботи використовуються в навчальному процесі на кафедрі електронних обчислювальних машин Харківського національного університету радіоелектроніки під час проведення лекційних та лабораторних занять з дисциплін «Комп'ютерні мережі» та «Корпоративні комп'ютерні мережі».

Особистий внесок здобувача. Всі наукові результати дисертаційної роботи отримано автором самостійно і опубліковані в роботах [1-29]. У роботах, опублікованих у співавторстві, автору належать такі результати: [2] –

досліджено модель доступу; [3] – проведено класифікацію і формальні моделі апаратних закладних пристроїв; [4] – проведено імітаційне моделювання хмарних сервісів з урахуванням самоподібних властивостей вхідних потоків; [5] – запропоновано модель і методи балансування самоподібного трафіка; [7] – запропоновано метод розподілу навантаження з урахуванням мультифрактальних властивостей трафіка; [8] – виконано аналіз дисбалансу розподіленої системи при самоподібному навантаженні; [9] – запропоновано методи розрахунку дисбалансу розподіленої системи; [14] – проведено моделювання хмарних сервісів з урахуванням самоподібних властивостей вхідних потоків; [16] – досліджено самоподібні властивості адитивних потоків; [17] – проведено аналіз методів балансування навантаження в розподілених системах; [18] – виконано огляд алгоритмів балансування в розподілених системах; [19] – запропоновано алгоритм балансування для мультифрактального трафіка; [20] – проведено дослідження зашумлених мультифрактальних рядів; [21] – проведено моделювання хмарних сервісів для самоподібних потоків; [23] – досліджено динамічне балансування фрактального трафіка; [24] – розглянуто алгоритми балансування самоподібного трафіку; [26] – запропоновано метод балансування самоподібного трафіку; [27] – досліджено мультифрактальні властивості адитивних потоків; [28] – досліджено стан розподілених систем при самоподібному навантаженні; в роботі [29] запропоновано метод розрахунку дисбалансу розподіленої системи.

Апробація результатів дисертації. Основні результати проведених досліджень доповідалися і обговорювалися на таких конференціях:

- 8-й та 9-й міжнародній науково-технічній конференції «Проблеми інформатики і моделювання» (Харків, 2008, 2009);

- 1-й та 2-й науково-технічній конференції «Інформаційні технології в навігації і управлінні: стан та перспективи розвитку» (Київ, 2010, 2011);

- 1-й, 6-й та 7-й науково-технічній конференції «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління» (Київ, 2010; Полтава, 2015; Харків, 2016);

- 18-му Міжнародному молодіжному форумі «Радіоелектроніка і молодь в XXI столітті» (Харків, 2014);

- Xth International Scientific and Technical Conference «Computer science and information technologies» (Lviv, 2015);

- II Міжнародній науково-практичній конференції «Інформаційні технології у інноваційному бізнесі (ITIB 2015)», (Харків, 2015);

- Second and third International Scientific Practical Conference "Problems of Infocommunications Science and Technology" (Kharkiv, 2015, 2016);

- Міжнародній науково-технічній конференції "Інформаційні системи і технології" (Харків, 2015);

- 3-й міжнародній науково-технічній конференції «Проблеми автоматизації», (Черкаси, 2015);

- 7-й всеукраїнській науково-практичній конференції «Free and Open Source Software» (Харків, 2015);

- Міжнародній науково-технічній конференції «Інформаційні технології

в металургії та машинобудуванні (ITMM - 2016)» (Дніпропетровськ, 2016);

– XIIIth International Conference TCSET'2016 «Modern problems of radio engineering, telecommunications, and computer science» (Lviv-Slavsko, 2016);

– 17-й міжнародній науково-практичній конференції «Сучасні інформаційні та електронні технології» (Одеса, 2016);

– 2016 IEEE First International Conference on data mining and processing. (Львів, 2016);

– Міжнародній науково-практичній конференції «Інформаційні управляючі системи та технології» (Одеса, 2016).

Публікації. Основні результати дисертаційної роботи опубліковано в 29 роботах: з них 6 статей у наукових фахових виданнях України з технічних наук, 3 статті - в міжнародних наукових журналах за кордоном та 20 матеріалів і тез доповідей конференцій (з них 6 входить до наукометричної бази Scopus).

Структура дисертації. Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, переліку використаних джерел зі 158 найменувань на 19 сторінках та 3 додатків на 10 сторінках, а також містить 33 рисунки та 10 таблиць. Загальний обсяг дисертації складає 184 сторінки, включаючи 133 сторінки основного тексту.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтовується актуальність теми дослідження, розглядаються проблеми, які виникають через властивості самоподібності в телекомунікаційних системах, формулюється мета та завдання дослідження, визначаються об'єкт, предмет і методи досліджень, висвітлюється наукова новизна та практичне значення отриманих результатів, а також особистий внесок автора в роботах, виконаних у співавторстві, відзначається апробація результатів дисертації та кількість публікацій, виконаних за темою дисертаційної роботи.

У першому розділі розглянуто основні поняття балансування навантаження в розподіленій комп'ютерній мережі. Визначено основні властивості основних алгоритмів балансування навантаження. Наведено їхню класифікацію за кількома типами. Виділено основні показники продуктивності: пропускна здатність, витрати, відмовостійкість, час міграції, час відгуку, використання ресурсів, масштабованість, продуктивність.

У ході роботи розглянуто особливості інфокомунікаційного мережного трафіка, що має самоподібні та мультифрактальні властивості. Самоподібність випадкових процесів полягає у збереженні статистичних характеристик при зміні масштабу часу. Параметр, що називається показником Херста H , є ступенем самоподібності. Поряд з цим показник H характеризує міру довгострокової залежності, тобто спадання автокореляційної функції за степеневим законом. Самоподібний трафік має особливу структуру, яка зберігається на багатьох масштабах – у реалізації завжди присутня деяка кількість дуже великих сплесків за відносно невеликим середнім рівнем трафіка. В останнє десятиліття інтенсивно вивчаються мультифрактальні властивості трафіку. Мультифрактальний трафік визначається як розширення

самоподібного трафіка за рахунок урахування масштабованих властивостей статистичних характеристик другого і вище порядків.

