

## ВІДГУК ОФІЦІЙНОГО ОПОНЕНТА

доктора технічних наук, професора **Малахова Євгенія Валерійовича**  
на дисертаційну роботу **Філімончук Тетяни Володимирівни**  
«Методи та інформаційна технологія розподілу завдань в гетерогенних GRID-системах», подану на здобуття наукового ступеня кандидата технічних наук за спеціальністю **05.13.06 – інформаційні технології**

### *Актуальність теми.*

На сьогодні є потреба забезпечення надійного, стійкого та недорогого доступу до високопродуктивних обчислювальних ресурсів, тому перед науковими організаціями ставиться задача об'єднання національних, регіональних та тематичних GRID-розробок в єдину GRID-інфраструктуру, яка використовувалася б для підтримки наукових досліджень. Однією із задач, яка знаходиться в рамках цих робіт, є задача розподілу обчислювальних ресурсів між завданнями, які надходять в систему на виконання. Для ефективного використання виділених ресурсів необхідно вирішувати задачу планування. Метою планування є розробка плану розподілу завдань, ефективність якого базується на різних критеріях. Процес планування може здійснюватися динамічно, за допомогою спеціальної програми, яка носить назву планувальник. Він керує потоком завдань, що надходять від користувачів, і проводить їх розподіл на обчислювальні ресурси Grid-системи за допомогою інформації, яку надають постачальники завдань та обчислювальних ресурсів. Результатом роботи планувальника є план розподілу, який повинен задовольнити всі вимоги постачальників завдань і оптимізувати загальний час виконання пулу завдань, вартість та ефективність використовуваних обчислювальних ресурсів GRID-системи.

Нині існує низка систем управління і розподілу ресурсів, проте планувальники, які реалізовані в них, не дають ефективного способу розподілу завдань, який дозволив би мінімізувати час простою обчислювальних ресурсів, скоротити обсяги і час переданої інформації між ресурсами. Тому розробка методів та технології 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.

***Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації, їх достовірності і новизни.***

Викладені в дисертаційній роботі положення, отримані автором теоретичні та практичні результати мають належний ступень обґрунтованості, який було





досягнуто за рахунок обґрунтованого використання математичних засобів процесної алгебри, обчислення виразів тощо. Достовірність результатів підтверджується тим, що отримані автором результати або узагальнюють результати, незалежно отримані раніше, або такі результати впливають з результатів автора.

Новизна наукових результатів, які отримані у роботі, полягає у наступному:

- підвищення ефективності використання обчислювальних ресурсів GRID-системи за рахунок скорочення часу виконання пулу завдань та простою обчислювальних ресурсів, що досягається за допомогою метод пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів, який, на відміну від існуючих, використовує узагальнений критерій оцінки завдання та імітаційне середовище моделювання GRASS;

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

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

### ***Важливість результатів дисертаційної роботи для науки і практики.***

Практична значимість отриманих теоретичних результатів дисертаційної роботи підтверджена поліпшенням якості функціонування системи GRASS, а також актами впровадження результатів. Зокрема, практичне вирішення теоретичних досліджень полягає у наступному: модифіковано математичну модель розподілу завдань у GRID-системі за рахунок введення множини методів розподілу та додаткових параметрів у поданні обчислювальних ресурсів і завдань; модифіковано метод розподілу завдань Backfill, який враховує інтенсивність і обсяг потоків даних між задачами в завданні; запропоновано метод пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простоем обчислювальних ресурсів; розроблено інформаційну технологію розподілу завдань, яка була впроваджена в середу моделювання GRASS.

Також у роботі запропоновано блок, який відповідає за формування додаткових параметрів для ефективного розподілу завдань. Цей блок реалізовано за



допомогою двох модулів, які дозволяють більш продуктивно управляти процесом розподілу завдань на обчислювальні ресурси та здійснювати підбір обчислювальних ресурсів з урахуванням скорочення часу, який витрачається на обмін між задачами у завданні. У імітаційне середовище GRASS додано базу даних, яка фіксує інформацію про кожне запущене завдання, що дозволяє планувальнику отримувати результати попередніх розподілів завдань за різними методами розподілу. Маючи цю інформацію, планувальник обирає такий план розподілу, який матиме вираш за рядом критеріїв, заздалегідь встановлених користувачем.

