НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
"КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"

СІМОНЕНКО ВАЛЕРІЙ ПАВЛОВИЧ

УДК 681.324

МЕТОДИ ТА ЗАСОБИ ЕФЕКТИВНОГО РОЗВ'ЯЗАННЯ ЗАДАЧ СТАТИЧНОГО І ДИНАМІЧНОГО ПЛАНУВАННЯ В РОЗПОДІЛЕНИХ ОБЧИСЛЮВАЛЬНИХ СИСТЕМАХ

Спеціальність 05.13. ~~13~~ 08 — Обчислювальні машини, системи і мережі.

Автореферат дисертації на здобуття наукового ступеня
доктора технічних наук

Київ - 1997

ЛННБ України ім.В.Стефаника
00751330 (J)

Дисертація є рукопис.

Робота виконана в Національному технічному університеті України (НТУУ)
"Київський політехнічний інститут"
На кафедрі обчислювальної техніки

Наукові консультанти — член-кореспондент НАН України,
доктор технічних наук, професор Самофалов К. Г.,
професор кафедри обчислювальної техніки НТУУ
доктор технічних наук, професор Луцький Г. М.
зав. кафедрою обчислювальної техніки НТУУ

Офіційні опоненти:

Член-кореспондент НАН України доктор технічних наук, професор Палагін О. В., зам. директора інституту Кібернетики імені В. М. Глушкова.

Доктор технічних наук, професор Нагорний Л. Я., Київський міжнародний Університет цивільної авіації, зав. кафедрою обчислювальних машин, систем та мереж.

Доктор технічних наук, професор Павлов О. А., Національний технічний Університет України "Київський політехнічний інститут" декан факультету інформатики та обчислювальної техніки.

Провідна установа:
Інститут проблем моделювання в енергетиці НАН України
Відділи - теорії моделювання, математичного і економометричного моделювання.

Захист відбудеться 8. 12 1997 р. В 14 30 годин на засіданні спеціалізованої ради Д26.002.02 в Національному технічному університеті України "Київський політехнічний інститут"
(м. Київ, пр. Перемоги, 37, корп. 18, ауд 306)
Відгуки на автореферат в двох примірниках, завірені печаткою установи, просимо направити за адресою: 252056, м. Київ, пр. Перемоги, 37,
вченому секретарю НТУУ "КПІ".
З дисертацією можна ознайомитись у бібліотеці
Національного Технічного університету України "КПІ"

Автореферат розісланий 30.10 1997 р.

Вчений секретар
спеціалізованої вченої ради М. М. Орлова

АНОТАЦІЯ

Метою дисертаційної роботи є підвищення ефективності використання ресурсів та скорочення накладних витрат на організування обчислювань у гомогенних та гетерогенних розподілених системах обробляння інформації за рахунок розробляння високоефективних методів та засобів проектування адаптивних статичних та динамічних планувальників, що забезпечують в умовах просторово-часових обмежень організовання обчислювального процесу, адекватну стану обчислювального середовища та вхідному потоку робіт.

Для досягнення поставленої мети у дисертації вирішуються наступні задачі:

1. Аналіз архітектурних особливостей паралельних обчислювальних систем (ПОС) та систем розподіленого обробляння даних (СРОД), виділення та класифікування узагальнених архітектурних характеристик, що визначають принципи управління та організування обчислювальних процесів в ПОС та СРОД.
2. Класифікування та аналіз стратегій, принципів, методів та алгоритмів планування робіт в СРОД, розробляння загальної методики опису, аналізу, визначення вимог та вибору системи планування робіт у СРОД.
3. Теоретичне обгрунтування та розробляння на структурно-логічному рівні загального підходу до організування обчислювальних процесів у розподіленом обчислювальном середовищі на всіх етапах проходження робіт, зокрема в умовах реорганізування обчислювальних процесів та перерозподіляння ресурсів в умовах підвищених вимог до життєздатності СРОД.
4. Розробляння загальної схеми системи планування робіт у СРОД, що дозволить виконати математичне формулювання задачі планування та розподілення робіт, визначити вимоги та критерії оптимізування системи планування за динамічного керування та диспетчіювання робіт у СРОД.
5. Теоретичне обгрунтування та розробляння методу спрямованого пошуку рішення задач статичного та динамічного планування робіт у СРОД.
6. Теоретичне обгрунтування та розробляння методу структурно-семантичного аналізу вхідної інформації з метою підвищення ефективності планування робіт у СРОД.
7. Розробляння алгоритмів статичного та динамічного планування та

ІНВ ім. В. Стефаника
АН України

розподілу робіт у СРОД зниженої часовій складності.

8. Обгрунтування загальної методики дослідження моделей та розробляння алгоритмів аналізу для оцінювання ефективності процедур планування та розподілу робіт, що пропонуються, а також для порівняння розроблених алгоритмів з відомими.

На захист виноситься наступне.

1. Загальна класифікація паралельних систем обробляння даних, що враховує особливості архітектури сучасних СРОД та організації обчислювальних процесів у розподілених середовищах.
2. Загальна модель планування та організування обчислень в розподілених системах обробляння даних.
3. Математичне формулювання задач планування, розподілу та оптимізування процесів за статичного та динамічного планування робіт у СРОД.
4. Теоретичне обгрунтування причин неефективності (великої часової складності) відомих алгоритмів рішення задачі складання розкладу у ПОС.
5. Базові теореми, твердження та слідства, призначені в основу розробленого методу спрямованого пошуку рішення задач планування та розподіляння робіт у ПОС.
6. Аналитико-графічний метод локалізування зони пошуку варіанту просторово-часового розміщення робіт у ПОС із заданою функцією мети.
7. Адаптивний алгоритм спрямованого пошуку рішення задачі розподілу робіт та оптимізування просторово-часових характеристик розкладу за динамічного планування у неоднорідних, півнеоднорідних та суворо-неоднорідних ПОС та ПОС реального часу.
8. Сканувальний алгоритм просторово-часового розподіляння робіт у СРОД з урахуванням обраних обмежень до системи планування робіт.
9. Метод структурно-семантичного аналізу вхідної інформації та вироблення стратегії балансового оптимізування просторово-часового відображення зв'язаного графа задач у ПОС, на основі розроблених процедур вертикально-горизонтального кластеризування та виділяння вершин, що мають признаки "явної", "прихованої" та "мультиплікативної" транзитності.
10. Результати статистичного дослідження розроблених методів динамічного планування процесів у СРОД та порівняння їх із відомими.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми.

Підвищування продуктивності обчислювальних систем (ОС) є постійним предметом досліджень у течії всього періоду розвитку засобів ОТ. Як відомо, підвищення продуктивності можна досягнути як за рахунок вдосконалення технічних засобів, так і за рахунок підвищування ефективності організування, планування та диспетчіювання робіт у ОС. Іншими словами рішення визначеної проблеми можна шукати наступними шляхами: реалізацією істинного або удаваного суміщення робіт на однорідних або неоднорідних ресурсах однопроцесорної обчислювальної системи; організацією паралельного виконання окремих частин програм; ефективного використування ресурсів у ПОС під час багатопрограмного режиму її роботи; застосуванням ефективних процедур планування робіт і диспетчіювання процесів у розподілених операційних системах за динамічної та статичної організації обчислювальних процесів.

Складність задач, що вирішуються у рамках цієї проблеми, зумовлена тим, що більшість із них носить комбінаційний характер та належить до класу NP-повних.

Поряд із підвищенням продуктивності процесорів, інтенсивно ведуться роботи із вдосконалювання паралельних систем обробляння даних, для яких питання організації планування та обчислень ще більш ускладнюються. Слід відзначити, що технічна реалізація високопаралельних ОС випереджає розробляння системного програмного забезпечення, що вирішує задачі оптимального занурювання паралельних програм у обчислювальне середовище.

Отримали розвиток системи розподіленого обробляння даних (СРОД) у обчислювальних мережах під керуванням розподіленої операційної системи (РОС). Однак, наявне системне програмне забезпечення для вирішення задач планування та диспетчіювання заздалегіть знижує ефективність ОС в цілому за рахунок додаткових витрат машинного часу на вірішування системних задач управління, планування та диспетчіювання. Залишають також бажати кращого характеристики безпеки та життєздатності ОС.