У роботі показано, що найбільш важливою задачею є розробка методів управління потоками даних і ресурсами розподілених комп'ютерних мереж, які враховують самоподібні та мультифрактальні властивості трафіка. Особливу увагу приділено задачі балансування, яка забезпечує більш високу утилізацію ресурсів мережі. Проведено огляд публікацій, присвячених аналізу впливу методів балансування на якість обслуговування в розподілених комп'ютерних мережах, що функціонують в умовах самоподібного трафіка. Показано, що розробка методів балансування фрактальних інформаційних потоків, які дозволяють забезпечити високий рівень якості обслуговування (QoS) в комп'ютерних мережах, є актуальним завданням.

У другому розділі розглянуто основні механізми забезпечення якості обслуговування мережі. Більшість із них враховує і використовує у своїй роботі факт існування в мережі трафіка різного типу, оскільки кожен тип трафіка висуває різні вимоги до характеристик продуктивності та надійності мережі. Проведено аналіз механізмів збору інформації щодо завантаженості системи. На основі зібраної інформації про завантаженість системи та вхідного навантаження розраховується дисбаланс кожного вузла та розподіленої системи взагалі. Під дисбалансом розуміють ступінь рівномірності розподілу навантаження між серверами. Однією з основних задач балансування навантаження є зменшення дисбалансу.

Набув подальшого розвитку метод розрахунку дисбалансу ресурсів розподіленої системи (п.3 наукової новизни), який, на відміну від існуючих, передбачає комплексне вимірювання загального рівня дисбалансу системи і враховує вагові коефіцієнти ресурсів серверів, що дозволяє рівномірно розподіляти навантаження в неоднорідній системі. Також запропонований метод розрахунку дисбалансу враховує розподіл потоків за класами обслуговування. У комп'ютерних мережах на маркування класів обслуговування виділяється 8 бітів, таким чином, можливі 24 класи, але зазвичай використовуються 8 класів. Метод розрахунку дисбалансу полягає у тому що:

1. Проводимо розрахунок розподілу потоків за серверами на основі перелічених значень трудомісткості, інтенсивності трафіка, стану завантаженості серверів і каналів зв'язку.

1.1. Оцінюємо $CPU_i^u(T)$ – середнє завантаження кожного u -го процесора i -го сервера, $u=1, N_{cpu}; i=1, N_S$;

1.2. $RAM_i^r(T)$ – середнє завантаження кожної r -ї пам'яті i -го сервера, $r=1, N_{ram}; i=1, N_S$;

1.3. $Net_i^k(T)$ – середнє завантаження k -го каналу i -го сервера, $k=1, N_{net}; i=1, N_S$ протягом періоду спостереження T .

2. Обчислюємо середнє завантаження кожного процесора, пам'яті і каналів для кожного сервера потоками класу $qs = [0, \dots, 7]$.

2.1. Обчислюємо завантаження процесора потоком класу qs :

$$CPU_i^{qs}(T) = CPU_i^u \times f_{CPU}^{qs},$$

де f_{CPU}^{qs} – частка сумарного використання u -го процесора, яка може належати класу qs : $f_{CPU}^{qs} = CPU_i^{qsv} / \sum_{\forall qs} CPU_i^{qsv}$; CPU_i^{qsv} – завантаження i -го процесора потоками класу qs , яка була виміряна системою обліку або монітором операційної системи.

2.2. – 2.3. Виконуємо аналогічно для RAM_i^{qs} та Net_i^{qs} .

3. Обчислюємо середній коефіцієнт використання кожного ресурсу в системі.

3.1. Розраховуємо CPU_u^{All} – середній коефіцієнт використання всіх процесорів у системі:

$$CPU_u^{All} = \frac{\sum_i^N CPU_i^u CPU_i^{n_i}}{\sum_i^N CPU_i^{n_i}},$$

де $CPU_i^{n_i}$ – середнє завантаження ЦПУ i -го сервера, N_S – загальна кількість фізичних серверів у системі, N_{cpu}^i – кількість ЦПУ на i -му сервері.

3.2. – 3.3. Виконуємо аналогічно для RAM_i^r , Net_i^k .

4. Обчислюємо значення дисбалансу всіх ресурсів в системі.

4.1. Розраховуємо IMB_{CPU} – значення дисбалансу всіх процесорів в системі:

$$IMB_{CPU} = 1/N \sum_i^N (CPU_i^u - CPU_u^{All})^2.$$

4.2. – 4.3. Виконуємо аналогічно для IMB_{RAM} , IMB_{Net} .

5. Визначимо комплексне значення дисбалансу навантаження i -го сервера, що враховує всі три ресурси серверу:

$$IMB_i = K_{CPU} (CPU_i^u - CPU_u^{All})^2 + K_{RAM} (RAM_i^r - RAM_r^{All})^2 + K_{Net} (Net_i^k - Net_k^{All})^2,$$

де $K_{CPU}, K_{RAM}, K_{Net}$ – вагові коефіцієнти для процесора, пам'яті і пропускної здатності мережі.

6. Розраховуємо середнє значення дисбалансу всіх серверів в системі:

$$IMB_{tot} = \frac{1}{N} \sum_i^N IMB_i.$$

7. Середня тривалість роботи при однаковій кількості задач дозволяє порівнювати різноманітні алгоритми планування.

8. Період обробки на i -му сервері визначається як максимальне навантаження на i -му сервері. Період обробки у системі визначається як середнє завантаження на усіх серверах.

9. Ефективність використання визначається як середнє навантаження на будь-якому сервері.

Таким чином, для планування ресурсів було розроблено методику комплексного вимірювання загального рівня дисбалансу системи, а також середнього рівня дисбалансу кожного сервера.

У третьому розділі запропоновано удосконалену модель балансування

самоподібного навантаження розподіленої системи (п.2 наукової новизни). Для цього були вивчені властивості мультифрактальних, зокрема адитивних фрактальних потоків даних, і проведено порівняльний аналіз алгоритмів балансування навантаження.

Проведено дослідження фрактальних властивостей сумарних потоків різного типу. Оскільки сучасні інформаційні мережі побудовані на основі мультиплексування потоків даних, одним із важливих завдань для підвищення якості обслуговування мереж є дослідження властивостей адитивних трафіків. Числові дослідження показали, що для визначення фрактальних властивостей адитивного трафіка необхідно враховувати показники Херста і співвідношення коефіцієнтів варіації $\sigma_{\text{var}}[Y_i]$ сумованих потоків $Y_i(t)$.

