

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

**ФІЛІМОНЧУК ТЕТЯНА ВОЛОДИМИРІВНА**

УДК 004.942: 004.272.26

**МЕТОДИ ТА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ РОЗПОДІЛУ  
ЗАВДАНЬ В ГЕТЕРОГЕННИХ GRID-СИСТЕМАХ**

05.13.06 – інформаційні технології

Автореферат дисертації на здобуття наукового ступеня  
кандидата технічних наук

Харків – 2017

Дисертацією є рукопис.

Роботу виконано в Харківському національному університеті радіоелектроніки Міністерства освіти і науки України.

Науковий керівник – кандидат технічних наук, доцент

**Волк Максим Олександрович,**

Харківський національний університет  
радіоелектроніки, доцент кафедри електронних  
обчислювальних машин

Офіційні опоненти:

доктор технічних наук, професор

**Гамаюн Ігор Петрович,**

Національний технічний університет  
«Харківський політехнічний інститут»,  
декан факультету комп'ютерних наук і  
програмної інженерії;

доктор технічних наук, професор,

**Малахов Євгеній Валерійович,**

Одеський національний університет  
імені І.І. Мечникова,  
завідувач кафедри математичного забезпечення  
комп'ютерних систем.

Захист відбудеться « 31 » травня 2017 р. о 13:00 годині на засіданні спеціалізованої вченої ради Д 64.052.08 Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, пр. Науки, 14.

З дисертацією можна ознайомитись у бібліотеці Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, пр. Науки, 14.

Автореферат розісланий « 27 » квітня 2017 р.

Учений секретар  
спеціалізованої вченої ради

І.П. Плісс

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** На сьогодні широкого поширення набули ресурсомісткі завдання, які для свого вирішення вимагають великої кількості обчислювальних ресурсів. Для таких задач використовуються територіально розподілені обчислювальні системи. Одним із прикладів таких систем є GRID. В основі GRID-систем лежить забезпечення стабільної роботи певного набору служб, побудованих на відкритих стандартах, і проміжного програмного забезпечення. Будь-яка GRID-система є неоднорідною (гетерогенною), тому що будується на множенні розподілених в просторі обчислювальних ресурсів. Завдання, що надходять на виконання в дану систему, також мають властивість неоднорідності, що ускладнює завдання розподілу. Тому важливе місце в GRID-системах при розподілах відводиться планувальнику (брокеру). Він виступає посередником між постачальниками завдань та обчислювальних ресурсів і реалізується за допомогою певного методу розподілу. Нині існує низка систем управління й розподілу ресурсів, проте планувальники, які реалізовані в них, не дають ефективного способу розподілу завдань, з їх допомогою користувач може лише скористатися одним простим методом.

Ряд досліджень направлено на створення методу розподілу завдань по обчислювальних ресурсів, який дозволив би мінімізувати час простою обчислювальних ресурсів, скоротити обсяги і час переданої інформації між ресурсами. Значну роль в дослідженнях цієї області грають роботи вітчизняних та зарубіжних авторів: А.М. Бершадського, І.В. Бичкова, І.А. Голубєва, Ж.Б. Кальпеевой, Н.В. Покусіна, Є.Ю. Селіверстова, Б.А. Телесніна, D. Agrawal, A. Amar, R. Bolze, E. Voix, B. Amedro, V. Vodnartchouk, L. Baduel, H. Zhao, R. Sakellariou. Тому розробка методів та технологій розподілу завдань (з урахуванням ряду додаткових параметрів) на обчислювальних ресурсах є актуальною задачею, оскільки вибір обчислювальних ресурсів визначає ефективність використання всієї системи. Правильний вибір впливає на час простою обчислювальних ресурсів, скорочує час і обсяги переданої між пристроями інформації тощо.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційна робота виконувалася відповідно до плану науково-дослідних робіт Харківського національного університету радіоелектроніки в рамках таких держбюджетних тем:

- «Розробка структури Харківського ресурсно-операційного GRID-центру та його ресурсів», договір №9 (28.09.07-31.10.07, 01.12.07-31.12.07) між ХНУРЕ і «ІПСА» НТУУ «КПІ», що виконувалася на підставі договору «ІПСА» НТУУ «КПІ» з Міністерством освіти і науки України №ІТ/506-2007, Державної програми «Інформаційні та телекомунікаційні технології в освіті і науці» на 2006-2010 рр. (№ДР 0107U010616);

- «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 рр. (№ДР 0108U008261).

У рамках вказаних тем здобувачем проведено аналіз концепцій, архітектури, конфігурацій обчислювальних ресурсів, програмного забезпечення сучасних національних проектів GRID, розглянуті цілі та завдання групи NA4. Здійснено підключення до суперкомп'ютерного центру «КПІ» та тестування потужності бібліотек інтерфейсу паралельного програмування PVM на прикладі ряду задач. Результати наукових досліджень використані в науково-технічних звітах про НДР №08-22, НДР №9, НДР №08-22/9.

**Мета та задачі дослідження.** Метою дисертаційної роботи є розробка методів та інформаційної технології розподілу завдань в гетерогенних GRID-системах, які дозволяють зменшити час простою обчислювальних ресурсів системи за рахунок скорочення часу виконання вхідного пулу завдань.

Згідно зі сформульованою метою в дисертаційній роботі необхідно розв'язати такі задачі:

- аналіз існуючих методів розподілу завдань в розподілених обчислювальних системах;
- модифікація математичної моделі розподілу завдань в гетерогенних GRID-системах;
- модифікація методу розподілу завдань (Backfill) на обчислювальні ресурси;
- розробка методу пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів;
- розробка інформаційної технології розподілу завдань на обчислювальні ресурси на основі запропонованої математичної моделі розподілу завдань;
- проведення ряду експериментів з розподілу завдань на основі запропонованої інформаційної технології в системі імітаційного моделювання;
- апробація моделі та методів при вирішенні практичних завдань.

**Об'єкт дослідження** – процес розподілу завдань в гетерогенних GRID-системах.

**Предмет дослідження** – методи планування та інформаційна технологія розподілу завдань в гетерогенних GRID-системах.

**Методи дослідження** базуються на використанні: теорії множин (для розробки моделі розподілу завдань в гетерогенних GRID-системах, моделей завдань та обчислювальних ресурсів), загальної теорії систем (для дослідження та розробки методів розподілу завдань в розподілених обчислювальних системах), теорії імітаційного моделювання (для моделювання процесу розподілу завдань в гетерогенних GRID-системах).

**Наукова новизна отриманих результатів.** Основні результати, які визначають наукову новизну дисертаційної роботи:

*а) вперше*

- запропоновано метод пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простим обчислювальних ресурсів, який, на відміну від існуючих, використовує узагальнену критерій оцінки завдання та імітаційне середовище моделювання GRASS, що дозволяє підвищити ефективність використання обчислювальних ресурсів GRID-системи за рахунок скорочення часу виконання пулу завдань та простою обчислювальних ресурсів;

*б) удосконалено*

- метод розподілу завдань на обчислювальні ресурси Backfill з урахуванням трафіку всередині завдання, який, на відміну від існуючих методів, враховує інтенсивність та обсяг потоків даних між задачами у завданні, що дозволяє підвищити ефективність використання GRID-системи за рахунок скорочення часу виконання пулу завдань на розподілених гетерогенних обчислювальних ресурсах;

*в) набули подальшого розвитку*

- моделі подання завдань і ресурсів у GRID-системі, які, на відміну від існуючих, враховують обсяг мережного трафіку та затримку передачі даних в процесі розподілу і виконання завдань, що дозволяє скоротити час виконання завдання в системі за рахунок усунення втрат за часом, які викликані обміном даних між окремими задачами у завданні.