Відомо безліч робіт у області теорії распаралелювання обчислень та реалізації практично працюючих систем распаралелювання алгоритмів та програм, що передбачають можливості виконання попереднього підготування робіт у багатопрограмном режимі роботи системи [1,2]. Однак дослідження у області теорії побудови систем планування робіт у ПОС ще не дали достатньо ефективних практичних алгоритмів та програм для

СРОД, тобто таких, що дозволять з незначними накладними витратами застосовувати їх для створювання РОС. Відсутня класифікація стратегій, принципів, методів та алгоритмів з єдиних методологічних позицій, що враховують особливості управління паралельними процесами. Практично відстуне програмне інтерфейсне середовище роботи користувача з багатопроцесорними (кластерними) робітничими станціями. Відсутні системи автоматичної підготовки паралельних програм для ПОС. Аналіз розвитку ОС у останні роки показує, що створення та застосування стандартів для високошвидкісного зв'язку, таких як FDDI, HiPP, SONET, дозволяє будувати глобальні ОС, що об'єднують різноманітні потужні усунені обчислювальні машини для спільної, ефективної та ощадливої обробки різноманітних задач. Однак, широке застосування таких систем як OMOEBA, PVM, MPI, Express, P4, Linda, Panda, Pirhana стримується достатньо великої часткою висококваліфікованої ручної праці під час підготування програм для ПОС, а засоби що застосовуються у багатозадачних операційних системах OS/2, WINDOWS 3.11, WINDOWS-95, WINDOWS NT, UNIX та ін. для організації багатопрограмних режимів роботи в ПОС не забезпечують ефективного використування ресурсів розподіленої ОС. Окрім цього, практично відстуні ефективні засоби, які дозволяють програмісту перевірити або змоделювати занурення розпаралеленої програми в обчислювальне середовище високопаралельних систем.

Виконаний у цієї роботі аналіз архітектурних особливостей ПОС та вимог до системного програмного оточення, а також детальний аналіз принципів, стратегій та алгоритмів планування і диспетчеріювання, що використовуються, дозволяє сформулювати дві основні задачі, які значною мірою визначають якість функціонування ПОС загального користування: просторово-часове складання розкладу виконання робіт та їх адресного розподілення у конкретному обчислювальному середовищі.

Статичне та динамічне планування та розподіляння ресурсів та задач є відображенням у *розкладі* виконання множин завдань трьох макрочинників: *системи ресурсів* (архітектура, структура та характеристики), *потоку завдань* (вимоги та структура) та *часу* (в системах реального часу або за обмеження ресурсів системи). Окрім цього, кожний з вище зазначених макрочинників рекурсивно являє собою сукупність мініта мікрочинників під час деталізації розкладу у просторі ресурсів ОС та у часі виконування цього розкладу на кожному ресурсі окремо. Рішення власне задачі планування та розподіляння робіт з урахуванням усіх вимог, що ставляться до системи планування та параметрів ОС та/або вхідного

потоку завдань є критично важке. Вона являє собою сукупність декількох задач, що важкорозв'язуються та належать до класу *NP-повних.*

Таким чином, досліджування та розробляння моделей, стратегій методів, та засобів організації обчислювальних процесів у розподілених системах обробляння даних та їх конструкторсько-технологічне та програмне проектування, моделювання, практичне реалізування . та апробування є актуальними, мають теоретичний та практичний інтерес.

Об'єктами дослідження є — математичні моделі алгоритмів статичного та динамічного планування, що орієнтувалися на застосування адаптивних засобів спрямованого пошуку рішення у розподілених однорідних та неоднорідних ОС.

Методи досліджень грунтуються на використуванні: теорії систем, теорії графів, теорії множин, теорії моделювання, методів комбінаторного оптимізування, а також теорем, затвердженнь та слідств, що доведені у дисертації. Основні положення та теоретичні оцінки підтверджено результатами імітаційного моделювання на ЕОМ.

Наукова новизна. Запропоновано новий підхід до рішення задач статичного та динамічного планування та диспетчіювання робіт у розподіленій обчислювальній системі обробляння даних, суть якого полягає у попередньому структурному аналізі вхідного стану системи та визначенні ефективної, адаптивної стратегії планування, яка гарантує отримання оптимізованого рішення за істотного зменшення часу планування. Під час цього отримані такі результати:

– Розроблена загальна модель планування для систем розподіленого обробляння даних, яка дозволила виділити етапи проходження робіт через ПОС, виконати математичне порушення задач, що вирішуються на кожному етапі для вибраного класу ОС і моделі планування.

– Сформульовані теореми, твердження та слідства, що доводять коректність та високу ефективність розроблених методів та засобів.

– Розроблені алгоритми спрямованого пошуку розкладу виконання робіт у СРОД різноманітного класу, які базуються на стратегії покрокового конструювання розкладу для задач планування та оптимізування розподілу робіт у ПОС. Переваги алгоритмів, що пропонуються підтверджені аналітичними розрахунками часової складності усіх процедур та статистичними експериментами на спеціально розробленій імітаційної системі.

Практична цінність результатів дисертаційної роботи полягає у тому,

що застосування запропонованих методів та засобів дозволяє підвищити ефективність планування в СРОД типу PVM, MPI, SmartNet.

Крім того, нові алгоритми та програми адаптивних планувальників та диспетчерів які слугують для вирішення широкого класу задач динамічного та статичного планування робіт у ПОС з урахуванням як особливостей паралельних алгоритмів рішення практичних задач, так і особливостей топології та характеристик обчислювального середовища або мережі. Нові алгоритми та програми відрізняються від відомих істотно меншою часовою складністю та дозволяють одержувати оптимізований розклад у реальному часі.

У дисертаційну роботу входять результати, отримані автором з 1971 по 1997 рр. на кафедрі обчислювальної техніки Київського політехнічного інституту: під час виконання госпдоговорних (тема №647 "Розробка багатомашиного обчислювального керуючого комплексу"), та держбюджетних НДР та ОКР, договорів про науково-технічне співробітництво з політехнічним інститутом м.Ченстохова (Республіка Польща) та університетом м. Валенсії (Іспанія), під час роботи у якості радника Головного управління планування та статистики Міністерства Освіти Республіки Куба, радника факультету Інформатики Університету Лас-Віл'яс (Куба), а також під час підготовки та науковому керівництві роботами аспірантів та стажистів.

Методологічні та наукові аспекти дисертаційних досліджень знайшли свою реалізацію у навчально-методичному забезпеченні лекційних дисциплін "Організація обчислювальних процесів в ЕОМ, комплексах, мережах та системах" та "Операційні системи", які викладаються автором студентам спеціальності 22.01 факультету інформатики та обчислювальної техніки Національного технічного університету України "КПІ".

Достовірність теорем, тверджень та висновків обгрунтовується точними теоретичними доказами та результатами статистичного моделювання.

Апробація роботи. Основні наукові результати дисертаційної роботи були надані та обговорювалися на державних, міжнародних та міждержавних, а також союзних та республіканських наукових симпозіумах, конференціях і семінарах, в тому числі: школі-семінарі "Розподілена обробка інформації (м. Новосибірск, 1982), науковій конференції з вищої освіти (Куба, м. Гавана, 1983), національній конференції з інформатики (Куба, м. Санта Клара, 1984), другій всесоюзній нараді з конвейєрних обчислювальних систем (м. Київ, 1988), VII всесоюзній школі-семінарі

"Распаралелювання обробки інформації", (м. Львів, 1989), First international conference on parallel processing and applied mathematics (Польща, 1994), міжнародної конеренції "Проблеми алгебри та кібернетики" (м. Гомель 1995), другої української конференції з автоматичного управління" (м. Львів, 1995), 10 European Simulation Multiconference (ESM-96, м. Будапешт, Угорщина, 6 червня 1996), 3 International Conference on Parallel and Distributed Processing Techniques and Applications — PDPTA'96 (м. Каліфорнія, США, 9-11 серпня 1996), міжнародному форумі інформатизації МФІ-96 (м. Москва, 1996), 11 International Parallel Processing Symposium (м. Женева, Швейцарія, 1-5 квітня 1997), XXIV міжнародній конференції та дискусійному науковому клубі "Нові інформаційні технології в науці, освіті та бізнесі", New Information Technology in Science, Education and Business, IT+SE`97 (Україна, Крим, Ялта-Гурзуф, 15-24 травня 1997), II міжрегіональному науково-технічному семінарі "Застосування пластиковых карт та захист інформації" (м.Сергієв-Посад, Росія, 10-11 червня 1997), The 2 International Conference on Parallel Processing & Applied Mathematics, PPAM' 97 (м. Закопане, Польща, 2-5 вересня 1997).