Для суми двох потоків: самоподібного $Y_1(t)$ і не самоподібного $Y_2(t)$ оцінювання показника Херста сумарного потоку показало, що у разі, коли відношення $R = \sigma_{\text{var}}[Y_1]/\sigma_{\text{var}}[Y_2]$ приблизно дорівнює одиниці, сумарний потік набуває самоподібних властивостей, а його значення показника Херста в середньому дорівнює показнику Херста $H(Y_1)$. У випадку, коли відношення R зменшується, показник Херста сумарного процесу також поступово зменшується, досягаючи значення 0.5.

Значення показника Херста для суми двох самоподібних потоків визначається максимальним значенням показника Херста потоків і співвідношенням коефіцієнтів варіації. При $R_1 \approx 1$ сумарний потік є самоподібним з максимальним показником Херста. У випадку, коли відношення R стає більше одиниці, показник Херста сумарного потоку прагнучим до значення $H(Y_1)$.

В ході підсумовування декількох самоподібних потоків, з різними коефіцієнтами варіації та показниками Херста, доцільно ввести коефіцієнт

$$R = \sigma_{\text{var}}(H \max) / \frac{1}{N-1} \sum_{i=1}^{N-1} \sigma_{\text{var}}[Y_i], \text{ де } N - \text{кількість підсумованих потоків,}$$

$\sigma_{\text{var}}(H \max)$ – коефіцієнт варіації потоку з найбільшим показником Херста.

Числові дослідження показали, що в цьому випадку показник Херста сумарного процесу залежить від того, наскільки значення R менше одиниці. При значеннях $R \approx 1$ показник Херста сумарного потоку збігається з максимальним значенням $H(Y_i)$.

У роботі проведено числові дослідження, які виявили взаємно однозначну відповідність між значенням коефіцієнта варіації σ_{var} і діапазоном узагальненого показника Херста Δh , що дає можливість переходу до мультифрактальних характеристик потоку. Таким чином, знаючи фрактальні характеристики потоків що сумуються, які подаються на вхід розподіленої системи, можна розрахувати навантаження адитивних потоків та обрати найкращу стратегію балансування.

Для створення такої стратегії удосконалено модель балансування самоподібного навантаження розподіленої системи, яка, на відміну від існуючих, враховує обмеження на заданий набір характеристик мережі для

різних класів обслуговування та зміну в часі фрактальних параметрів трафіка, що дозволяє розробити методи балансування з урахуванням певних вимог якості обслуговування.

Балансувальники навантаження використовують різноманітні алгоритми управління трафіком для розподілу навантаження та/або максимального використання усіх серверів у кластері. Проведено порівняльний аналіз алгоритмів балансування навантаження за такими показниками продуктивності: пропускна здатність, витрати, відмовостійкість, час міграції, час відгуку, використання ресурсів, масштабованість, продуктивність. Вказано їх переваги та недоліки. На основі проведеного аналізу було обрано чотири найпродуктивніші алгоритми: Round Robin Weight, Least Connection, Compare and Balance, Connection mechanism. Також у ході роботи було проведено порівняння різноманітних типів сценаріїв розподілу навантаження у розподіленому середовищі.

Розглянемо модель балансування самоподібного навантаження розподіленої системи. У кожний момент $t \in T$ на балансувальник LB поступає трафік, що належить до qs -го класу обслуговування, який необхідно доставити на i -й сервер $Serv_i$ для обробки, без перевищення заданих максимально допустимих значень затримки τ_{qs} і максимально допустимого процента втрат l_{qs} залежно від поточного завантаження серверів і реальної пропускної здатності у конкретний момент часу.

Вхідний трафік описується множиною характеристик: $V = \{ \lambda, h, \mu \}$, де $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_\sigma]$, $\sigma = 1, 2, \dots$ – потоки заявок (пакетів) різноманітної інтенсивності; $h = [H, h(q), \Delta h]$, де $h(q)$ – вибіркове значення функції узагальненого показника Херста, $H = h(2)$ – значення параметра Херста, $\Delta h = h(q_{\min}) - h(q_{\max})$ – діапазон значень узагальненого показника Херста для ділянки трафіка, μ – трудомісткість запиту. Трудомісткість запиту визначається як вектор ресурсів, які необхідні $\mu = (CPU, Net, RAM)$ для виконання запиту. Кожному qs -му класу обслуговування відповідає набір векторів необхідних ресурсів $\mu_{qs} = (CPU, Net, RAM)$.

Балансувальник навантаження та сервера поєднані між собою двобічними мережними зв'язками $Link_{lk} = \{ L_{lk} \}$, $lk = 1, 2, \dots$, що поділяються на k каналів із пропускною здатністю $Net_{lk}^k(t) = \{ Net_{lk}^k \}$ у момент часу t . Кожен i -й сервер характеризується такими параметрами: $CPU_i^{n_i}$ – середнє завантаження ЦПУ i -го сервера за період часу T , де n_i – кількість ЦПУ (ядер) на i -му сервері; середній коефіцієнт використання пам'яті i -го сервера $RAM_i^{m_i}$ за період часу T , де m_i – кількість пам'яті на i -му сервері; середній коефіцієнт використання пропускної здатності каналу $Net_i^{k_i}$ за період часу T , де k_i – кількість каналів у лінії зв'язку між балансувальником та на i -му сервері.

На вхід балансувальника навантаження LB поступає декілька незалежних мультифрактальних потоків пакетів з різною інтенсивністю $\lambda_1, \lambda_2, \dots, \lambda_\sigma$, кожен з

яких відправляється у чергу Q_w обмеженої ємності. Час обслуговування заявок залежить від класу обслуговування qs , тобто враховується пріоритетність заявок (найвищий пріоритет – першим). Система, що розглядається, має механізм, яка працює на основі ймовірності та виконує функцію виштовхування. Підсистема балансування завантаження LB , відповідно до закладених у неї алгоритмів, виконує вилучення задач із черг Q_w та призначає їх на вільні обчислювальні ядра доцільних серверів.

Балансувальник LB у t -й момент характеризується коефіцієнтом втрат $X_{LB}^{qs}(t) \in X$, середнім часом очікування пакета у черзі $T_{LB}^{qs}(t) \in T$. Змінна $X_{LB}^{qs}(t) \in X$ дорівнює проценту втрат на балансувальнику трафіка з класом обслуговування qs , що передається шляхом $Net_{lk}^k(t)$ на сервер $Serv_i$ у момент t . Передбачається, що ймовірністю викривлення пакета у тракті можна знехтувати, а втрати відбуваються виключно у балансувальнику через переповнення буфера.