### ***Повнота викладу наукових положень, висновків і рекомендацій в опублікованих працях.***

За результатами дисертаційних досліджень опубліковано 25 наукових праць. З них 7 статей у виданнях, які зазначені в переліку фахових видань України з технічних наук та 18 тез доповідей конференцій. Всі наукові праці входять до міжнародних науково-метричних баз, одна наукова праця до Scopus.

Основні положення дисертаційної роботи повністю розкрито в публікаціях і відповідають змісту автореферату.

### ***Оцінка змісту дисертації, її завершеності й оформлення.***

Побудова дисертації відповідає прийнятим для наукового дослідження нормам. Дисертація складається з вступу, чотирьох розділів, висновків, списку використаних джерел зі 150 найменувань на 17 сторінках та 2 додатків на 8 сторінках. Загальний обсяг роботи складає 180 сторінок, включаючи 131 сторінку основного тексту.

Для вирішення питань, зазначених у меті роботи, дисертант провів аналіз предметної області, розглянув питання розвитку розподілених систем як середовища для розв'язання задач великої обчислювальної складності. В роботі наведено огляд технологій розподілених обчислень та Cloud-технологій, існуючих планувальників завдань, мов опису завдань в GRID-системах (розділ 1). Аналізуючи сучасний стан предметної області, автор дійшов висновку щодо необхідності розробки нових моделей та методів розподілу завдань на обчислювальні ресурси в гетерогенних GRID-системах.

У другому розділі здобувачем була досліджена існуюча модель розподілу завдань у GRID-системі. На основі проведеного аналізу було виявлено, що під час розподілу завдань на обчислювальні ресурси не використовується ряд параметрів, які дозволять підвищити ефективність використання GRID-системи за рахунок зменшення часу виконання пулу завдань: зв'язність задач в завданні, пропускну здатність кагалу та затримка часу передачі пакету по каналу.



У третьому розділі запропонована модифікована модель розподілу завдань на обчислювальні ресурси за рахунок введення в неї множини методів розподілу, що дозволяє проведення серії експериментів для подальшого вибору плану розподілу з мінімальним часом виконання пулу завдань та мінімальним простим обчислювальних ресурсів. 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.

Розподіл завдань на обчислювальні системи здійснюється за рядом параметрів, які характеризують завдання. Однак це викликає труднощі у процесі розподілу. В роботі пропонується введення узагальненого критерію для оцінки завдання, який дозволяє збільшити ефективність роботи планувальника за рахунок зменшення простою обчислювальних ресурсів GRID-системи.

Метод Backfill жодним чином не враховує час, який витрачається на передачу вхідних та вихідних даних завдання, тому в роботі запропонована модифікація методу Backfill, яка враховує трафік обміну даними між задачами в завданні, що дозволяє скоротити час на пересилання результатів між окремо взятими ресурсами і зменшити трафік у мережі за рахунок розвантаження каналів зв'язку.

На сьогодні актуальною задачею у GRID-системі є задача ефективного розподілу завдань по обчислювальних ресурсах. Здобувач запропонував метод пошуку розподілу завдань з мінімальним часом виконання пулу завдань та мінімальним простим обчислювальних ресурсів. Даний метод було впроваджено в імітаційне середовище моделювання GRASS, що дозволяє підвищити ефективність використання обчислювальних ресурсів GRID-системи за рахунок скорочення часу виконання пулу завдань та простою обчислювальних ресурсів.

На основі модифікованої запропонованій математичної моделі розподілу завдань в GRID-системі, розширенні кортежів подання завдань та обчислювальних ресурсів, модернізації методу Backfill, розробці методу пошуку найкращого розподілу для пулу завдань була запропонована інформаційна технологія розподілу завдань, яка була впроваджена в імітаційне середовище моделювання GRASS.

У четвертому розділі наведена практична реалізація імітаційної технології розподілу завдань на обчислювальні ресурси GRID-системи. Запропонована технологія об'єднує процеси передачі, зберігання, збору даних, спостережень за швидкоплинними процесами та процес їх обробки з використанням запропонованих у дисертаційній роботі методів.

Додатки до дисертації містять акти про впровадження результатів дисертаційної роботи та перелік використаних джерел здобувача.