**Практичне значення отриманих результатів.** Практична значимість отриманих теоретичних результатів дисертаційної роботи підтверджено поліпшенням продуктивності GRID-системи, за рахунок скорочення простою обчислювальних ресурсів. Зокрема, практичне вирішення теоретичних досліджень полягає у такому.

Запропоновано інформаційну технологію розподілу завдань на обчислювальні ресурси в GRID-системі, яка на підставі використання множини методів розподілу дозволяє здійснювати вибір розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простим обчислювальних ресурсів.

Запропоновано блок, який відповідає за формування додаткових параметрів для ефективного розподілу завдань. Він включає 2 модуля: модуль згортки кортежу та модуль аналізу зв'язності. Модуль згортки кортежу здійснює обчислення узагальненого критерію оцінки для кожного завдання, що дозволяє більш продуктивно управляти процесом розподілу завдань на обчислювальні ресурси. Модуль аналізу зв'язності дозволяє здійснювати підбір обчислювальних ресурсів з урахуванням скорочення часу, що витрачається на обмін між задачами у завданні.

Додана база даних (БД), яка фіксує інформацію про кожне запущене завдання: час виконання завдання на кожному з обчислювальних ресурсів, дані про відмови, дані про стан ресурсів. Завдяки використанню запропонованої БД планувальник має можливість отримувати результати

попередніх розподілів завдань за різними методами розподілу, що дозволяє обрати такий план розподілу, який матиме вигоду за рядом критеріїв, заздалегідь встановлених користувачем. У процесі моделювання в імітаційному середовищі GRASS всі дані з log-файлів додаються до БД, яку можна використовувати для аналізу статистичної інформації в міру необхідності.

Результати дисертаційної роботи впроваджені у Радіоастрономічному інституті НАН України (відділ космічної радіофізики), м. Харків (акт від 30.11.2015 р.) та Харківському національному університеті радіоелектроніки, кафедра електронних обчислювальних машин, м. Харків в процесі проведення лекційних занять і лабораторних робіт з курсів «Паралельне моделювання на НРС системах» та «Інтерфейси паралельного програмування» (акт від 10.01.2017 р.).

**Особистий внесок здобувача.** Всі наукові результати дисертаційної роботи, що виносяться на захист, отримані автором самостійно. У роботах, опублікованих у співавторстві здобувачеві належать такі результати: у роботі [1] класифікація з точки зору областей застосування методів для різних класів завдань та конфігурацій обчислювальних ресурсів; у роботі [2] архітектура системи моніторингу трафіку в комп'ютерній мережі, яка забезпечує ефективний розподіл завдань за обчислювальними ресурсами GRID-системи; у роботі [3] моделі подання завдань та обчислювальних ресурсів у GRID-системі, модуль розподілу завдань на обчислювальні ресурси, який підтримує можливість проведення серії експериментів; у роботі [4] класифікація сучасного прикладного програмного забезпечення, яке використовується у GRID-системах для розподілу завдань; у роботі [5] формальний опис процесів у розподілених системах імітаційного моделювання, розглянуто процеси управління локальними та глобальними даними моделей; у роботі [6] інформаційна технологія розподілу завдань для GRID-систем, яка ґрунтується на використанні імітаційного середовища моделювання GRASS, а також метод пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простом обчислювальних ресурсів; у роботі [7] удосконалений метод розподілу завдань Backfill (mod) з урахуванням трафіку всередині завдання; у роботі [8] досліджено сучасний стан штучних імунних систем і можливостей їх використання в розподілених обчислювальних системах типу GRID; у роботі [9] класифікація методів розподілу ресурсів у гомогенних GRID-структурах; у роботі [10] досліджено найбільш поширені методи планування, які дозволяють оптимізувати час виконання послідовності завдань; у роботі [11] механізм підключення модулів розподілу заявок до імітаційної системи моделювання GRID на основі плагінів; у роботі [12] структура імітаційного середовища GRASS, проаналізовано склад та функціональне призначення модулів системи; у роботі [13] програмні реалізації декількох методів розподілу завдань на обчислювальні ресурси, наведені результати аналізу ефективності даних методів; у роботі [14] обґрунтовано необхідність розробки та вдосконалення

методів і засобів віртуалізації при моніторингу трафіку комп'ютерних мереж; у роботі [15] наведені результати аналізу сучасного стану і перспективи розвитку GRID-систем, запропоновано розмежування існуючого прикладного програмного забезпечення за критеріями, які класифікують GRID-систему залежно від поставленої перед користувачем мети; у роботі [16] наведено обґрунтування доцільності введення узагальненого критерію для оцінки завдань, який дозволяє підібрати ресурси обчислювальної системи з урахуванням вимог постачальників завдань та власників ресурсів; у роботі [17] запропоновано підхід, який ґрунтується на використанні інформаційної технології розподілення завдань, що дозволяє використовувати сукупність заданих методів розподілу для отримання інформації про розподіл потоку завдань; у роботі [18] запропоновано інформаційну технологію розподілу завдань на обчислювальні ресурси GRID-системи, обґрунтовано використання модуля, який формує додаткові параметри для розподілу завдань; у роботі [19] обґрунтовано доцільність використання існуючих технологій з точки зору підвищення продуктивності системи, описано функціонування технології розподілу завдань на обчислювальні ресурси; у роботі [20] запропоновано інформаційну технологію розподілу завдань на обчислювальні ресурси, яка дозволяє отримати план розподілу з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів GRID-системи; у роботі [21] удосконалена інформаційна технологія розподілу завдань на обчислювальні ресурси в GRID-системі, яка дозволяє здійснювати вибір плану розподілу з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів, а також реалізує можливість розширення множини методів розподілу за рахунок додавання нових програмних модулів; у роботі [22] запропоновано алгоритм функціонування імітаційної середовища моделювання GRASS, яке відтворює всі процеси, що відбуваються у реальній GRID-системі; у роботі [23] запропонована інформаційна технологія розподілу завдань на обчислювальні ресурси з великим обсягом вхідних та вихідних даних для GRID-систем, яка заснована на ряді існуючих технологій, що працюють у комплексі та дозволяють знизити трудомісткість процесів використання обчислювальних ресурсів; у роботі [24] запропоновано використання інформаційної технології розподілу завдань для обробки великих масивів вхідних даних у віртуальних приватних хмарах; у роботі [25] запропонована реалізація методу Backfill (mod), який, на відміну від існуючого (Backfill), враховує інтенсивність і обсяг потоків даних між задачами у завданні, що дозволяє підвищити ефективність використання GRID-системи за рахунок зменшення часу виконання пулу завдань на розподілених гетерогенних обчислювальних ресурсах.