Публікації. За матеріалами дисертації опубліковано 76 наукових роботи. Основні результати викладені у 40 роботах, у тому числі в 3-х монографіях (3; 9.5 та 19 д.л.), та 6 навчальних посібниках. Основний зміст дисертації втілений у роботах [1,2,5,9,10,23,30.34,35,40] що опублікувалися без співавторів. Окремі питання, стосовні застосування отриманих наукових результатів, опубліковані у співавторстві з співробітниками у роботах [6,7,11-16,24-33,36-39], а практичні аспекти застосування результатів розглянуті в роботах [4,11,17-23].

Структура та обсяг роботи. Дисертаційна робота складається з вступу, п'яті розділів, додатків, що викладені на 296 сторінками друкованого тексту, містить 170 малюнків та таблиць на 43 сторінках, список літератури з 245 найменувань та п'ять додатків.

У вступі обгрунтовується актуальність теми дисертаційної роботи, формулюється мета та задачі досліджування, основні положення, що виносяться на захист.

В першому розділі розглянуті питання організації обчислень у паралельних обчислювальних системах, досліджено вплив архітектури обчислювальних систем паралельного обробляння інформації (ПОС) на організацію та управління обчисленнями, принципи побудови розподілених операційних систем, засобів, структур та алгоритмів планування та диспетчеріювання, визначені вимоги та критерії оптимізації

рішення задач планування та розподілу. Запропонована узагальнена класифікувальна схема принципів, алгоритмів та стратегій рішення задач планування та розподілу. Виконано дослідження питань організації обчислень в ПОС, визначена загальна модель системи планування, особливості рішення задач планування на кожному рівні.

У другому розділі викладена загальна математична модель рішення задач планування та розподілу, математичне порушення задачі динамічного планування для вибраної моделі ОС та системи планування. Виконане теоретичне обгрунтування засобу спрямованого пошуку рішення, наведені теореми, твердження та слідства, призначені у основу процедур адаптивного алгоритму мультианаліза (АМА) та алгоритму спрямованого пошуку (АСП) рішення задачі динамічного розподілу робіт у неоднорідній, півнеоднорідній та строгонеоднорідній розподіленої ОС.

У третьому розділі виконано огляд стратегій рішення задач планування у неоднорідній СРОД та розглянуто метод та алгоритм АСП розподілу робіт у неоднорідній обчислювальній системі з використовуванням алгоритмів АМА, Угорського методу, який було модифіковано. Наведено можливість застосування АСП для динамічного планування робіт в НСРОД реального часу. Виконано розрахунок часової складності запропонованого АСП.

У четвертому розділі виконано обгрунтування засобу структурно-семантичного аналізу вхідної інформації, наведені твердження, призначені в основу процедур локалізування зони пошуку та алгоритму сканівного просторово-часового розподілу взаємопов'язаних робіт між ресурсами під час вирішування задач статичного планування для систем із загальною пам'яттю та розподілених систем. Розглянутий евристичний алгоритм оптимізує пакування робіт під час вирішування задачі горизонтального кластеризування у алгоритмі сканівного просторово-часового розподілу взаємопов'язаних робіт між ресурсами. Розглянутий підхід до розробляння алгоритмів предпроцесорного розбивання вхідної множини взаємопов'язаних задач на підмножини підзадач, відповідних до структури кластерних робітничих станцій.

У п'ятому розділі виконано обгрунтування вибраної системи моделювання, критеріїв порівняння, наведені результати статистичного порівняльного аналізу розроблених алгоритмів та програм з найбільш відомими, що підтверджено результатами теоретичних досліджень.

У висновках наведені основні теоретичні та практичні результати роботи, рекомендації щодо їхнього використовування.

ОСНОВНИЙ ЗМІСТ РОБОТИ

Система паралельного обробляння даних (СПОД) — це сукупність технічних засобів (ТЗ) та програмного забезпечення (ПЗ), яка призначена для інформаційного паралельного обслуговування користувачів та технічних об'єктів. Існують різноманітні підходи до класифікування СПОД або паралельних обчислювальних систем (ПОС). Основними ознаками для класифікування є: ступінь однорідності обчислювальних машин та процесорів, які входять у обчислювальну систему, механізм управління обчислювальним процесом у системі та взаємозв'язком між обчислювальними вузлами, організація пам'яті та її доступність, структура зв'язків в ОС та можливість її спрямованої реконфігурації. Серед СПОД особливий клас складають неоднорідні системи. Типовим прикладом неоднорідних систем є система розподіленого обробляння даних (СРОД), побудована на різнотипових засобах, які обробляють. Під час рішення задач планування та диспетчеріювання має значення засіб організації пам'яті ПОС: ПОВС з розподіленою або з пам'яттю, що поділяється (загальною). У загальному випадку архітектуру СРОД можна уявити як PARA-PC - комп'ютер з різноманітними засобами доступу до пам'яті за взаємодії паралельно виконуваних процесів (UMA, NUMA).

У ПОС (багатопроцесорних, багатомашиних комплексах, мережах) проблема планування для розподілу заявок (завдань, процесів...) на ресурсах системи які її задовольняють є однією із задач, що важкорозв'язуються у обчислювальній теорії та теорії розкладів. У цілому, обчислювальну систему та обчислювальний вузол ПОС у взагалом можна розглядати з різноманітним ступенем деталізації як сукупність ресурсів, а об'єкти планування обчислень або "завдань" на ресурси можуть бути уявлені як програми, процедури, паралельні дільниці програм, окремі блоки команд, команди та окремі макро- або мікрооперації.

Розширення поняття "ресурсу" в системах паралельного обробляння інформації та функціональних можливостей обчислювального вузла до можливостей обчислювальної системи викликало необхідність зміни підходів до організування та планування обчислювального процесу та можливості формування єдиної методології планування незалежно від ступіня деталізації ОС. При розширюванні поняття ресурсу, обчислювальні системи, традиційно що вважалися однорідними за структурою (архітектурою), стають неоднорідними з точки зору системи планування. Для розгляду питань планування у таких системах введене поняття напівнеоднорідної та суворонеоднорідної ОС [1,16]. У роботі вказано, що для чисельної оцінки ступеня гетерогенності СРОД можна використати

відносну вагу потужності самої швидкої або самої повільної машини (Wi=1) під час виконання задачі P на всіх інших.

$$H_f = \frac{\sum_{j=0}^{n}(1 - W_j^f(P))}{n} \qquad H_s = \frac{\sum_{j=0}^{n}(1 - W_j^s(P))}{n}$$

Окрім цього, використовування таких характеристик як прискорення, ступінь розпаралелювання для НСРОД дозволяють визначити якісну характеристику використання ОС під час укладання розкладу розпаралелених робіт між ресурсами. Цією характеристикою є ефективність неоднорідної системи (EF). Вона визначається з виразу —

$$EF_p = \frac{E_p}{\sum_{i=0}^{n} W_i} \text{ або } EF_p = \frac{E_p}{n \times H_{гом}}$$ Ер-продуктивність НСРОД.

Де: $$E_p = \frac{S_p}{n}$$ або $$E_p = \frac{\sum_{i=0}^{n} \frac{W_i \times size(P_i)}{S_i^P}}{\sum_{i=0}^{n}(t_H - t_i^{own}) \times W_i}$$

size (Pi) — це розмір програми, а t_i^{own} - це час, протягом якого машина Mi безпосередньо зайнята доцільною роботою;

Sp — прискорення для НСРОД

$$S_p = \frac{\min\{t_P^{M_1}, t_P^{M_2}, \ldots, t_P^{M_n}\}}{t_P^H}$$

Якщо код P можна поділити на паралельно виконувані задачі $P_1, P_2, .., P_n$ на машинах M_i, той t_p^H можна визначити як найбільш великий час, необхідний для виконання окремої задачі.

$$t_P^H = \max\{t_{P_1}^{M_1}, t_{P_2}^{M_2}, \ldots, t_{Pn}^{M_n}\},$$

Де кожна задача виконується на найбільш вдалій для неї конфігурації.

Тоді Sp визначається як:

$$S_p = \frac{\min\{t_P^{M_1}, t_P^{M_2}, \ldots, t_P^{M_n}\}}{\max\{t_{P_1}^{M_1}, t_{P_2}^{M_2}, \ldots, t_{Pn}^{M_n}\}}.$$

Крім того Sp можна визначити з виразу

$Sp = H_{гом} \times P_{пар}$

Де $H_{гом}=W_{ср}$, а $$P_{п\,ар} = \frac{\sum_{i=0}^{n} t_i^a}{t_P^H}$$

де t_i^a - це активний час обчислювання у кожному обчислювальному вузлі, а t_p^H - це загальний час виконання задачі у НСОД.