На розмір втрат накладається таке обмеження:

$$0 \leq X_{LB}^{qs}(t), \sum_{i=1}^N X_{LB}^{qs}(t) \leq l_{qs},$$

для усіх вузлів мережі. Втрати визначаються як відношення кількості відкинутих даних до кількості отриманих на обслуговування. Мінімізації підлягає значення $X_{LB}^{qs}(t) \rightarrow \min$.

Обмеження, що накладаються на час затримки, аналогічні:

$$0 \leq T_{LB}^{qs}(t), \sum_{i=1}^N T_{LB}^{qs}(t) \leq \tau_{qs},$$

де $T_{LB}^{qs}(t)$ – середній час очікування пакета класу обслуговування qs у черзі на вузлі i .

Також мінімізації підлягають значення дисбалансу $IMB_{tot} \rightarrow \min$, $IMB_i \rightarrow \min$ (1).

У четвертому розділі запропоновано метод балансування навантаження (п.1 наукової новизни), який враховує мультифрактальні властивості адитивного трафіка, і розрахунок дисбалансу ресурсів розподіленої системи, що дозволяє підвищити ступінь використання ресурсів системи за рахунок спрямування неоднорідних інформаційних потоків на менш завантажені ресурси.

Наведемо поетапний опис динамічного методу балансування навантаження:

1. У трафіку, що поступає на балансувальник, виділяється вікно X , фіксованої довжини T .

2. Знаходимо за трафіком інтенсивність, вибіркоче значення функції узагальненого показника Херста $h(q)$, значення параметра Херста $H = h(2)$ та діапазон значень узагальненого показника Херста $\Delta h = h(q_{\min}) - h(q_{\max})$ для ділянки трафіка у виділеному вікні.

3. Проводимо збір та аналіз статистичної інформації за серверами та каналами: середній коефіцієнт використання пропускної здатності каналу $Net_i^{k_i}$ за період часу T , стани серверів $CPU_i^{n_i}$, $RAM_i^{m_i}$, середнє завантаження ЦПУ та середній коефіцієнт використання пам'яті i -го сервера за період часу T .

4. На основі мультифрактальних властивостей трафіка та значень трудомісткості запитів обчислюємо набір векторів необхідних ресурсів $\mu_{qs}^{new} = (CPU, Net, RAM)$ для кожного qs -го класу трафіка.

$$\mu_{qs}^{new} = \begin{cases} \mu_{qs}, & H \leq 0,5; \\ \mu_{qs} + (H - 0,5)\mu_0, & 0,5 < H < 0,9, \Delta h \leq 0,4; \\ \mu_{qs} + (H - 0,5)(\Delta h - 0,4)\mu_0, & 0,5 < H < 0,9, 0,4 < \Delta h < 1; \\ \mu_{qs} + \mu_0, & H \geq 0,9 \text{ or } H > 0,5, \Delta h \geq 1, \end{cases}$$

де μ_{qs} визначається відповідно до класу обслуговування та необхідними ресурсами, значення μ_0 обирається адміністратором мережі з урахуванням стану мережі. Щоб відобразити зміну мультифрактальних властивостей потоків, вектори потрібних ресурсів μ_{qs}^{new} оновлюються у регулярні проміжки часу та перераховуються за формулою.

Кількість потрібних ресурсів не змінюється ($\mu_{qs}^{new} = \mu_{qs}$), якщо трафік є звичайним пуассонівським потоком ($H = 0,5$) або має антиперсистентні властивості ($H < 0,5$). При значенні $0,5 < H < 0,9$ та малому розкиді даних ($\Delta h \leq 0,4$) значення μ_{qs} збільшується пропорційно значенню показника Херста. При значенні показника Херста $0,5 < H < 0,9$ та великому розкиді даних ($0,4 < \Delta h < 1$) значення μ_{qs} збільшується пропорційно обом характеристикам. Кількість потрібних ресурсів з максимальним значенням $\mu_{qs} + \mu_0$ отримується при значенні $H \geq 0,9$ або при персистентному трафіку ($H > 0,5$) з діапазоном значень узагальненого показника Херста $\Delta h \geq 1$. Після перерахування вартості усіх шляхів повідомлення про стан шляхів розсилається поміж маршрутизаторами.

5. Проводимо розрахунок розподілу потоків по серверах на основі перелічених значень трудомісткості, інтенсивності трафіка, стану завантаженості серверів і каналів зв'язку.

6. На основі отриманих даних визначається завантаженість серверів на наступному етапі.

7. Розподіляємо трафік за серверами, згідно з алгоритмом балансування у межах кожного класу потоку.

8. Якщо весь трафік не вдалося розподілити, то проводимо розподіл трафіка, що залишився, серед кількості ресурсів, що є в наявності $CPU_i^u(T)$, $RAM_i^r(T)$, $Net_i^k(T)$. Переоцінка не враховується алгоритмом, тому що не вносить суттєвих змін.

9. Проводимо збір даних про завантаженість серверів $CPU_i^u(T)$,

$RAM_i^r(T)$, $Net_i^k(T)$ та передачу їх у систему балансування навантаження для розрахунку нового розподілу потоків.

10. Зсуваємо вікно X довжини T вперед на задану величину зсуву ΔT ; здійснюємо аналіз трафіка та розрахунок наступного значення завантаженості серверів.

Проведено імітаційне моделювання роботи запропонованого методу балансування навантаження за допомогою програмного продукту, написаного мовою Python.

На рис. 1 та 2 показано зміну дисбалансу процесорів, пам'яті та пропускної здатності у випадку двох кластерів, а також комплексне значення дисбалансу навантаження кожного кластера за формулою (1) під час роботи алгоритму балансування Compare Balance. Структура кожного кластера складається з шести серверів, які мають різні параметри моделі балансування (див. розділ 3).