**Апробація результатів дисертації.** Основні положення дисертаційної роботи доповідалися на таких міжнародних конференціях і форумах:

- 8-й, 9-й Міжнародній науково-технічній конференції «Проблеми інформатики і моделювання» (Харків: 2008, 2009 pp.);

- 10-й, 11-й, 12-й Міжнародній науково-технічній конференції «Системний аналіз та інформаційні технології» (Київ: 2008, 2009, 2010 рр.);

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

- 1-й, 3-й, 5-й Міжнародній науково-технічній конференції «Сучасні напрямки розвитку інформаційно-комунікаційних технологій і способів управління» (Київ – Харків 2010 р., Полтава – Белгород – Харків – Київ – Кіровоград 2013 р., Полтава – Баку – Кіровоград – Харків 2016 р.);

- 3-й, 4-й, 5-й Міжнародній науково-практичній конференції «Інформатика, математическое моделирование, экономика» (Смоленськ: 2013, 2014, 2015 рр.);

- 1-й, 2-й, 3-й, 4-й Міжнародній науково-технічній конференції «Проблеми інформатизації» (Черкаси – Київ – Тольятті – Полтава 2013 р., Київ – Полтава – Катовіце – Париж – Белгород – Черкаси – Харків 2014 р., Черкаси – Баку – Бельсько-Бяла – Полтава 2015 р., Черкаси – Баку – Бельсько-Бяла – Полтава 2016 р.);

- 20-й Международный молодежный форум «Радиоэлектроника и молодежь в XXI веке» (Харків: 2016 р.);

- 5-й Міжнародній науково-технічній конференції «Інформаційні системи та технології» (Коблево: 2016 р.).

**Публікації.** За результатами дисертаційних обчислень опубліковано 25 наукових праць, у тому числі: 7 статей, які зазначені в переліку фахових видань України з технічних наук (всі наукові праці входять до міжнародних науково-метричних баз – 1 наукова праця у Scopus) та 18 тез доповідей конференцій

**Структура дисертації.** Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел зі 150 найменувань на 17 сторінках та 2 додатків на 8 сторінках, а також містить 21 рисунок і 4 таблиці (рисунки та таблиці, що займають окрему площу на 2 сторінках). Загальний обсяг роботи складає 180 сторінок, включаючи 131 сторінку основного тексту.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтовано актуальність обраної теми дисертації, сформульовано мету та задачі дослідження, визначено об'єкт, предмет і методи дослідження, визначені наукова новизна та практичне значення отриманих результатів, а також особистий внесок здобувача в роботи, виконані у співавторстві, наведено відомості про апробацію результатів дисертації та кількість публікацій за темою дисертаційної роботи.

У першому розділі на базі вивчення літературних джерел проведено



аналіз предметної області. Розглянуто питання розвитку розподілених систем як середовища для розв'язання задач великої обчислювальної складності. Наведено огляд технологій розподілених обчислень та Cloud-технологій, існуючих планувальників завдань, мов опису завдань в GRID-системах, а також сформульована мета та задачі дослідження.

У другому розділі досліджена існуюча модель розподілу завдань у GRID-системі  $G = \{R, Z\}$ , яка будується на основі двох множин: множині обчислювальних ресурсів  $R$ , множині завдань  $Z$ , а також методі розподілу  $q$ .

Завдання, які надходять у GRID-систему, утворюють потік  $\{Z_i, i = 1, 2, \dots, M\}$ , де  $i$  – порядковий номер завдання, а  $M$  – кількість завдань. Кожне завдання містить ряд параметрів, необхідних для їх запуску в даному середовищі (1):

$$Z_i = \{ar_i^z, os_i^z, pc_i^z, ps_i^z, ms_i^z, dc_i^z, pr_i^z\}, \forall i = 1..M, \quad (1)$$

де  $ar_i$  – архітектура процесора;  $os_i$  – операційна система;  $pc_i$  – кількість процесорів;  $ps_i$  – швидкодія процесорів;  $ms_i$  – обсяг оперативної пам'яті;  $dc_i$  – доступний обсяг вінчестера;  $pr_i$  – пріоритет завдання.

Ресурси в GRID-системі також утворюють множину  $\{R_j, j = 1, 2, \dots, N\}$ , де  $j$  – номер обчислювального ресурсу, а  $N$  – число ресурсів у GRID-системі.

Обчислювальний ресурс, який надходить у GRID-систему, має ряд характеристик, які представлені кортежем (2):

$$R_j = \{ar_j^r, os_j^r, pc_j^r, ps_j^r, ms_j^r, dc_j^r\}, \forall j = 1..N. \quad (2)$$

У ході аналізу моделі розподілу завдань в гетерогенних GRID-системах було виявлено що, в процесі розподілу не враховується ряд параметрів, які можуть підвищити ефективність використання її ресурсів (пропускна здатність каналу, затримка часу передачі пакету по каналу). Також під час розподілу завдань на обчислювальні ресурси планувальники не враховують характер пакета завдання (зв'язність завдань у завданні), що дозволить зменшити час перебування завдання в системі. Проведений аналіз показав, що метод Backfill, також не враховує час, витрачений на передачу вхідних і вихідних даних завдання, і під час його використання немає важелів впливу на хід розподілу конкретного завдання, тому що розподіл здійснюється за принципом «перший відповідний».

Обґрунтована необхідність розробки нових методів розподілу завдань у GRID-системах. Для цього у роботі вирішується науково-технічна задача розробки нових методів до розподілу завдань, що орієнтовані на використання нової інформаційної технології розподілу завдань, який дозволяє враховувати й усувати недоліки в існуючих рішеннях задачі розподілу завдань, які були виявлені в ході порівняльного аналізу.

У третьому розділі запропонована модифікована модель GRID-системи за рахунок введення в неї множини методів розподілу (3):

$$Q_k = \{mn_k, lp_k\}, \forall k = 1..k, \quad (3)$$

де  $mn$  – алгоритм розподілу завдань за обчислювальними ресурсами;  $lp$  – перелік вхідних параметрів, що враховуються в процесі розподілу.

Модифікована модель подання завдань і ресурсів у GRID-системі будується на підставі трьох множин: множині обчислювальних ресурсів  $R$ , множині завдань  $Z$  та множині методів розподілу  $Q$ :  $G = \{R, Z, Q\}$ .

Також пропонується розширити кортежі (1) і (2) за рахунок параметрів, які дозволять усунути зазначені у другому розділі недоліки та підвищити ефективність використання обчислювальних ресурсів GRID-системи (4, 5):

$$Z_i = \{ar_i^z, os_i^z, pc_i^z, ps_i^z, ms_i^z, dc_i^z, pr_i^z, ca_i^z, rt_i^z\}, \forall i = 1..M, \quad (4)$$

$$R_j = \{ar_j^r, os_j^r, pc_j^r, ps_j^r, ms_j^r, dc_j^r, bw_j^r, d_j^r\}, \forall j = 1..N, \quad (5)$$

де  $ca_i$  – критерій зв'язності задач у завданні;  $rt_i$  – час виконання завдання;  $bw_j$  – пропускна здатність каналу;  $d$  – затримка часу передачі пакету.