Класична трьохрівнева схема планування обчислювального процесу прийнятна у ОС з невеликим числом процесорів складається з 3 типів планувальників: високого рівня, проміжного та нижнього рівнів. Інтенсивне впровадження розподілених систем та систем масового розпаралелювання вимагає розробляння нової схеми, зміни та доповнення системи організації та планування обчислювального процесу. Повну систему планування у цьому випадку можна уявити у вигляді семирівневої схеми, [1,2,3,5,6] що містить такі етапи: 1 — введення, 2 — аналізування, 3 — розпаралелювання, 4 — адаптаціювання, 5 — розподіляння, 6 — оптимізування, 7 — перерозподіляння. (Мал.1)

Розподіл завдань (робіт, процесів) на ресурси у системах масового розпаралелювання (СМР — однорідні або у розподілених системах (РС — неоднорідні) має наступне загальне твердження [1,2]:

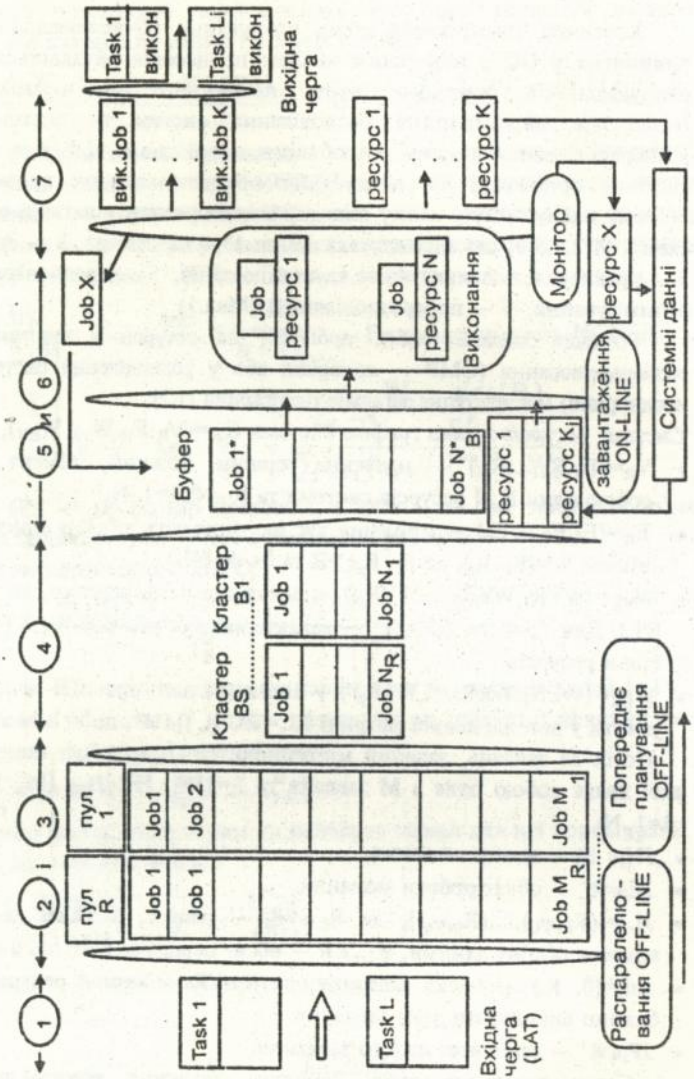
Система ресурсів задана графом системи $G_R=(V_R,E_R,W_{VR},W_{ER})$, де:

- $V_R=\{R_1,R_2,...,R_N\}$ — множина вершин, кожний елемент якої являє собою один із N ресурсів системи та $R_i\in N|$, i=1..N.
- $E_R=\{E_1,E_2,...,E_d\}$ — множина дуг що описують зв'язки одного ресурсу з іншим $E_i=\{R_i, R_j\}$, де $R_i, R_j\in VR$ та $0\le d\le N^2$.
- $W_{VR}=\{WVR_1,WVR_2,...,WVR_N\}$ — множина вагів вершин, де $WVR_i=\{RE_i, RT_i\}$. Для $\forall i=1..N$, $RE_i\in\Re^+$ — характеристика ресурсу Ri, $RT_i\in\{0$ і $\Re^+\}$ — стани ресурсів.
- $W_{ER}=\{WER_1,WER_2,...,WER_p\}$ — множина вагів дуг. Цю множину можна уявити у вигляді деякої матриці RC=RC [i, j]$\in\Re^+$, де i=1..N та j=1..N.

Потік M завдань, заданий множиною $V_J=\{J_1,J_2,...,J_M\}$, кожний елемент якої являє собою одне з M завдань та $J_i=\{JN_i, JE_i, JL_i, JM_i, JP_i\}$, де для $\forall i=1..N$:

- $JN_i\in N|$ — номер завдання,
- $JE_i\in\Re^+$ — обсяг роботи завдання.
- $JL_i=\{(R_L,\varphi_L),...,(R_q,\varphi_q)\}$, де $R_L\in VR$ — ресурс, з яким дане завдання вимагає обміну даними, $\varphi_{L+}\in\Re$ — обсяг передачі, L=1.. q, $q\in N|$.
- $Jm_i=\{0, R_i\}$ — маска завдання, де $R_i\in VR$ — номер ресурсу, на якому бажано виконувати дане завдання.
- $JP_i\in\Re^+$ — пріоритет даного завдання.

Серед виділених задач особливе значення мають дві: статичне безадресне укладання розкладу розподілу робіт та динамічне призначання (розподіл або перерозподіл) або реконфігурація паралельно виконуваних робіт між ресурсами.



Мал. 1 Загальна схема проходження задач скрізь ПОС

Для математичного формулювання задачи дінамичного розподілу завдань зостасовані наступні твердження:

Твердження 1: Г є відображення множини завдань $V_J=\{Job_1, Job_2,.., Job_M\}$ на множину ресурсів $V_R=\{R_1, R_2,.., R_N\}$ графу системи $G_R=(V_R,E_R,W_{VR},W_{ER})$, якщо результат відображення $\Gamma(V_J,V_R)$ є деяка множина А: $A=\{a_1, a_2,.., a_n\}$, де $a_i=(R^i, J^i)$, $R^i\in V_R$, $J^i\in V_J$, i=1..n, n∈N|.

Позначимо $AR=\{R^1, R^2,.., R^n\}$, $AJ=\{J^1, J^2,.., J^n\}$. Таким чином, $|A|=|AR|\cap|AJ|$, $AR\subseteq V_R$, $AJ\subseteq V_J$.

Твердження 2: відображення Г є розподіл завдань V_J на ресурси V_R, якщо його результат $\Gamma(V_J,V_R)=A$, де $A=\{(R^1, J^1), (R^2, J^2),.., (R^n,J^n)\}$ задовольняє наступної умові: для $\forall i=1..n$, $R^i\notin AR\setminus R^i$, $J^i\in AJ\setminus J^i$, де $AR=\{R^1, R^2,.., R^n\}$, $AJ=\{J^1, J^2,.., J^n\}$. Розміром наданого розподілення $\Gamma(V_J,V_R)$ є число n. Тоді $\Gamma(V_J,V_R)\to A$, $n=|A|$.

Твердження 3: результат розподілення завдань на ресурси $A=\Gamma(V_J,V_R)$ призначимо розкладом даного розподілення Г. Пару $a_i=(R^i, J^i)$, i=1..n будемо визначати призначенням для ресурса $R^i\in V_R$, та для завдання $J^i\in V_J$.

Твердження 4: нехай $X=\{A^1, A^2,.., A^z\}$, z∈N| є множина результатів усіх можливих розподілень для множини завдань V_J та для множини ресурсів V_R. Тоді $\Gamma(V_J,V_R)\equiv X$. Розподілення завдань на ресурси $\Gamma(V_J,V_R)\to A^*$ є максимальне розподілення для даних множин завдань V_J та множини ресурсів V_R якщо:

1) $n^*=|A^*|$;

2) $n^*=\max\{|A^1|, |A^2|,.., |A^z|\}$.

Твердження 5: нехай Δ є деяка функція від призначення $a_s=(R^s, J^s)$ (тобто призначення завдання J^s на ресурс R^s, $R^s\in V_R$ та $J^s\in V_J$). Тоді $\Delta(a_s)=\Phi$ або $\Phi=\Delta(R^s, J^s)$ та $\Phi_i=\Delta(a_i)=\Delta(R^i, J^i)$, де i=1..n будемо називати вагою призначення $a_i=(R^i, J^i)$ за Δ.