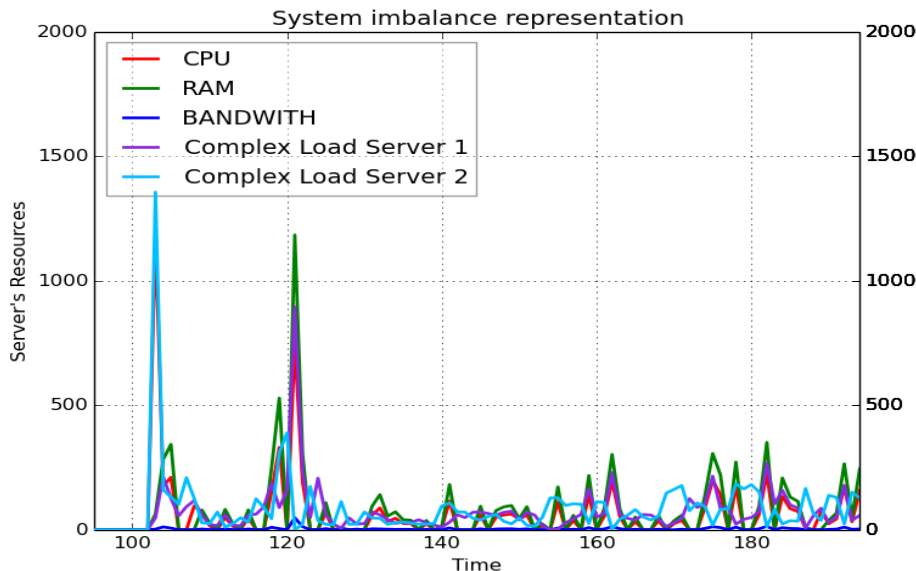


Рисунок 1 – Дисбаланс системи при параметрах трафіка $H = 0.6$ і $\Delta h = 1.5$

У першому випадку (рис. 1) на балансувальник надходить згенерований трафік з параметром $H = 0.6$ та діапазоном узагальненого показника Херста $\Delta h = 1.5$. Рис. 2 демонструє дисбаланс системи для трафіка з більшою довгостроковою залежністю ($H = 0.7$) і неоднорідністю $\Delta h = 2$. З рисунка 2 видно, що перші 230 секунд роботи система знаходиться у нестійкому стані та ресурси серверів використовуються неефективно, але починаючи з 230-ї секунди роботи, дисбаланс системи починає зменшуватися, також зменшується середнє значення використання ресурсів. Починаючи з 400-ї секунди роботи, система приходить до стабільного стану, та значення дисбалансу практично не змінюється.

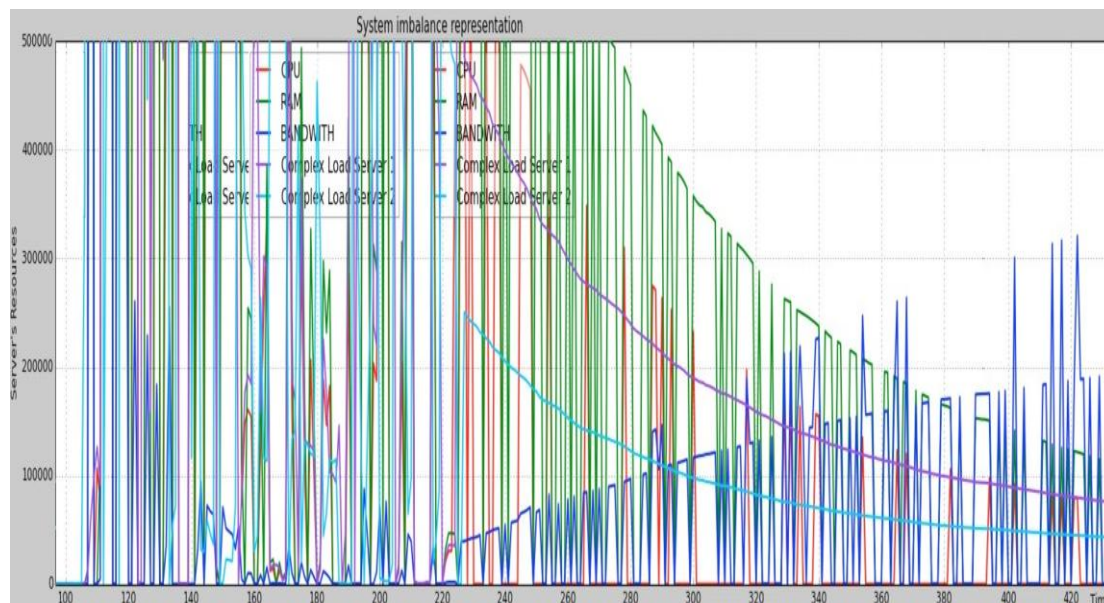


Рисунок 2 – Дисбаланс системи при параметрах трафіка $H=0.7$ і $\Delta h=2$

Дослідження показали, що дисбаланс системи суттєво залежить від мультифрактальних характеристик трафіка. При невеликих значеннях H та невеликій неоднорідності система балансування приходить у рівноважний стан, а значення дисбалансу прагне до нуля. Зі збільшенням показника Херста з часом дисбаланс системи не затухає, а система балансування не приходить до рівноважного стану. З більшими значеннями показника Херста та великою неоднорідністю система балансування знаходиться у нестійкому стані та значення дисбалансу змінюється у декілька разів, що призводить до максимального завантаження ресурсів.

У ході імітаційного моделювання проводився аналіз параметрів якості обслуговування під час запропонованого динамічного методу балансування навантаження, стандартного методу та без методів підвищення QoS. У таблиці 1 наведено середні оцінки параметрів якості обслуговування без впровадження методів підвищення QoS, зі стандартним методом балансування та запропонованим в роботі методом.

Таблиця 1 – Середні оцінки параметрів якості обслуговування мережі

Методи	Оцінки	Дисбаланс системи	Середнє значення втрат, %	Середній час очікування, мс
Без методів підвищення QoS		0,7	3,6	9,7
Стандартний метод балансування навантаження		0,65	2,5	7,9
Динамічний метод балансування навантаження		0,52	1,8	4,6

Із таблиці видно, що запропонований метод балансування навантаження за допомогою аналізу та обліку самоподібних властивостей вхідного потоку забезпечує статистично рівномірний розподіл навантажень на серверах, високі показники продуктивності та пропускну здатності, а також низький час очікування завдань в системі та низьку кількість втрачених даних.

У додатках наведено акти про впровадження результатів дисертаційної роботи у виробництво та навчальний процес, список публікацій за темою дисертації та відомості про апробацію результатів дисертації.

ВИСНОВКИ

Проведені науково-технологічні дослідження в рамках дисертаційної роботи характеризуються успішним вирішенням актуальної науково-практичної задачі розробки методів балансування навантаження, які враховують властивості самоподібності трафіка і використання ресурсів кожного вузла та всієї розподіленої системи.