На цей час існує ряд завдань, які необхідно аналізувати не за одним параметром, що в свою чергу викликають труднощі у процесі розподілу. У таких випадках раціонально використовувати методи розв'язання багатокритеріальних задач оптимізації. Застосування цих методів передбачає об'єднання ряду параметрів у єдиний критерій. У роботі пропонується згортка параметрів у єдиний критерій оцінки завдання, за рахунок розподілу параметрів на два класи: якісні та кількісні. Узагальненим критерієм для якісних параметрів запропоновано прийняти мультиплікативний критерій виду (6), де окремі критерії  $P_i$  визначаються виразом (7):

$$F_{\text{quality}} = \prod_{i=1}^n P_i. \quad (6)$$

$$P_i = \begin{cases} 1, & \text{якщо } P_i^Z = P_i^R \\ 0, & \text{якщо } P_i^Z \neq P_i^R \end{cases}. \quad (7)$$

Узагальненим критерієм для кількісних параметрів прийнято адитивний критерій виду (8):

$$F_{\text{quantity}} = \begin{cases} \sum_{i=1}^n C_i P_i, & \text{якщо } P_i \leq 1, C_i > 0 \\ 0, & \text{якщо } P_i > 1 \end{cases}. \quad (8)$$

де  $P_i$  – нормоване значення  $i$ -го параметра,  $P_i = \frac{P_i^Z}{P_i^R}$ .

Підсумковим узагальненим критерієм оцінки завдання, що враховує якісні та кількісні параметри, є вираз (9), який дозволяє відібрати множину ресурсів, на яких може бути запущено завдання, а також показує, яку частину ресурсу займає завдання в процесі виконання.

$$F = F_{\text{quality}} \cdot F_{\text{quantity}}. \quad (9)$$

На сьогодні у розподілених системах широко поширений метод Backfill, який має безліч переваг і дозволяє запобігти дефрагментації обчислювальних ресурсів. Проте більш раціонально його використовувати або у зв'язці з додатковими евристикami розподілу, або його модифікацію, що враховує трафік обміну даними між задачами у завданні, тобто під час розподілу завдань слід враховувати оцінку сумарної затримки часу передачі пакету по каналу і пропускну здатність каналу. Ці два параметри дозволяють скоротити час на пересилання результатів між окремо взятими ресурсами і зменшити трафік у мережі за рахунок розвантаження каналів зв'язку.

Вхідними даними для запропонованої модифікації методу Backfill є:

- множина завдань у черзі із зазначенням вимог до обчислювальних ресурсів  $\{Z_i, i = 1, 2, \dots, M\}$ ;

- кількість завдань, які виконуються у поточному часі, із зазначенням задіяних обчислювальних ресурсів та очікуваним часом завершення виконання  $z_i^{\text{run}} \in \{Z_i, i = 1, 2, \dots, M\}$ ;

- рекомендований список ресурсів, які були відібрані для запуску кожного завдання  $\{R_j^r, j = 1, 2, \dots, K\}$  за узагальненим критерієм оцінки завдання.

Результатом роботи модифікованого методу Backfill є отримання для кожного завдання  $Z_i$  вікна для запуску.

На першому кроці резервування відбувається сортування списку запущених завдань  $z_i^{\text{run}}$  відповідно до їх очікуваним часом закінчення  $rt$ ,  $\{Z/\text{sort}(rt_i, \forall i = 1..M)\}$ . Час виконання завдань зі списку  $t$  розбивається на періоди, відповідні часу закінчення завдань  $\tau$ . Для кожного періоду записується кількість CPU, які використовуються у цьому періоді,  $t \subset \tau, \forall \tau := pr/R_j, \forall j = 1..N$ . Як результат – профіль використання  $Pr of \subset \{\tau\}$ .

На другому кроці відбувається такий розподіл завдання:

- завдання з черги упорядковано відповідно до критерію зв'язності;

- на підставі множини  $\{R_j^r, j = 1, 2, \dots, K\}$  формується список вікон, які можуть бути використані для запуску завдання –  $W_i^{\text{run}} ;$   
 $i \in 1..M$

- для кожного отриманого вікна формується перелік доступних

обчислювальних ресурсів (CPU):

а) для кожної пари (CPU-брокер) визначається затримка часу передачі пакету по каналу ( $\delta_n$ );

б) для кожної пари (CPU-брокер) визначається пропускна здатність каналу ( $C_n$ );

- для запуску задач із завдання ( $a, z_a \in Z_i$ ) набирається необхідна кількість CPU, мінімізуючи один з параметрів:

а) якщо завдання має великий обсяг вхідних даних і відповідно великий обсяг вихідних даних, то відбираються канали зв'язку з найбільшими значеннями пропускної здатності –  $C_n$ ;

б) якщо для запуску завдання не потрібна передача великого обсягу інформації, то відбираються канали зв'язку з найменшим значенням  $\delta_n$ ;

- за кожним отриманим вікном підсумовуються значення попарних відстаней і визначається найменше з них –  $D_{\min} = \min_{i \in 1..M} D_{W_i^{\min}}$  – це найкращий результат розподілу.

Після того для завдання було підібрано вікно на виконання, відбувається оновлення профілю з урахуванням того, що ці процесори будуть зайняті.

Актуальною задачею у GRID-системі є задача розподілу завдань за наявними обчислювальними ресурсами. У роботі вперше запропоновано метод пошуку розподілу з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів. Імітаційне середовище GRASS дозволяє промоделювати роботу методів розподілу, які реалізовані в модулі Algorithm Loader, та обрати з них той, що мінімізує ряд таких критеріїв як час виконання пулу завдань, час простою обчислювальних ресурсів та час перебування заявок у черзі. Ефективність вибору розподілу з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів для вхідного пулу завдань визначається системою критеріїв і правилами їх порівняння. Вибір розподілу заснований на аналізі даних імітаційних експериментів у середовищі GRASS та порівнянні значень запропонованих критеріїв по кожному з методів. Метод, який дає кращий результат за сукупністю всіх критеріїв, буде обраний результуючим для розглянутого пулу завдань.

Метод пошуку розподілу з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів включає в себе такі етапи.

Етап 1: завантажити пули завдань і ресурсів з реальних GRID-систем:

-  $Z_i = \{ar_i^z, os_i^z, pc_i^z, ps_i^z, ms_i^z, dc_i^z, pr_i^z, ca_i^z, rt_i^z\}$ ,  $\forall i = 1..M$ ;

-  $R_j = \{ar_j^r, os_j^r, pc_j^r, ps_j^r, ms_j^r, dc_j^r, dc_j^r, bw_j^r, d_j^r\}$ ,  $\forall j = 1..N$ .

Етап 2: обрати методи розподілу з множини  $Q_k = \{mn_k, lp_k\}$ ,  $\forall k = 1..K$ .