Для вирішення задач оптимізацування та розподіляння у динамічному режимі введемо функцію Δ для виміру ваги призначення завдання J_j на ресурс R_i ($R_i\in V_R$ та $J_j\in V_J$), яку в загальному випадку можна визначити таким чином [1]:

$$\Delta(Ri, Jj)=\delta_{i,j}=\prod_{k=1}^{K} P_k^{i,j}\times\prod_{x=1}^{H} C_x^{i,j}\times\sum_{y=1}^{G} PR_y O_y^{i,j}.$$

Де,

- $\prod_{k=1}^{K} P_k^{i,j}$ — величина абсолютного пріоритету призначення (R_i,J_j).
- $\prod_{x=1}^{H} C_x^{i,j}$ — результат аналізу всіх H обов'язкових вимог, де $C_x^{i,j}$ — ступінь виконання обов'язкової вимоги x для призначення завдання J_j на ресурс R_i, $C_x^{i,j}\in\{0,1\}$.

- $\sum_{y=1}^{G} PR_y O_y^{i,j}$ — результат аналізу всіх G оптимізувальних вимог, де $O_y^{i,j} \in \Re^+$ та $O^d \le O_y^{i,j} \le O^u$ — ступінь виконання оптимізівної вимоги у для призначення завдання J_j на ресурс R_i; $PR_y \in \Re^+$ та $L^d \le L_y \le L^u$ є ваговий коефіцієнт оптимізівної вимоги у.

Твердження 6: суму ваг усіх призначень $\{a_1, a_2, .., a_n\}$ будемо називати вагою D(A) розкладу A. Тобто: $D(A) = \sum_{i=1}^{n} \Delta(a_i)$.

Твердження 7: нехай $Xm=\{A_1, A_2, .., A_m\}$, $m \in N|$ є множина усіх максимальних розподілень для множини завдань V_J та для множини ресурсів V_R. Тоді розклад $A^*=\Gamma(V_J, V_R)$ є оптимальний розклад розподілення (завдань V_J на ресурси V_R) Γ за виміром Δ, якщо A^* задовольняє наступним умовам:

1) $A^*=\{(R^1, J^1), (R^2, J^2), .., (R^n, J^n)\}$ є результатом максимального розподілення для наданих множини завдань V_J та множини ресурсів V_R, тобто $|A^*|=\max\{|A^1|, |A^2|, .., |A^z|\}$; (твердження 5).
2) Вага розподілу $A^*=\{a_1, a_2, .., a_n\}$ повинна бути максимальною тобто:

$$D(A^*) = \sum_{i=1}^{n} \Delta(a_i) = \max\{D(A_1), D(A_2), \ldots, D(A_m)\} = \max_{j=1}^{m} \{D(A_j)\}$$

Ця задача, у загальному вигляді, у комбінаторної теорії формулюється як задача призначення та є NP-повною. Однак, у формулюванні що пропонується, вона зводиться до задачі пошуку максимального паросполучення у зваженому двочастковому графі.

Існує декілька поліноміальних методів для вирішування задачі призначення для зваженого двочасткового графа $G=(V_R, V_J, E, WE)$:

Де, $V_R=\{R_1, R_2, ..., R_{N_t}\}$ та $V_J=\{J_1, J_2, ..., J_M\}$,

$E=\{E_1, E_2, ..., E_d\}, E_k=\{R^*, J^*\}$, де $R^* \in V_R$ и $J^* \in V_J$,

Де k=1..d , $0 \le d \le N_t \times M$.

Найбільш відомим алгоритмом вирішування цієї задачі є Угорський, що містить у собі як підзадачу — пошук максимального паросполучення для незваженого двочасткового графа. У тому випадку, якщо планування виробляється для суворонеоднорідної ОС, то вхідна інформація надається у вигляді незваженого двочасткового графу.

Аналіз та властивості середовища реалізування задачі та двочасткового графа дозволив виділити деякі особливості, сформулювати та довести ряд тверджень та слідств, що були покладені у основу розробленого методу спрямованого пошуку плану розподіляння заявок за ресурсами, що характеризується меншою часовою складністю у порівнянні з Угорським методом.

Досліджування методів та алгоритмів пошуку максимального паросполучення [2] показують, що найбільші труднощі, які впливають на кількість кроків, а звідси і на час пошуку максимального паросполучення, виникають у графах, у яких перманент матриці зв'язуваності близький до "1" або рівний "0". В роботі доведено, що використання як теоретичної бази теорем Фробеніуса-Кеніга та Мінка, що стосуються досліджування властивостей перманентів (0,1) матриць, дозволяє виконати теоретичне обгрунтування нового підходу до розробляння алгоритмів спрямованого пошуку за формування розкладу в просторових планувальниках СРОД [2.3].

Провідна ідея нового підходу, що спрощує рішення задачі пошуку максимального паросполучення, полягає в розподілі його на декілька етапів, коли власне рішенню передує швидке аналізування вхідної інформації та вироблення стратегії її подальшого вірішування. Додання додаткових кроків значно меншої часової складності, ніж сам алгоритм, не впливає на теоретичну оцінку часової складності алгоритму в цілому, однак, дозволяє наступне: зменшити вимірність рішення задачі за рахунок виділяння призначень, що обов'язково треба зробити, та виділити призначення, що робити не можна, та за рахунок цього уникнути зайвих перевірок на можливість входження їх у рішення. Окрім цього, на етапі підготовки вхідної інформації можливо обчислення потужності максимального паросполучення. Запропоновані процедури базуються на базових теоремах та слідствах з них, наведених та доведених у роботах [2,3,4,5].

Для доказу теорем та слідств прийнято наступне визначення:

Множина A: $A=\{a_1, a_2, ..., a_n\}$, де $a_k=(R^k, J^k)$, $R^k \in V_R$, $J^k \in V_J$, $k=1..\ n$, $n \in N$| називається паросполученням графа G або його матриці зв'язуваності RJ якщо $A=\{(R^1,J^1),(R^2,J^2),..,(R^n,J^n)\}$ задовольняє таким умовам: Для $\forall$ $k=1..\ n$:

- $R^k \notin AR \backslash R^k$,
- $J^k \notin AJ \backslash J^k$,

Де $AR=\{R^1,R^2,...,R^n\}$, $AJ=\{J^1,J^2,...,J^n\}$.

Таким чином, $AR \subseteq V_R$, $AJ \subseteq V_J$.

Або підмножина A ребер графа G=(V, E) називається паросполученням, якщо жодні з двох ребр з A не мають спільної вершини.

Теорема 1. Якщо у матриці RJ[i, j], i=1..N, j=1..N графу $G=(V_R,V_J,E)$, де $V_R=\{1,2,...N\}$, $V_J=\{1,2,..N\}$, існує рішення A потужністю n=N та існують такі вершини p, q що:

RJ [p, q] =1,

RJ[p,j]=0 $\forall$ $j \in \{1,...,N\} \backslash q$ та/або

RJ[i,q]=0 $\forall$ $i \in \{1,...,N\} \backslash p$.

Тоді ця пара (p, q) завжди бере участь в рішенні A, (p, q)∈ A.

Теорема 2. Якщо у матриці RJ[i, j], i=1.. N, j=1.. N,∃ FA (віяло):

- FA={(R^1,q), (R^2,q),..., (R^f,q)}, R^k∈{1,..,N}, 2≤ f ≤N, де RJ[R^k,q]=1 для ∀k=1..f і RJ[R^k,J^k]=0, ∀J^k ∈{1,..,N}\q або
- FA={(p,J^1), (p,J^2),..., (p,J^f)}, J^k∈{1,..,N}, 2≤ f ≤N, де RJ[p,J^k]=1 ∀k=1..f і RJ[R^k,J^k]=0, ∀ R^k ∈{1,..,N}\ p,

тоді будь-яка з вершин FA входить у один з варіантів максимального паросполучення, задача призначення не має повного рішення та потужність максимального паросполучення M<N-f+1.

Слідство 1. Вимірність рішення задачі пошуку максимального паросполучення може бути зменшена на кількість пар вершин, що визначаються за Теоремами 1 та 2, а пошук паросполучення повинен проводитись на основі нового сурграфа, одержаного після редукції вхідного графу.

Следствие 2. Суміжні ребра, що інцідентні вершинам, що визначенні за Теоремами 1, 2 повинні бути усунуті з подальшого розгляду, а вхідний граф редуцирован та перетворен в новий сурграф.