Автором одержано такі наукові та практичні результати:

1. Удосконалено математичну модель системи балансування навантаження, в якій балансувальник навантаження описується за допомогою системи масового обслуговування. Стани серверів описуються обсягом вільних ресурсів ЦПУ і обсягом вільної оперативної пам'яті. Всі значення параметрів моделі мають залежність від часу. Така модель дозволяє описувати поведінку розподіленої мережі в часі для різних класів обслуговування вхідного трафіка, при заданих обмеженнях на час очікування пакета в черзі і кількість втрачених пакетів.

2. Набув подальшого розвитку метод розрахунку дисбалансу системи на основі оцінки завантаження вузлів розподіленої системи. Як оцінка завантаження ресурсів вузлів запропоновано характеристики завантаження процесора, пам'яті і пропускної здатності каналу. Розраховується середнє завантаження процесора, пам'яті і пропускної здатності каналу на основі завантаження, яка виміряна системою обліку або моніторингом операційної системи. Запропонований метод дозволяє проводити розрахунок завантаження для потоків різних класів обслуговування, як для кожного сервера окремо, так і для всієї системи. Введено комплексне значення дисбалансу навантаження сервера, що враховує вагові коефіцієнти для процесора, пам'яті і пропускної здатності мережі. Введені вагові коефіцієнти дозволяють визначити значущість кожної характеристики сервера по відношенню одна до одної. Таким чином, цей метод дозволяє обчислити дисбаланс всіх серверів системи й ефективність використання ресурсів системи.

3. Проведено дослідження зміни фрактальних властивостей для адитивних інформаційних потоків у випадках, коли хоча б один з сумованих потоків має самоподібні або мультифрактальні властивості. Результати дослідження дозволяють визначити характеристики фрактального трафіку при мультиплексуванні потоків даних в комп'ютерній мережі.

4. На основі розробленої математичної моделі системи балансування навантаження запропоновано динамічний метод розподілу навантаження, який враховує мультифрактальні властивості трафіка і задані обмеження на час очікування і кількість втрачених пакетів. Запропонований метод балансування навантаження завдяки аналізу та обліку мультифрактальних властивостей вхідного потоку забезпечує статистично рівномірний розподіл навантаження на

серверах, високі показники продуктивності і пропускної здатності, а також зниження часу відгуку і кількості втрачених даних.

5. Розроблено програмне забезпечення для виконання імітаційного моделювання роботи балансувальника розподіленої системи. Методами імітаційного моделювання виконано порівняння показників якості обслуговування мережі для найбільш затребуваних алгоритмів балансування навантаження. На підставі результатів досліджень запропоновано рекомендації щодо вибору алгоритмів балансування навантаження в комп'ютерних мережах із фрактальним трафіком залежно від мережних параметрів і характеристик.

6. Розроблені моделі та методи були реалізовані та впроваджені в корпоративній мережі компанії «Іпра-софт», де вони використовуються для попередження перевантаження вузлів мережі за рахунок перерозподілу потоків даних. Також результати дисертаційної роботи впроваджені в навчальний процес у Харківському національному університеті радіоелектроніки, що підтверджено відповідними актами.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Иванисенко И.Н. Проблемы моделирования дискретно-событийных систем // Вісник Харківського Університету. – Серія: Актуальні проблеми сучасної науки в дослідженнях молодих вчених м. Харкова. – 2002. – № 551. – С. 195–199.

2. Горбачов В.О. Суб'єктно-об'єктна модель доступу з використанням апаратних закладок до комп'ютерної інформації / В.О. Горбачов, В.В. Степаненко, І.М. Іванісенко // Радиоэлектроника и информатика. – 2006. – Том 6, №3. – С. 47–50.

3. Горбачев В.А. Классификация и формальные модели аппаратных закладных устройств // В.А. Горбачев, И.Н. Иванисенко // Прикладна радіоелектроніка та інформатика. – 2007. – Том 6, № 2. – С. 306–310.

4. Иванисенко И.Н. Имитационное моделирование облачных сервисов с учетом самоподобных свойств входящих потоков. Информатика, математическое моделирование, экономика / И.Н. Иванисенко, Ю.А. Кобицкая // Сб. научных статей по итогам Четвертой Междунар. науч.-практ. конф. – 23–25 апр. 2014. – Россия, Смоленск, – Том 1. – С. 79–83.

5. Иванисенко И.Н. Методы балансировки с учетом мультифрактальных свойств загрузки / И.Н. Иванисенко, Л.О. Кириченко, Т.А. Радивилова // Information content and processing. – 2015. – Vol. 2 (4). – P. 345–368.

6. Ivanisenko Igor. Methods and Algorithms of load balancing / Igor Ivanisenko // Information technologies and knowledge. – 2015. – Vol. 9, N. 4. – P. 340–375.

7. Иванисенко И.Н. Об одном методе распределения нагрузки с учетом мультифрактальных свойств трафика / И.Н. Иванисенко, Л.О. Кириченко, Т.А. Радивилова // Системні технології. – 2016. – Вип. 3(104). – С. 125–135. (Входить до міжнар. наукометричних баз Index Copernicus, Google Scholar).

8. Иванисенко И.Н. Анализ дисбаланса распределенной системы при самоподобной нагрузке / И.Н. Иванисенко, Л.О. Кириченко, Т.А. Радивилова // Вісник Херсонського національного технічного університету. – 2016. – Вип. 3(58). – С. 224–231. (Входить до міжнар. наукометричних баз:

РИНЦ (eLibrary, Google Scholar).

9. Радивилова Т.А. Динамический метод оценки загрузки узлов распределенной системы [Электронный ресурс] / И.Н. Иванисенко, Т.А. Радивилова // Проблемы телекоммуникацій. – 2016. – № 1(18). – С. 42–51. – Режим доступу до

журн.: http://pt.journal.kh.ua/2016/1/1/161_radivilova_utilization.pdf.

10. Иванисенко И.Н. О проблеме решения задач оптимального распределения ресурсов в вычислительной системе / И.Н. Иванисенко // Проблемы информатики и моделирования: Восьмая междунар. науч.-техн. конф., 26-28 нояб. 2008: матер. конф. – Харьков, Украина, 2008. – С. 53.