Етап 3: запустити процес моделювання в середовищі GRASS за кожним методом розподілу ( $mn \in Q$ ):

а) на першому кроці як результат отримати множину планів розподілу

$(Plan_k) f : Q_k \rightarrow Plan_k, \forall k = 1..K;$

б) на другому кроці як результат отримати множину часів виконання планів ( $T_k$ ) за кожним методом розподілу ( $mn \in Q$ )  $f : Plan_k \rightarrow T_k, \forall k = 1..K;$

в) на третьому кроці визначити відсоток простою обчислювальних ресурсів і середній час очікування завдання в черзі за кожним із запропонованих методів розподілу.

Етап 4: для всіх методів розподілу ( $mn \in Q$ ) проаналізувати log-файли  $f : Q_k \rightarrow T_k, \forall k = 1..K.$

Етап 5: обрати розподіл з мінімальним часом виконання пулу завдань та мінімальним простим обчислювальних ресурсів на підставі прийнятих правил і обмежень:  $plan = \min t_k, \forall k = 1..K.$

Запропоновано інформаційну технологію розподілу завдань на обчислювальні ресурси для GRID-систем (рис. 1), яка впроваджена в імітаційне середовище моделювання GRASS.

Дане рішення дозволяє проводити експерименти, які реалізують різні методи розподілу, з подальшим вибором найбільш ефективного рішення на основі збору, аналізу та інтерпретації результатів моделювання. Інформаційна технологія містить такі етапи.

Етап 1: завдання параметрів експерименту. Для запуску експерименту налаштовується конфігураційний файл `plugins.xml`, який описує взаємозв'язки між іменами модулів у системі, назвами файлів і бібліотек. Результатом роботи першого етапу є зчитування та декодування конфігураційного файлу `plugins.xml`, завантаження основних плагінів системи для запуску процесу моделювання.

Етап 2: завантаження в середовище моделювання інформації про обчислювальні ресурси та завдання. Будь-яке завдання, яке надходить у середовище моделювання, можна розбити на дві складові: характеристики завдання та тіло завдання (у вигляді `.exe` файлу, файли вхідних даних, БД та ін.). На початковому етапі для системи головне значення має інформація про завдання (4), яка оперує даними, необхідними для підбору обчислювальних ресурсів, придатних для виконання завдання. Водночас із завданнями в систему надходить інформація про доступні обчислювальні ресурси (зв'язок 2, рис. 1).

Етап 3: формування додаткових параметрів для найбільш ефективного розподілу. Усі завдання додаються до черги завдань (зв'язок 5, рис. 1), і паралельно відбувається передача інформації для кожного завдання (у вигляді кортежу 4) у модуль згортки кортежу та модуль аналізу зв'язності (зв'язок 6, рис. 1). Модуль згортки кортежу здійснює обчислення узагальненого критерію оцінки для кожного завдання, який дозволить більш продуктивно управляти процесом розподілу завдань на обчислювальні ресурси і покаже, яку частину ресурсу займає завдання в процесі його виконання. Водночас на даний модуль відбувається передача інформації про обчислювальні ресурси, які наявні у системі (зв'язок 7, рис. 1).

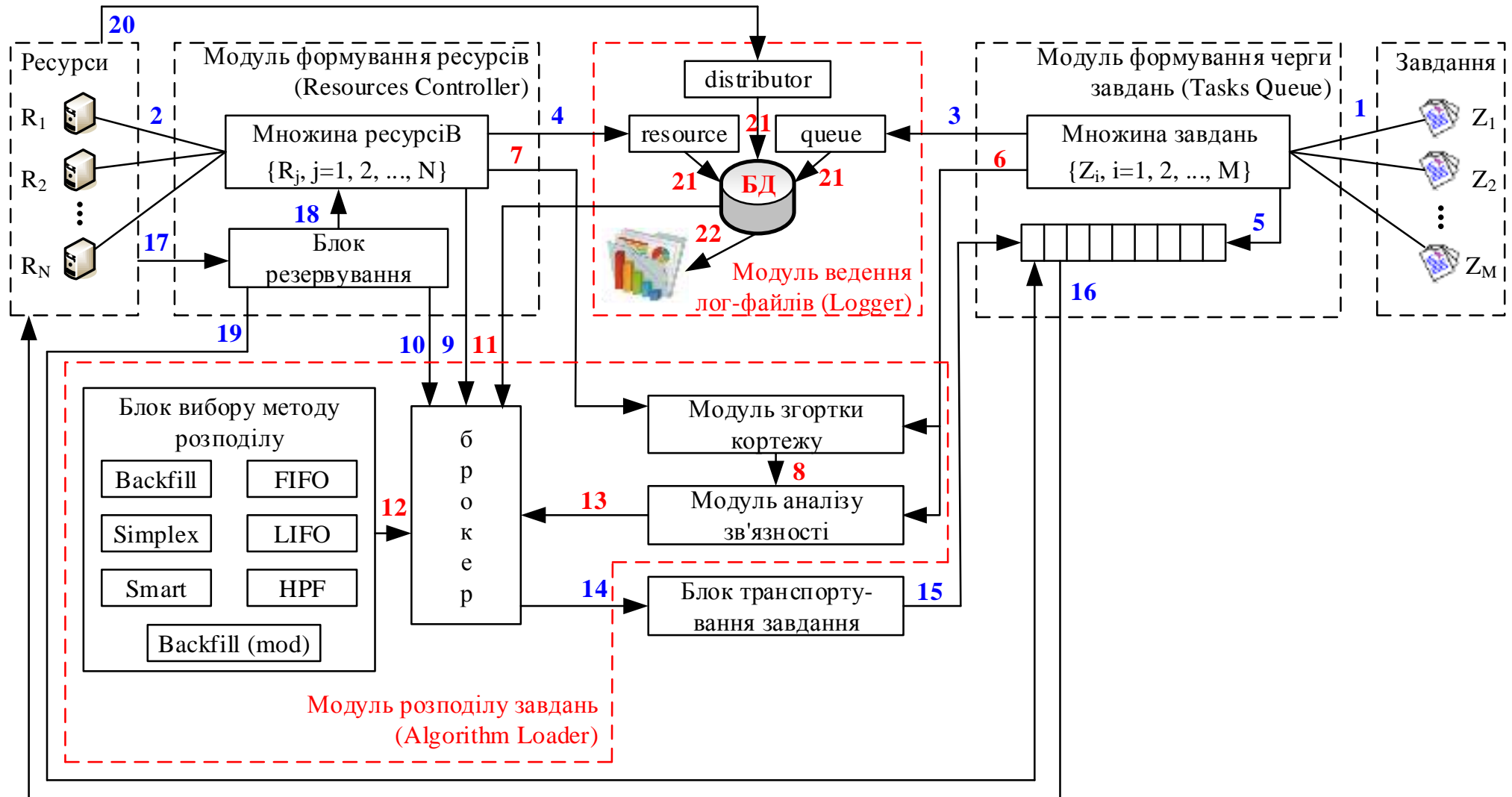


Рисунок 1 – Інформаційна технологія розподілу завдань у гетерогенній GRID-системі, яка використовує імітаційне середовище моделювання GRASS

Результатом роботи модуля згортки кортежу є перелік ресурсів, на яких завдання може бути розподілено.