Теорема 3. Якщо у матриці RJ[i, j], i=1..N, j=1..N можна виділити підматрицю MM розміром T*S, де MM[k, l] =RJ[k, l], k=1..T, l=(N-S+1)..N, S+T>N і MM=θ (θ - нульова матриця), тоді задача призначення не має повного рішення.

Теорема 4. Якщо у матриці RJ[i, j], i=1..N, j=1..N, можна виділити підматрицю MN[k, l], k=1..T, l=(N-S+1)..N, де S+T=N і MN=θ, тоді для Ri{1∈,..,(N-S)}, Jj∈ {(T+1),..,N}: ∀(Ri, Jj) ∉A. І всі ці (Ri, Jj)∉A повинні бути обнулені при пошуку повного рішення.

Теорема 5. Якщо в матриці RJ[i, j], i=1..N, j=1..N, можна виділити декілька підматриць MN, які задовольнять Теоремі 4, то всі відповідні симетричні їм відносно головної діагоналі підматриці є "конфліктними" та повинні бути обнулені.

Теорема 6. Якщо в матриці RJ існує підматриця MT така, що ∀k=1..T, l=(N-S+1)..N: окрім MT[T, (N-S+1)] =1, всі інші елементи MT[k, l]=0, де S+T=N+1, тоді призначення (T,(N-S+1)) завжди беруть участь в одному з варіантів рішення.

Теорема 7. Якщо в матриці зв'язності (МЗ), відповідній двочастковому графу вимірності N, можна виділити нульову підматрицю S×T, де S+T>N, то задача не має варіанту повного рішення та потужність паросполучення (M) рівна 2N-(S+T).

На підставі наведених вище теорем рішення о призначенні завдання на ресурс виконується у відповідності з наступним твердженням:

Твердження 8: якщо в графі:

1. відсутні обов'язкові призначення (Теореми 1, 2, 6, Слідства 1, 2);
2. усунуті конфліктні призначення (Теореми 4, 5);
3. per(MC)>1 (Теореми 3, 7)

тоді будь-яка "1" МС, тобто ребро двочасткового графу приймає участь в одном з варіантів паросполучення та разом з інцидентними їй вершинами повинне бути взяте у решення, а граф редуцирован та пошук наступної пари повинен вестися з нового сургафа.

Очевидно, що виконання декомпозиції вхідного графа на підставі наведених теорем, зменшує часову складність пошуку рішення за застосування більшості відомих алгоритмів, наприклад, під час пошуку рішення алгоритмом Карпа-Хопкрофта $O(n_1{}^{1.5}m) >> O(n_1{}^{1.5}m_1)+O(n_2{}^{1.5}m_2)$, де n1+n2=n, а (m1+m2)≤m. Окрім цього, реалізування наведених вище теорем, за чинності обмежень на показник неоднорідності РОС, дозволяє для зменшення часу очікування використати більш простий алгоритм пошуку максимального паросполучення, заснований за принципом вилучального планування, та виконувати активізацію розподілених процесів під час рішення задачі планування. Вилучальне планування у поєднанні з алгоритмом просторового планування дозволяє наблизити часову складність динамічного планування до лінійної з покращеними характеристиками оптимізації у порівнянні з алгоритмами планування за списками.

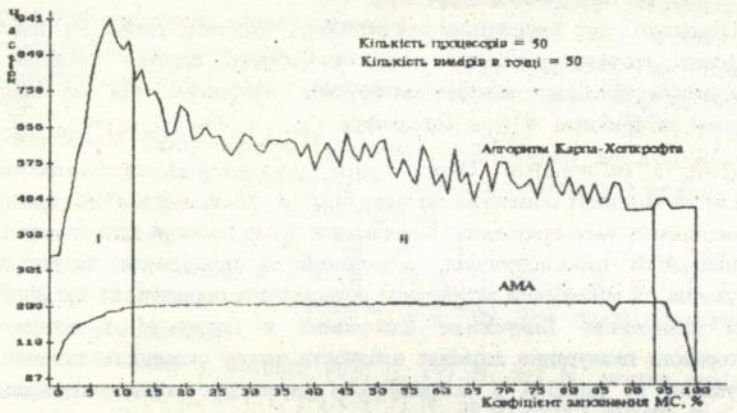
На підставі наведених теорем та слідств розроблено адаптивний алгоритм мультианаліза (АМА), який виконує пошук максимального паросполучення у незваженому двочастковому графі.

АМА є складником запропонованого методу спрямованого пошуку рішення та АСП, який реалізує його для динамічного розподіляння заявок за ресурсами у гетерогенній ОС. Основою АСП є Угорський метод, що має часову складність $O(n^3)$, де n — кількість вершин графа відношення завдань-ресурсів JR. Застосування АМА у АСП, за заданих обмежень на неоднорідність (1≥Нгет≥0.7) дозволило отримати розрахункову часовую складність АСП $O(n^{1.5}\log(n))$). У випадку однорідної системи при Нгет⇒ 0 часова складність прагне до $O(n^2)$.

Для підтвердження теоретичних висновків та розрахункової часової складності АСП та АМА у роботі розроблена моделююча система, що дозволяє порівнювати запропоновані та альтернативні алгоритми. Як альтернативні алгоритми взяті: алгоритм, який базується на застосуванні

засобу Отжега (SA) та алгоритм із застосуванням оцінкової функції (EF). Окрім цього, виконалося порівняння запропонованого просторового планувальника з планувальниками, яки базуються на використовуванні дисциплін обслуговування FIFO та EASY. Порівняння алгоритмів планування у НРСОД виконувалося за наступними параметрами: абсолютний час планування, часова складність алгоритмів

На мал. 2 показані результати порівняльного статистичного моделювання залежності часу вирішення задачі пошуку максимального паросполучення алгоритмом Карпа-Хопкрофта та АМА.

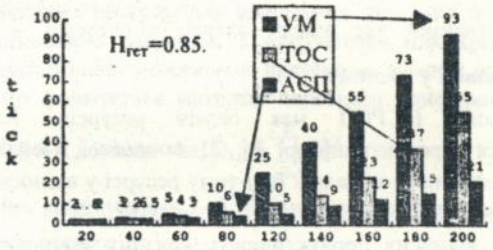


Мал. 2 Порівняння абсолютного часу планування за алгоритмами Карпа-Хопкрофта та АМА

На мал. 3 показані результати порівняльного статистичного моделювання залежності часу вирішення задачі розподіляння ресурсів від неоднорідності ОС (мал. 3а) та вимірності задачі (мал. 3б) Угорським методом (УМ), АСП, та за теоретично визначенною обчислювальною складністю (ТОС).

Незважаючи на більші значення часу планування у порівнянні з EF, АСП за рахунок оптимізування розподілу зменшується середній час очікування заявок у черзі, у порівнянні з лінійними планувальниками (ЛП), яки використовують дисципліни обслуговування типу FIFO або EASY (мал. 3). За критерій порівняння був прийнятий відносний коефіцієнт випередження Ко за коефіцієнту неоднорідності 0.8 та кванті активізації просторового планувальника рівного середньому часу тривалості оброблювання заявок, які надходять.

Мал.3а Абсолютний час планування



Вимірность рішення задачі

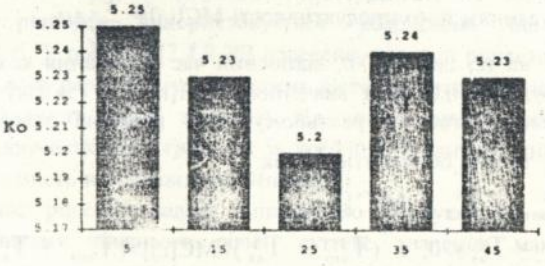
Мал.3.б Порівняння часової складності

Ко=(n2-n1)/n*100%

Де: n1 - кількість заявок в черзі планувальника АСП на момент часу t_i

n2 - кількість заявок в черзі планувальника ЛП на момент часу t_i

n - загальна кількість заявок, які надійшли в ОС до моменту часу t_i



Мал. 4

Низька часова складність АСП та малий абсолютний час роботи просторового планувальника, побудованого на його основі, дозволяють використовувати цей алгоритм у системах реального часу. У роботі запропоновано метод динамічного розподілу заявок за ресурсами неоднорідної розподіленої обчислювальної системи у реальному часі, що вирішується планувальником нижнього рівня. Планування у *реальному* часі характеризується формуванням плану розподіляння сукупності задач з заданим часом виконання та обмеженнями по часу виходу задач із системи. Система планування повинна мінімізувати сумарний час відхилення реального виходу задач з системи (виконання термінів рішення) від вхідних часових обмежень: системі планування повідомляється час рішення — $T^i_{\text{п рац}}$ (час виконання і-ої роботи на обчислювальному вузлі СРОД, що має максимальну продуктивність, повинно бути задано у відносних одиницях) та крайній час виходу заявки з системи — $T^i_{\text{вих}}$, та час надходження заявки у систему — $T^i_{\text{вх}}$.