11. Иванисенко И.Н. Модель оптимального использования ресурсов IP-сети путем оптимизации трафика / И.Н. Иванисенко // Проблемы информатики и моделирования: Девятая междунар. науч.-техн. конф., 26-28 нояб. 2009: матер. конф. – Харьков, Украина, 2009. – С. 44.

12. Иванисенко И.Н. Методы решения задачи балансировки вычислительной нагрузки в сети распределенных вычислений / И.Н. Иванисенко // Информационные технологии в навигации и управлении: состояние и перспективы развития: Первая научно-техн. конф., 5–6 июля 2010: матер. конф. – Киев, Украина, 2010. – С. 26.

13. Иванисенко И.Н. Обзор и анализ программного обеспечения для решения задачи балансировки нагрузки компьютерной сети / И.Н. Иванисенко // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: Перша наук.-техн. конф., 13–14 груд. 2010: матер. конф. – Київ, Україна, 2010. – С. 79.

14. Ivanisenko Igor. Load Balancing in Cloud Services Considering the Self-Similar Properties of Incoming Flows / Igor Ivanisenko, Ludmila Kirichenko, Tamara Radivilova // Modern Problems of Radio Engineering, Telecommunications and Computer Science: XII th International Conference, February 25 – March 1, 2014: Abstract Book. – Lviv-Slavske, Ukraine. – P. 571–572.

15. Иванисенко И.Н. Балансировка нагрузки в облачных сервисах с учетом самоподобных свойств входящих потоков / И.Н. Иванисенко // Радиоэлектроника и молодежь в XXI веке: Междунар. молодеж. форум, 14–16 апр. 2014: матер. форума. – Харьков, Україна. – Том 5. – С. 74.

16. Ivanisenko Igor. Investigation of Self-Similar Properties of Additive Data Traffic / Igor Ivanisenko, Ludmila Kirichenko, Tamara Radivilova // Computer Science and Information Technologies (CSIT 2015): X th International Scientific and Technical Conference, 14–17 September 2015: Abstract Book. – Lviv, Ukraine. – P. 169–172. (Входит до міжнар. наукометричної бази Scopus).

17. Иванисенко И.Н. Анализ методов балансировки нагрузки в распределённых системах / И.Н. Иванисенко, Т.А. Радивилова // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: П'ята міжнар. наук.-техн. конф., 23–25 квіт. 2015: матер. конф. – Полтава-Баку-Кіровоград-Харків, Україна. – С. 20–21.

18. Ivanisenko Igor. Survey of Major Load Balancing Algorithms in Distributed System / Igor Ivanisenko, Tamara Radivilova // Інформаційні технології у інноваційному бізнесі (ITIB 2015): II Міжнар. наук.-практ. конф., 7–9 жовтня

2015: матер. конф. – Харків, Україна. – С. 89–92. (Входить до міжнар. наукометричної бази Scopus).

19. Ivanisenko Igor. The multifractal load balancing method / Igor Ivanisenko, Tamara Radivilova // Problems of Infocommunications Science and Technology: Second Intern. Scien. Pract. Conf., October 13–15 2015: Abstract Book. – Kharkiv, Ukraine. – P. 122–123. (Входить до міжнар. наукометричної бази Scopus).

20. Иванисенко И.Н. Исследование зашумленных мультифрактальных рядов / И.Н. Иванисенко, Л.О. Кириченко, А.Ю. Хабачева // Информационные системы и технологии: Междунар. науч.-техн. конф., 21–27 сент. 2015: матер. конф. – Харьков, Украина. – С. 62–63.

21. Ivanisenko Igor. Using cloud storage services in decision support system / Igor Ivanisenko, Yu Kobytka // Проблеми автоматизації: Третя міжнар. наук.-техн. конф., 12–13 лист. 2015: тези доп. – Черкаси-Баку-Бельсько-Бяла-Полтава, Україна. – С. 29.

22. Іванісенко І.М. OPEN SOURCE продукти балансування навантаження / І.М. Іванісенко // Free and Open Source Software: VII Всеукр. наук.-практ. конф., 24–27 лист. 2015: матер. конф. – Харків, Україна. – С. 85.

23. Иванисенко И.Н. Динамическая балансировка фрактального трафика / И.Н. Иванисенко, Л.О. Кириченко, Т.А. Радивилова // Інформаційні технології в металургії та машинобудуванні (ІТММ - 2016): Міжнар. наук.-техн. конф., 29–31 бер. 2016: матер. конф. – Дніпропетровськ, Україна. – С. 58.

24. Kirichenko Lyudmila. Dynamic load balancing algorithm of distributed systems / Lyudmila Kirichenko, Igor Ivanisenko, Tamara Radivilova // Modern Problems of Radio Engineering, Telecommunications and Computer Science: XIIIth Intern. Conf. TCSET'2016, February 23–26, 2016: Abstract Book. – Lviv-Slavsko, Ukraine. – P. 515–518. (Входить до міжнар. наукометричної бази Scopus).

25. Іванісенко І.М. Балансування навантаження з урахуванням рівня дисбалансу системи / І.М. Іванісенко // Сучасні напрямки розвитку інформаційно-комунікаційних технологій та засобів управління: П'ята міжнар. наук.-техн. конф., 21–22 квіт. 2016: матер. конф. – Полтава-Баку-Кіровоград-Харків, Україна. – С. 32.

26. Радивилова Т.А. Алгоритм балансировки самоподобной нагрузки / Т.А. Радивилова, И.Н. Иванисенко, Л.О. Кириченко // Современные информационные и электронные технологии: XVII междунар. науч.-практ. конф., 23–27 мая 2016: труды конф. – Одесса, Украина – С. 115–117.

27. Kirichenko Lyudmila. Investigation of multifractal properties of additive data stream / Lyudmila Kirichenko, Igor Ivanisenko, Tamara Radivilova // 2016 IEEE First International Conference on Data Mining and Processing: First Intern. Conf., 23–27 August 2016: Abstract Book. – Lviv, Ukraine. – P. 305–308. (Входить до міжнар. наукометричної бази Scopus).

28. Кириченко Л.О. Аналіз стану розподіленої системи при самоподібному навантаженні / Л.О. Кириченко, И.Н. Иванисенко, Т.А. Радивилова // Інформаційні управляючі системи та технології: Міжнар. наук.-практ. конф., 20-22 вер. 2016: матер. конф. – Одеса, Україна. – С. 115–117.