Далі інформація надходить у модуль аналізу зв'язності (зв'язок 8, рис. 1), де відбувається аналіз задач із завдання. Якщо задачі у завданні мають високу зв'язність або для даного завдання необхідне пересилання великого обсягу вхідних / вихідних даних підбір обчислювальних ресурсів буде спрямований на зменшення часу, який пов'язаний з пересиланням даних між задачами в завданні.

Етапи 4-6: робота модуля розподілу завдань. На етапі 4 у модуль Algorithm Loader надходить інформація, що дозволяє прийняти рішення для розподілу завдання на обчислювальні ресурси, виконання на яких буде для нього оптимальним, як за апаратними показниками, так і за характеристиками продуктивності.

Для розподілу брокеру необхідно отримати таку інформацію:

- інформацію про обчислювальні ресурси (6), які на даний момент присутні у системі (зв'язок 9, рис. 1);
- інформацію про обчислювальні ресурси, які на даний момент зарезервовані (зв'язок 10, рис. 1) і час їх звільнення;
- інформацію з БД про попередні запуски завдання (зв'язок 11, рис. 1);
- метод розподілу (зв'язок 12, рис. 1);
- інформацію про зв'язність задач у завданні (зв'язок 13, рис. 1).

На основі отриманих даних здійснюється побудова плану розподілу завдань на обчислювальних ресурсах GRID-системи. Результатом роботи етапу 4 є множина планів розподілу за обраними методами розподілу.

Далі (етап 5) отримані плани розподілу запускаються у середовищі моделювання GRASS на виконання. Результатом роботи даного етапу є множина часів виконання пулу завдань з кожного методу розподілу, а також ряд значень критеріїв, які обчислюються за кожним методом.

Вибір розподілу з мінімальним часом виконання пулу завдань та мінімальним простим обчислювальних ресурсів здійснюється на етапі 6 на підставі прийнятих заздалегідь критеріїв відбору (зв'язок 14, рис. 1). Будь-яке завдання, яке надійшло в середовище GRASS, буде розподілено, винятком може бути тільки випадок, коли в системі не виявиться необхідних обчислювальних ресурсів. У даному випадку завданням буде відмовлено у виконанні, і воно буде повернуто постачальнику для зміни вимог до обчислювальних ресурсів.

Етап 7: процес пересилання завдання на обчислювальний ресурс або ресурси, зазначені в плані розподілу. З блоку транспортування здійснюється запит у чергу завдань (зв'язок 15, рис. 1), який відповідно до плану розподілу обирає конкретне завдання і привласнює йому стан Running. Далі завдання (друга складова) відправляється на обраний ресурс для подальшого його запуску (зв'язок 16, рис. 1). Якщо в системі сталися події з ресурсами, то система оповіщається про це (зв'язок 17, рис. 1). Блок резервування відсилає

інформацію системі, про те, що обчислювальний ресурс не може бути використаний для подальшого запуску до моменту закінчення виконання на ньому завдання (зв'язок 18, рис. 1). Коли знову відбудуться дії з ресурсами (наприклад, звільнення ресурсу), блок резервування отримає інформацію про подію, що відбулася (зв'язок 17, рис. 1) і сповістить систему про можливість використання ресурсу для його подальшого використання в розподілі (зв'язок 18, рис. 1), а також відправить повідомлення в чергу завдань для видалення завдання з черги в зв'язку з закінченням його виконання (зв'язок 19, рис. 1).

Етап 8: збір статистичної інформації про експеримент. У середовищі GRASS за збір статистики відповідає модуль ведення log-файлів (Logger). Даний модуль надає можливість гнучкого та централізованого ведення логів для всіх плагінів середовища моделювання GRASS. У системі є кілька типів log-файлів, які фіксують ряд дій, що дозволяє за необхідності швидко отримати потрібну інформацію з БД.

З надходженням завдань до системи відбувається автоматичний запис даних про завдання у файл `queue.log` (зв'язок 3, рис. 1), додавання обчислювальних ресурсів у систему призводить до формування файлу `resources.log` (зв'язок 4, рис. 1), у момент запуску завдання на обчислювальному ресурсі або закінчення виконання завдання відбувається фіксація інформації з даної дії у файлі `gaspred.log` (зв'язок 20, рис. 1), тобто відбувається дозапис даного файлу. Всі записи з log-файлів пересилаються в БД системи (зв'язок 21, рис. 1), дані з якої можуть бути використані в подальшому для здійснення вибірки за заданими параметрами або для графічного відображення інформації за певний проміжок часу (зв'язок 22, рис. 1). Процес моделювання з кожного методу розподілу триває до тих пір, поки в черзі присутні завдання.

У четвертому розділі наведена практична реалізація запропонованої інформаційної технології та показано, що після її впровадження в імітаційне середовище моделювання GRASS стало можливим:

- паралельне моделювання реального процесу розподілу завдань на обчислювальні ресурси в GRID-системі за різними методами розподілу;
- вибір плану розподілу з мінімальним часом виконання пулу завдань та мінімальним простим обчислювальних ресурсів;
- врахування зв'язності задач у завданні та класу завдань в процесі розподілу їх на обчислювальні ресурси.

В ході дослідження розробленої інформаційної технології було проведено ряд експериментів з розподілу завдань на обчислювальні ресурси для різних методів розподілу та з різними пулами завдань і обчислювальних ресурсів. Вони довели те що, застосування одного методу розподілу завдань в планувальнику є менш ефективним, тому що була отримана залежність результатів розподілу (час виконання пулу завдань) від класу задач. На сьогодні не існує методу розподілу, який для будь-якого пулу завдань мав би оптимальний варіант розподілу за часом виконання. Однак, якщо відомі



характеристики обчислювальних ресурсів GRID-системи, вимоги вхідного потоку завдань, то можна провести моделювання, обрати метод розподілу з мінімальним часом виконання пулу завдань та запропонувати його реальній GRID-системі в якості найбільш ефективного рішення. Результати, отримані в ході експериментів, свідчать про скорочення часу виконання пулу завдань до 24% і підвищення ефективності використання обчислювальних ресурсів до 32% для ряду методів розподілу, реалізованих в імітаційному середовищі GRASS.

У додатках наведено акти про впровадження результатів дисертаційної роботи у виробництво та навчальний процес.

## ВИСНОВКИ

У дисертаційній роботі розв'язана нова актуальна науково-практична задача розробки методів та створення на їх основі інформаційної технології розподілу завдань в гетерогенних GRID-системах. Внаслідок виконання роботи отримані нові наукові та практичні результати.

1. Модифіковано математичну модель розподілу завдань у GRID-системі за рахунок введення множини методів розподілу і додаткових параметрів у поданні обчислювальних ресурсів і завдань. Врахування зв'язності задач у завданні дозволяє скоротити час виконання пулу завдань у системі за рахунок усунення втрат за часом, які викликані обміном даних між окремими задачами в завданні. Множина методів розподілу дозволяє проведення серії експериментів для подальшого вибору плану розподілу з мінімальним часом виконання пулу завдань та мінімальним простом обчислювальних ресурсів.