Окрім цього НСРОД має безліч ресурсів. Кожний ресурс характеризується продуктивністю R^j. Для кожної і-ой роботи можна визначити час рішення і-ої задачі на j-тому ресурсі у відносних одиницях.

Для цього визначимо R_{max} як $R_{max} = \max\{R^1,..,R^n\}$.

Визначимо відносну продуктивність кожного обчислювального вузла R^j у відношенні до Rmax. Для цього обчислюємо множину $Z^j = \frac{R^j}{R_{max}}$.

Маючи відносну продуктивність кожного вузла, можна визначити відношення робота-ресурс з урахуванням часу виконання роботи та продуктивності ВУ.

Для цього формується матриця претендування (МП), де кожний елемент визначає відносний час і-ої роботи, що виконується на j-ому ресурсі з урахуванням його продуктивності $MC[i,j] = \frac{T^i_{\text{п рац}}}{Z^j}$.

Зважаючи на те, що $1 \geq Z^j > 0$, відносний час виконання кожної роботи на ресурсах буде визначатися значенням МП[i, j]. У зв'язку з тим, що задача, яка виконується у реальному часі, повинна завершитися до заданого $T^i_{\text{вих}}$, то слід обчислити $\Delta_t^{i,j}$ як:

$$\Delta_t^{i,j} = T^i_{\text{вих}} - T^i_{\text{вх}} - MC[i,j]$$

Бо $(T^i_{\text{вих}} - T^i_{\text{вх}}) > 0$, а $(T^i_{\text{вих}} - T^i_{\text{вх}}) > MC[i,j] > (T^i_{\text{вих}} - T^i_{\text{вх}})$, то $\Delta_t^{i,j}$ може набувати як позитивних так і негативних значень. Позитивні значення $\Delta_t^{i,j}$ означають, що робота і призначена на ресурс j буде

завершена до призначеного терміну на час $\Delta_t^{i,j}$, а негативні значення означають, що вихід заявки з системи відбудеться пізніше призначеного терміну. Під час обчислення $\Delta_t^{i,j}$ слідує враховувати, що на час планування($T_{план}$) всі заявки мають один й тойже базовий час початку планування незалежно від часу надходження заявок у систему. Окрім цього слід враховувати й час роботи планувальника, тобто час планування є.

Тому можна прийняти $T_{вх}^{i} = T_{план} + \varepsilon$.

Інформація у матриці $\Delta_t^{i,j}$ враховує тільки об'єм роботи та відносний час її виконання з урахуванням продуктивності обчислювальних вузлів. Однак неоднорідність СРОД зумовлює необхідність врахування індивідуальних характеристик кожного обчислювального вузла (наявності достатньої оперативної пам'яті, наявності необхідних даних у вузлі, наявності необхідних програм для виконання завдання й т.д.), зв'язаних з можливістю виконання завдання. З урахуванням цих критеріїв необхідне оцінювання принципової можливості виконання кожної роботи у кожному вузлі. Для цього формується матриця перевірки конфлікності призначень кожної і-ої роботи у j-ий вузол - $Q_{i,j} = \prod_{x=1}^{p} C_x^{i,j}$,

де: $C_x^{i,j}$ - ступінь виконування х-ої обов'язкової вимоги для призначення і-ої заявки на j-ий ресурс;

На наступному етапі формування вхідної інформації виконується фільтрація елементів матриці $\Delta_t^{i,j}$ у відповідності до значень елементів.

У такому порушенні призначення і-ої роботи на j-ий ресурс зводиться до задачі пошуку максимального паросполучення у зваженому двучастковому графі та може буте вирішане Угорським методом або АСП. Слід зазначити, що пошук варіанту розподілу виконується для неоднорідної обчислювальної ситеми (ступінь неоднорідності не нижче 0.7) та за пошуком рішення використовується розріджена бінарна матриця зв'язуваності. У роботах [2,4,9,20] доведено, що при рішенні задачі пошуку максимального паросполучення можна застосовувати адаптивний алгоритм, часова складність якого залежить від коефіцієнта заповнення матриці зв'язності двучасткового графа, а за коефіцієнту заповнення меньш 0.35 - часова складність не перевищує $O(n^{1.5}\log n)$.

Під час рішення задачі динамічного планування, що вирішується планувальниками нижнього рівня, суттєве значення мають результати рішення задач попереднього підготування паралельних програм для занурювання у паралельну систему, тобто задачі статичного планування.

Практично всі вони вирішуються засобами комбінаторіки та являють собою NP-повні задачі.

Для комбінаторних задач особливе значення має визначення кількісної оцінки якості рішення та оцінка ступеня наближення варіанту рішення до оптимального.

У роботі запропоновано метод локалізації зони пошуку розміщення робіт у СРОД із загальною пам'яттю, повнозв'язною ОС та із загальною шиною. Запропонований аналітико-графічний метод [16] дозволяє шляхом структурного аналізу вхідного графа задач, заданого у вигляді зваженого ациклічного орієнтованого графа (АОГ), визначити кількість поцесорів СРОД та теоретично оптимальний час планування для заданої ПОС.

У результаті структурного аналізування АОГ визначається нижнє (N_{low}) та верхнє (N_{high}) граничне число процесорів.

Під нижнім граничним числом процесорів N_{low} мається на увазі таке число процесорів, за яким теоретично можливо досягнення максимальних значень коефіцієнтів ефективності, використовування та якості рішення множини взаємопов'язаних робіт у T_{cr}. У тому випадку, якщо не можна вийти за межі критичного шляху, то кількість процесорів не може бути менше ніж N_{low}, тобто теоретично можливого граничного числа процесорів, та не повинно перевищувати N_{high}. Перевишення N_{high} призводить до різкого погіршення характеристик системи планування за кількістю необхідних пересилань.

N_{high} та N_{low} можуть розглядатися як критерії оцінювання системи планування у ОС із загальною пам'яттю, коли часами пересилок зневажають.

Виявлено, що традиційна оцінка N_{low} як $N^{I}_{low} = \left]\frac{T_z}{T_{cr}}\right[$,

де: $T_z = WG_{(n)} = \sum_{i=1}^{n} WV_i$ дає хибні оцінки для АОГ, що мають ромбоподібні фрагменти. У роботі запропонована нова оцінка N_{low}

$N_{low} = \max\left\{N^{I}_{low}, N^{II}_{low}\right\}$.

де: $N^{II}_{low} = \max\left\{\left]\frac{T_L}{\max\{t_i\}}\right[\right\}$ o(x) - порядкова функція АОГ,

$1 \le i \le o(x)$

N_{high} визначається з виразу $N_{high} = \max\limits_{1 \le i \le o(x)}\{N_{L_i}\}$,

де N_{L_i} — кількість вершин в рівні (ярусі) Li

Для визначення залежності комунікаційних витрат від числа

процесорів, що використовуються, застосовується алгоритм сканівного занурення АОГ з "зануленням" ваг вузлів та використовуванням стратегії виділяння, що модифікує критичний шлях на кожній ітерації та вилучує виділянні вершини із подальшого розгляду.