29. Kirichenko Lyudmila. Calculation of distributed system imbalance /

Lyudmila Kirichenko, Igor Ivanisenko, Tamara Radivilova // Problems of Infocommunications. Science and Technology (PICS&T-2016): Third Intern. Scien.-Pract. Conf., October 4–6 2016: Abstract Book. – Kharkiv, Ukraine. – P. 156–159 (Входить до міжнар. наукометричної бази Scopus).

АНОТАЦІЯ

Іванісенко І.М. Методи балансування навантаження у розподілених системах з урахуванням самоподібних властивостей вхідних потоків. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.05 – комп'ютерні системи та компоненти. – Харківський національний університет радіоелектроніки, Міністерство освіти і науки України, Харків, 2017.

У дисертаційній роботі запропоновано нове вирішення актуальної науково-практичної задачі розробки методів балансування навантаження, які враховують властивості самоподібності трафіка і використання ресурсів кожного вузла та всієї розподіленої системи.

Удосконалено модель балансування навантаження розподіленої системи, яка враховує обмеження на заданий набір характеристик мережі для різних класів обслуговування та зміну в часі фрактальних параметрів трафіка. Набув подальшого розвитку метод розрахунку дисбалансу ресурсів розподіленої системи, який включає в себе комплексний вимір загального рівня дисбалансу системи та враховує вагові коефіцієнти ресурсів серверів. Проведено дослідження фрактальних властивостей адитивного трафіка, що постає при мультиплексуванні потоків даних. На основі цих результатів запропоновано метод балансування навантаження, який враховує мультифрактальні властивості адитивного комп'ютерного трафіку. Використання даних методів дає можливість підвищити ступінь використання ресурсів системи за рахунок спрямування неоднорідних інформаційних потоків на менш завантажені ресурси.

Ключові слова: розподілені комп'ютерні системи, балансування навантаження, параметри якості обслуговування, дисбаланс обчислювальних ресурсів системи, самоподібність завантаження, мультифрактальний трафік.

АННОТАЦИЯ

Иванисенко И.Н. Методы балансировки нагрузки в распределенных системах с учетом самоподобных свойств входящих потоков. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.05 - компьютерные системы и компоненты. – Харьковский национальный университет радиоэлектроники, Министерство образования и науки Украины, Харьков, 2017.

В диссертационной работе предложено новое решение актуальной научно-

практической задачи разработки методов балансировки нагрузки, учитывающих свойства самоподобия трафика и использование ресурсов каждого узла и всей распределенной системы.

Усовершенствована модель балансировки нагрузки распределенной системы, которая учитывает ограничения на заданный набор характеристик сети для различных классов обслуживания и изменение во времени фрактальных параметров трафика. Состояния серверов описываются объемом свободных ресурсов ЦПУ и объемом свободной оперативной памяти. Также балансировщик и сервера соединяются каналами с задаваемой пропускной способностью. Все значения параметров модели имеют зависимость от времени и ограничения по параметрам качества обслуживания. Такая модель позволяет описывать поведение распределенной сети во времени для различных классов обслуживания входящего трафика, при заданных ограничениях на время ожидания пакета в очереди и количество потерянных пакетов.

Предложен метод расчета дисбаланса системы на основе оценки загрузки узлов распределенной системы. Оценкой загрузки ресурсов узлов являются характеристики загрузки процессора, памяти и пропускной способности каналов. Предложенный метод позволяет проводить расчет загрузки для потоков разных классов обслуживания, как для каждого сервера отдельно, так и для всей системы. Введено комплексное значение дисбаланса нагрузки сервера, учитывающего весовые коэффициенты для процессора, памяти и пропускной способности каналов. Введенные весовые коэффициенты позволяют определить значимость каждой характеристики сервера по отношению друг к другу. Таким образом, метод позволяет вычислить дисбаланс всех серверов системы и эффективность использования ресурсов системы.

Проведено исследование изменения фрактальных свойств для аддитивных информационных потоков в случаях, когда хотя бы один из суммируемых потоков обладает самоподобными или мультифрактальными свойствами. Результаты исследования позволяют определять характеристики трафика при мультиплексировании потоков данных в компьютерной сети. На основе разработанной математической модели предложен динамический метод распределения нагрузки, учитывающий мультифрактальные свойства аддитивного трафика и заданных ограничений на время ожидания и количество потерянных пакетов. Предлагаемый метод балансировки нагрузки на основе анализа мультифрактальных свойств входящего потока обеспечивает статистически равномерное распределение нагрузки на серверах, высокие показатели производительности и пропускной способности, а также снижение времени отклика и количества потерянных данных.

Разработано программное обеспечение для выполнения имитационного моделирования работы балансировщика распределенной системы. Методами имитационного моделирования выполнено сравнение показателей качества обслуживания сети для наиболее востребованных алгоритмов балансировки нагрузки. На основании результатов исследований предложены рекомендации по выбору алгоритмов балансировки нагрузки в компьютерных сетях с

фрактальным трафиком в зависимости от сетевых параметров и характеристик.

Ключевые слова: распределенные компьютерные системы, балансировка нагрузки, параметры качества обслуживания, дисбаланс вычислительных ресурсов системы, самоподобие нагрузки, мультифрактальный трафик.

ABSTRACT

Ivanisenko I.N. Methods of load balancing in distributed systems, taking into account the self-similar properties of input streams. – Manuscript.

A thesis for the candidate degree in technical sciences in the specialty 05.13.05 – computer systems and components. – Kharkiv National University of Radioelectronics, Ministry of Education and Science of Ukraine, Kharkiv, 2017.

The thesis proposes a new solution to the actual scientific and practical problem of developing load balancing methods that take into account the self-similarity of traffic and the use of resources of each node and the entire distributed system.

The distributed load balancing model is improved, which takes into account the limitations on a given set of network characteristics for different classes of service and time variation of fractal traffic parameters. The method of calculating the imbalance of resources in a distributed system has been further developed, which includes a comprehensive measurement of the overall system imbalance level and takes into account server resource weights. Investigation of the fractal properties of additive traffic that occur when multiplexed data streams are carried out. Based on these results, a load balancing method is proposed that takes into account the multifractal properties of additive computer traffic. The use of these methods makes it possible to increase the utilization of system resources by sending heterogeneous information flows to less loaded resources.

Keywords: distributed computer systems, load balancing, quality of service parameters, imbalance of system computing resources, self-similar loading, multifractal traffic.