2. Модифіковано метод розподілу завдань на обчислювальні ресурси Backfill з урахуванням трафіку всередині завдання, який на відміну від існуючих методів враховує інтенсивність і обсяг потоків даних між задачами в завданні, що дозволяє підвищити ефективність використання GRID-системи за рахунок зменшення часу виконання пулу завдань на розподілених гетерогенних обчислювальних ресурсах.

3. Запропоновано метод пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простом обчислювальних ресурсів, який, на відміну від існуючих, використовує узагальнений критерій оцінки завдання та імітаційне середовище моделювання GRASS, що дозволяє підвищити ефективність використання обчислювальних ресурсів GRID-системи за рахунок скорочення часу виконання пулу завдань та простою обчислювальних ресурсів. Впровадження узагальненого критерію оцінки завдання в роботу планувальника дозволяє збільшити ефективність його роботи за рахунок зменшення простою обчислювальних ресурсів GRID-системи.

4. На підставі запропонованої математичної моделі розподілу завдань в GRID-системі, розширенні кортежів завдань і обчислювальних ресурсів, модернізації методу розподілу завдань Backfill, розробки методу пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним

простим обчислювальних ресурсів розроблено інформаційну технологію розподілу завдань, яка була впроваджена в середу моделювання GRASS. Вона об'єднує процеси передачі, зберігання, збору даних, спостережень за швидкоплинними процесами та процес їх обробки з використанням запропонованих у дисертаційній роботі методів.

5. Розроблені програмні засоби, що реалізують запропоновані методи, можуть бути використані в наукових організаціях України та світу, де виникає необхідність оперативної передачі та паралельної обробки великих масивів даних.

Використання розроблених у дисертаційній роботі методів і засобів дає можливість підвищити ефективність використання обчислювальних ресурсів GRID-системи за рахунок зменшення їх простою. Достовірність отриманих практичних результатів підтверджена експериментальними дослідженнями створеного програмного забезпечення, результатами розподілу пулів завдань для реальних GRID-систем. Напрямки теоретичних і практичних досліджень дисертаційної роботи доцільно розвивати в області створення та використання методів розподілу в гетерогенних GRID-системах.

### **СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ**

1. Волк М. А., Филимончук Т. В., Гридель Р. Н. Методы распределения ресурсов для GRID-систем. Збірник наукових праць ХУПС. Харків: ХУПС, 2009. №1(19). С. 100–104.

2. Руденко О. Г., Волк М. А., Филимончук М. А., Филимончук Т. В. Архітектура системи моніторингу трафіку в GRID. Право і безпека. Харків: ХНУВС, 2010. №1(33). С. 226–229.

3. Волк М. А., Филимончук М. А., Филимончук Т. В. Модуль распределения заданий в GRID-системах. Системи обробки інформації. Харків: ХУПС, 2012. №2(100). С. 177–182.

4. Волк М. А., Филимончук Т. В., Ал Шиблак Муаз Анализ современного состояния и развития GRID-технологий и языков описания заданий. Збірник наукових праць ХУПС. Харків: ХУПС, 2013. №2 (35). С. 75–81.

5. Волк М. А., Гридель Р. Н., Филимончук Т. В., Ал Шиблак Муаз Формализация процессов распределенной имитации информационных систем. Системи обробки інформації. Харків: ХУПС, 2013. №4(111). С. 89–93.

6. Filimonchuk T., Volk M., Ruban I., Tkachov V. Development of information technology of tasks distribution for grid-systems using the GRASS simulation environment. Eastern-European Journal of Enterprise Technologies. Information and controlling system, 2016. Vol.3/9 (81). P. 45–53. (цитуються в Scopus).

7. Волк М. А., Филимончук Т. В. Разработка модифицированного метода обратного заполнения Backfill для консервативного резервирования. Системи обробки інформації. Харків: ХУПС, 2017. №1(147). С. 33–37.

8. Волк М. А., Филимончук М. А., Корниенко Т. В. (Филимончук Т. В.) Анализ использования искусственных иммунных систем в GRID-

инфраструктуре. Системный анализ та інформаційні технології: Матеріали X Міжнародної науково-технічної конференції. К.: НТУУ «КПІ», 2008. С. 290.

9. Волк М. А., Филимончук М. А., Филимончук Т. В. Анализ алгоритмов распределения ресурсов в имитационных системах моделирования, ориентированных на GRID системы. Проблемы інформатики і моделювання. Матеріали восьмої міжнародної науково-технічної конференції. Х.: НТУ «ХПІ», 2008. С. 20.

10. Волк М. А., Филимончук М. А., Филимончук Т. В. Исследование методов распределения заданий для GRID-систем. Системный анализ и информационные технологии: Материалы XI Международной научно-технической конференции (26-30 мая 2009 г., Киев). К.: УНК «ИПСА» НТУУ «КПИ», 2009. С. 423.

11. Дьяченко К. Ю., Филимончук Т. В. Подключение модулей распределения задач к имитационной модели GRID-системы. Проблемы інформатики і моделювання. Матеріали дев'ятої міжнародної науково-технічної конференції. Х.: НТУ «ХПІ», 2009. С. 43.

12. Волк М. А., Филимончук Т. В., Гридель Р. Н. Имитационная система моделирования GRID-инфраструктуры GRASS. Системный анализ и информационные технологии: Материалы XII Международной научно-технической конференции (25-29 мая 2010 г., Киев). К.: УНК «ИПСА» НТУУ «КПИ», 2010. С. 359.

13. Фесенко М. В., Филимончук Т. В. Анализ алгоритмов распределения заявок в Grid-системах. Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління. Матеріали першої науково-технічної конференції. Х.: ДП «ХНДІ ТМ»; К.: ДП «ЦНДІ НіУ», 2010. С. 76.

14. Филимончук М. А., Филимончук Т. В. Использование средств виртуализации при мониторинге трафика для GRID. Інформаційні технології в навігації і управлінні: стан та перспективи розвитку. Матеріали першої міжнародної науково-технічної конференції. К.: ДП «ЦНДІ НіУ», 2010. С. 28.

15. Волк М. А., Филимончук Т. В. Анализ существующего прикладного программного обеспечения GRID-систем и языков описания заданий. Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: Матеріали третьої міжнародної науково-технічної конференції. Полтава: ПНТУ; Белгород: НДУ «БілДУ»; Харків: ДП «ХНДІ ТМ», Київ: НТУ; Кіровоград: КЛА НАУ, 2013. С. 49.

16. Волк М. А., Филимончук Т. В. Обобщенный критерий оценки задания для технологии планирования заданий в GRID. Информатика, математическое моделирование, экономика: Сборник научных статей по итогам третьей международной научно-практической конференции (24-26 апреля 2013 г., Смоленск). В 3-х томах. Том 2. Смоленск: Смоленский филиал Российского университета кооперации, 2013. С. 172–176.

17. Волк М. А., Филимончук Т. В. Информационная технология распределения заданий по ресурсам. Проблемы інформатизації: Матеріали

першої міжнародної науково-технічної конференції. Черкаси: ЧДТУ; Київ: ДУТ; Тольятті: ТДУ; Полтава: ПНТУ, 2013. С. 20–21.