Обсяг та структура дисертаційної роботи відповідають вимогам щодо тексту дисертацій на здобуття наукового ступеня кандидата технічних наук. Автор використав сучасну бібліографію та наукову термінологію. Зміст та результати



досліджень викладено аргументовано, результати практичної реалізації підтверджують теоретичні висновки. Автореферат відповідає змісту дисертаційної роботи.

### ***Зауваження щодо змісту дисертації.***

Зміст дисертації послідовний та логічний. Мета роботи досягнута за рахунок чітко визначених і розв'язаних задач побудови моделі предметної області, визначення властивостей технологій розподілу завдань на обчислювальні ресурси, розробки методу пошуку розкладу, який орієнтований на мінімізацію часу виконання вхідного пулу завдань. Все це забезпечує можливість ефективного управління розподілом завдань за наявними обчислювальними ресурсами GRID-системи.

В дисертаційній роботі виявлено такі недоліки:

1) В процесі аналізу предметної області, не розглянуті рішення, які надає теорія масового обслуговування, наприклад рішення на базі марківських ланцюгів.

2) По ходу роботи не наведено переваги впровадження запропонованої технології розподілу завдань на обчислювальні ресурси в імітаційне середовище моделювання GRASS.

3) В роботі немає чіткої інструкції при формуванні якісних параметрів, що викликає ймовірність відмов у виконанні завдання, в тому випадку, коли це виконанні можливо, наприклад при розбіжності версії ОС.

4) Вимагає пояснення структура БД, що розроблюється для збереження результатів процесу розподілу в системі GRASS (рис. 3.8): на мій погляд зв'язки між таблицями Політики розподілу, План розподілу і Експеримент допускають суперечливість даних.

5) По ходу роботи не зрозуміло, що дає розширення моделі розподілу GRID-системи після введення додаткових параметрів.

6) У п. 4.10 у висновках до аналізу результатів застосування технології розподілу завдань відзначено, що тому що «не існує алгоритму розподілу, який на будь-якому пулі завдань одержував би найкращий план», то для вибору найкращого методу розподілу «можна провести моделювання ...». Однак, якщо класи завдань, належність до яких залежить від обсягу вхідних і вихідних даних (п. 4.9), розглядати як аналогію типізації завдань за часом виконання в системах обробки даних, то запроповану технологію можна було б випередити рекомендаціями з теорії масового обслуговування, наприклад: завданням з меншим часом виконання давати більший пріоритет і т.ін.

Однак наведені вище зауваження суттєво не впливають на загальну позитивну характеристику дисертаційної роботи, яка має наведені актуальність, наукову та практичну цінність.



**Висновок щодо відповідності дисертації встановленим вимогам.**

Дисертаційна робота виконана у відповідності до наступних пунктів, зазначених в паспорті спеціальності 05.13.06 – інформаційні технології:

2) розроблення інформаційних технологій для аналізу та синтезу структурних, інформаційних і функціональних моделей об'єктів і процесів, що автоматизуються;

8) побудова інформаційних технологій для ефективного розроблення програмного забезпечення комп'ютерних мереж і систем розподіленої обробки даних;

10) моделювання предметних галузей інформаційних систем (аналітичне, імітаційне, інфологічне, об'єктно-орієнтоване, тощо) на підґрунті створення і застосування відповідних інформаційних технологій.

**Загальні висновки.**

В цілому дисертаційна робота є завершеною науковою працею, в якій отримано нові обґрунтовані теоретичні і практичні результати, що є суттєвими для вирішення питань ефективності функціонування розробленої інформаційної технології розподілу завдань на обчислювальні ресурси гетерогенної GRID-системи.

Вважаю, що дисертаційна робота відповідає вимогам пунктів 9, 11, 12-14 «Порядку присудження наукових ступенів», затвердженого постановою Кабінету Міністрів України від 24.07.2013 № 567 (зі змінами, внесеними згідно з постановою КМУ № 567 від 27.07.2016) щодо кандидатських дисертацій, а її автор Філімончук Тетяна Володимирівна заслуговує на присудження наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Офіційний опонент,

доктор технічних наук, професор,  
завідувач кафедри математичного  
забезпечення комп'ютерних систем  
Одеського національного університету  
імені І.І. Мечникова, м. Одеса



Є.В. Малахов

Підпис Малахова Є.В. засвідчую  
Вчений секретар ОНУ імені І.І. Мечникова  
канд. хім. наук, доцент




Є.В. Курандо