18. Волк М. А., Филимончук Т. В. Информационная технология управления заданиями в распределенных системах. Проблемы інформатизації: Матеріали другої міжнародної науково-технічної конференції. Київ: ДУТ; Полтава: ПНТУ; Катовице: Катовицький економічний університет; Париж: Університет Париж VII Венсент-Сен-Дені; Білгород: НДУ «БДУ»; Черкаси: ЧДТУ; Харків: ХНДІТМ, 2014. С. 59–60.

19. Волк М. А., Филимончук Т. В. Информационная технология управления заданиями в GRID-системах. Информатика, математическое моделирование, экономика: Сборник научных статей по итогам четвертой международной научно-практической конференции (23-25 апреля 2014 г., Смоленск). В 2-х томах. Том 1. Смоленск: Смоленский филиал Российского университета кооперации, 2014. С. 379–383.

20. Филимончук Т. В., Ткачев В. Н. Информационная технология распределения заданий на вычислительные ресурсы в GRID-системах. Информатика, математическое моделирование, экономика: Сборник научных статей по итогам пятой международной научно-практической конференции (11-15 мая 2015 г., Смоленск). В 2-х томах. Том 1 Смоленск: Смоленский филиал Российского университета кооперации, 2015. С. 204–209.

21. Волк М. А., Филимончук Т. В. Информационная технология распределения заданий на вычислительные ресурсы для обработки радиоастрономических данных в GRID-системах. Проблемы інформатизації: Матеріали третьої міжнародної науково-технічної конференції. Черкаси: ЧДУТ; Баку: ВА ЗС АР; Бельсько-Бяла: УТіГН; Полтава: ПНТУ; 2015. С. 23.

22. Волк М. А., Филимончук Т. В. Использование системы моделирования GRASS в задачах распределения заданий в GRID-системах. 20-й Ювілейний міжнародний форум «Радіоелектроніка та молодь у ХХІ столітті». Харків: ХНУРЕ, 2016. Т.4. С. 171–172.

23. Volk M. O., Filimonchuk T. V. Information technology for job distribution in GRID-systems Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: матеріали шостої міжнародної науково-технічної конференції. Полтава: ПНТУ; Баку: ВА ЗС АР; Кіровоград: КЛА НАУ; Харків: ДП «ХНДІ ТМ», 2016. С. 45–46.

24. Ткачев В. Н., Филимончук Т. В., Митин Д. Е. Использование информационной технологии распределения заданий при обработке больших массивов данных в виртуальных частных облаках. Информационные системы и технологии: Материалы 5-й международной научно-технической конференции (2-17 сентября 2016 г., Харьков). Х.: ДРУКАРНЯ МАДРИД, 2016. С. 333–334.

25. Волк М. А., Филимончук Т. В. Модифицированный метод Backfill с консервативным резервированием. Проблемы інформатизації: Матеріали четвертої міжнародної науково-технічної конференції. Черкаси: ЧДУТ; Баку: ВА ЗС АР; Бельсько-Бяла: УТіГН; Полтава: ПНТУ; 2016. С. 20.

## АНОТАЦІЯ

Філімончук Т.В. Методи та інформаційна технологія розподілу завдань в гетерогенних GRID-системах. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Харківський національний університет радіоелектроніки, Міністерство освіти і науки України, Харків, 2017.

У дисертаційній роботі запропоновано нове вирішення актуальної науково-практичної задачі розробки методів та інформаційної технології розподілу завдань в гетерогенних GRID-системах.

Запропоновано модифіковану математичну модель розподілу завдань на обчислювальних ресурсах в GRID-системі, розширені кортежі подання завдань та обчислювальних ресурсів, модифіковано метод Backfill з консервативним резервуванням, розроблено та впроваджено в імітаційне середовище GRASS метод пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів. На підставі цих результатів розроблено інформаційну технологію розподілу завдань, яка дає можливість підвищити ефективність використання обчислювальних ресурсів GRID-системи за рахунок зменшення їх простою.

Ключові слова: розподілені системи, GRID-система, планувальник завдань, інформаційна технологія, методи розподілу, імітаційне середовище моделювання GRASS, обчислювальні ресурси, завдання.

## АННОТАЦИЯ

Филимончук Т.В. Методы и информационная технология распределения заданий в гетерогенных GRID-системах. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.06 – информационные технологии. – Харьковский национальный университет радиоэлектроники, Министерство образования и науки Украины, Харьков, 2017.

В диссертационной работе предлагается новое решение актуальной научно-практической задачи разработки методов и информационной технологии распределения заданий в гетерогенных GRID-системах.

Предложена модифицированная модель распределения заданий на вычислительные ресурсы в GRID-системе: добавление множества методов распределения позволяет осуществлять серии экспериментов для различных пулов заданий с последующим выбором плана распределения; использование коэффициента связности – осуществлять подбор вычислительных ресурсов с учетом минимизации времени на обмен данными между задачами в задании; введение пропускной способности и задержки времени передачи пакета по каналу – сократить время выполнения пула заданий.

Предложена модификация метода Backfill, которая, в отличие от

существующего, позволяет производить распределения заданий в зависимости от связности задач в задании. Метод оперирует двумя параметрами: пропускной способностью ( $\delta_n$ ) и задержкой времени передачи пакета по каналу ( $C_n$ ). В зависимости от класса задания осуществляется подбор ресурсов по одному из параметров, благодаря чему уменьшается время простоя вычислительных ресурсов за счет разгрузки каналов связи.

Разработан метод поиска распределения заданий с минимальным временем выполнения пула заданий и минимальным простоем вычислительных ресурсов, который, в отличие от существующих, использует обобщенный критерий оценки задания и имитационную среду моделирования GRASS, что позволяет повысить эффективность использования вычислительных ресурсов GRID-системы.

Предложена информационная технология распределения заданий на вычислительные ресурсы для GRID-систем, которая внедрена в имитационную среду моделирования GRASS, позволяющую воспроизводить процесс функционирования во времени элементарных событий, протекающих в реальной GRID-системе с сохранением логики их взаимодействия.

Ключевые слова: распределенные системы, GRID-система, планировщик заданий, информационная технология, методы распределения, имитационная среда моделирования GRASS, вычислительные ресурсы, задания.

## ABSTRACT

Filimonchuk T.V. Methods and information technology for task distribution in heterogeneous GRID-systems. – Manuscript.

A thesis for the candidate degree in technical sciences in the specialty 05.13.06 – information technology. – Kharkov National University of Radio Electronics, Ministry of Education and Science of Ukraine, Kharkov, 2017.

A new solution to the actual scientific and practical task of developing methods and information technology for the tasks distribution in heterogeneous GRID systems was proposed in this thesis.

Job distribution mathematical model for GRID system was proposed. Model used extended set of tasks and computing resources parameters and characteristics. On based model the modified Backfill method for job distribution was developed. The job distribution method with the minimum time for task pool executing and the minimal downtime of computing resources was developed. New information technology for job distribution on computer resources has been developed. This information technology is implemented in the GRASS simulation environment, that allows to reproduce the operation process in time for elementary logical events, occurring in a real GRID system with preserving interaction logic.

Keywords: distributed systems, GRID-system, scheduler, information technology, distribution methods, simulation modeling environment GRASS, computing resources, tasks.