Порівнювання розкладів складених для вирішення задачі вилучення Гауса та швидкого перетворення Фурье відомими алгоритмами EZ, ETF, DLS, MCP, DSC MD, DCP, алгоритмами кластерного планування Шаркара, Янга із запропонованим у роботі сканівним просторово-часовим алгоритмом (СПЧА) вказало на високу ефективність методу та точність теоретичного припущення знаходження оптимального варіанту. Так, наприклад, порівнюючи з планом занурювання рішення задачі швидкого перетворення Фурье були одержані такі наступні результати (час планування даний в умовних тактах):

	ПОРІВНЯЛЬНІ АЛГОРИТМИ				
Приклад	Лінійна кластерізація	Алгоритм OSCAR	Дво-прохідний алгоритм Кіма	SA алгоритм Метод Отжега	Запропонований алгоритм СПЧА
1	519	533	533	519	507
2	587	587	540	540	505

Основною метою статичного планування є знаходження оптимального або оптимізованого плану занурювання АОГ у задану кількість процесорів. У роботі наведено, що попередній структурний аналіз вхідної інформації та надання АОГ у ярусно-паралельній формі (ЯПФ) спільно з комбінуванням стратегій раннього та пізнього планування, дозволяють поліпшити алгоритми пошуку всіх вершин, яки входять у критичні шляхи Флойда-Уоршелла та Дейкстри, та отримати новий алгоритм із часовою складністю O(2E+2N). Для вершин яки не є критичними, запроваджено поняття "транзитні". Транзитна вершина може належати трьом категоріям:

1. "справжня транзитність" — зміна рівня ЯПФ приналежності вершини не збільшує критичний шлях;

   Умова визначення вершин "справжньої транзитності" —

   $$\forall X_i \in V_i \Leftrightarrow \exists X_m = succ(X_i) \in V_j \,/\, j > i+1$$

2. "несправжня транзитність" — зклеювання (кластерізація) вершин цієї категорії на одному рівні ЯПФ не збільшує критичний шлях;

   Вершини "несправжньої транзитності" володіють властивістю:

   Якщо $\forall x_i \in v_i \Leftrightarrow \exists e_i = (x_j, x_i) \in E_{i-1} \,/\, x_j \in v_{i-1}; x_i \in v_i$

Де: v_i множина вершин рівня "i" ЯПФ

E_{i-1} множина дуг між

$$(v_{i-1}, v_i) / \forall x_i, x_j \in v_i \Leftrightarrow T_{x_i} \geq T_{x_j}$$

3. "мультиплікативна транзитність" — зклеювання (кластерізація) вершин цієї категорії, що належать різним рівням ЯПФ не збільшує критичний шлях;

X_i^m — вершини володіють властивістю "мультиплікативної транзитності" між рівнями "i" та "m",

$$\text{Якщо } \exists X_p \in V_i / T_{x_p} > \sum_{K=i}^{m} T_{x_i} .$$

Введення транзитності та класифікація вершин за видом транзитності на етапі попереднього структурного аналізу АОГ дозволили запропонувати сканівний просторово-часовий алгоритм знаходження оптимізованого плану занурювання АОГ у задане число процесорів ПОС із спільною пам'яттю. Головна ідея нового підходу полягає не у покроковому багатопрохідному укладанні розкладу із використанням стратегій раннього планування та оцінювальних функцій, а у формуванні спочатку базового рішення за знятих обмеженнях на число процесорів, та після цього горизонтального та вертикального кластерізування вершин АОГ. Мета кластерізування — балансування АОГ за рівнями та за часом планування, рівному T_{cr}. Під час кластерізування враховується класифікаційна характеристика кожної вершини за "транзитністтю". Цей підхід дозволив звести NP-повноту задачі великої вимірності до NP-повноти задач малої вимірності, тому що виявлення "транзитності" вершин дозволяє вести пошук варіанту їх кластеризіації тільки у рамках виділених рівнів графа. Для виконання процедур горизонтального кластерізування у роботі запропоновано евристичний алгоритм, який оптимізує пакування робіт. У алгоритмі використане суміщення стратегій "найбільш" та "найменьш" доцільних для розрахунку пріоритетів робіт, що підлягають горизонтальному кластеризуванню.

Запропоновані у роботі методи та засоби орієнтовані на вирішення задач динамічного планування у НСРОД. Однак результати виконаних теоретичних досліджень можуть мати більш широке застосування. Наприклад їх можна використовувати під час планування розподілу робіт у кластерних робітничих станціях. Одною з переваг кластерних робітничих станцій є наявність простих засобів для збільшення загальної продуктивності ПОС шляхом додавання багатопроцесорних плат (кластерів). Однак "кластерна" архітектура робітничих станцій, де кожен

кластер являє собою багатопроцесорний модуль із загальною пам'яттю або повнозв'язну архітектуру, а кластери з'єднані між собою за допомогою загальної шини, вимагає застосування нових алгоритмів планування та диспетчіювання. Кластери, як правило, мають структуру, що дозволяє не враховувати час на обмін даними між обчислювальними елементами (спільна пам'ять) або звести ці витрати до мінімуму (повнозв'язна або симетрична система). Однак за такою архітектурою обмін даними між кластерами повинен враховуватися та минімізуватися, тому що обмін здійснюється через ресурси, які поділяються (у часі). На підставі теорем 3, 4 у роботі запропоновано алгоритм для вирішення задачі розбиття множини взаємопов'язних робіт заданих у вигляді графіка Ганта на підмножини.

У загальному випадку дана задача зводиться до рішення задачі пошуку мінімальної підмножини сполучень у наступному формулюванні.

Нехай задан зв'язаний граф G=(E, U),

де: E — множина вершин $E=\{V_1, \ldots, V_n\}$,

U — множина дуг $U=\{W_1, \ldots, W_m\}$,

$W_i = V_p V_q$; $Vp \in E$; $Vq \in E$; $i = \overline{1,m}$.

Встановлено деяке число розбиття графу $K \geq 2$

Необхідно знайти K підмножин вершин у графі G:

$G_1=(E_1,U_1)$;

$G_2=(E_2,U_2)$;

.....

$G_k=(E_k,U_k)$;

$E_1 \cup E_2 \cup .. \cup E_k = E$,

таких щоб: $|U \setminus (U_1 \cup U_2 \cup .. \cup U_k)|$ було мінімальним.

Запропоновано підхід, який заснован на аналізі вхідного графіка Ганта, за наслідком якого виконується редукція некритичних комунікацій та формування основних підматриць і на підставі їх аналізу провадиться виділяння робіт, що виконуються на кожному кластері робітничої станції.

Наведені в роботі результати моделювання дозволяють зробити висновок про вірогідність запропонованих методів та теоретичні дослідження. Часова складність запропонованих алгоритмів відповідає розрахунковій. Порівняльний аналіз результатів планування запропонованими алгоритмами з найбільш відомими за такими показниками як: час планування, число процесорів, що використовуються, пропускна спроможність ПОС, показує їхню високу ефективність за заданих обмежень.

ОСНОВНІ НАУКОВІ ТА ПРАКТИЧНІ РЕЗУЛЬТАТИ РОБОТИ ПОЛЯГАЮТЬ В НАСТУПНОМУ

1. Запропонована порівняльна класифікація систем паралельного обробляння даних. Виконано аналіз та порівняння методів та засобів організації обчислювальних процесів та аналіз структурної організації СРОД: виділені особливості паралельних ОС та на цій основі визначені можливості ефективного організування обчислювальних процесів на всіх етапах життєвого циклу робіт та процесів.

2. Проведено аналіз та досліджені особливості систем планування, що використовуваються для керування та диспетчіювання; вказано на необхідність розробляння нових багаторівневих систем планування, які базуються на стратегії декартово-покоординатного спуску, підвищення пропускної спроможності ОС та забезпечення безпеки функціонування розподілених систем обробляння інформації.

3. Запропонована та досліджена модель системи планування у СРОД, дано математичне формулювання задачі розподілу робіт за ресурсами. Доведено, що її можна звести до задачі пошуку максимального паросполучення у зваженому або незваженому двочастковому графі. Запропоновано метод покрокового конструювання розподілу робіт, який базується на спрямованому пошуку рішення шляхом декомпозування вхідної інформації та виробляння стратегії адаптивного налагоджування диспетчера. Для звуження області пошуку рішення запропоновано застосування теорем Мінка та Фробеніуса-Кеніга, які стосуються досліджень властивостей перманентів (0.1) матриць. Доведено, що попереднє (на етапі підготовки вхідної інформації) вилучення "обов'язкових" та "конфліктних" призначень, істотно скорочує часову складність алгоритмів пошуку максимального паросполучення за рахунок зниження вимірності рішення та можливості використання вилучального принципу планування.

4. Запропонована та досліджена модель системи планування робіт у неоднорідній розподіленій системи обробляння даних (НСРОД); виконано математичне фомулювання задачі динамічного планування. Запропонован новий підхід до вирішення задачі розподіляння робіт в НСРОД, який використовує модифікований Угорський метод, що дозволяє знизити часову складність, мінімізувати час планування та виконання завдань (зокрема зменшити витрати на комунікаційні пересилання). Показана можливість використовування запропонованого методу динамічного

планування для створення динамічних просторових планувальників реального часу.

5. Розроблено метод сканівного просторово-часового планування взаємопов'язаних робіт, у рамках якого реалізовано новий підхід до вирішення задачі укладання розкладу, оснований на ідеї структурно-семантичного аналізу графу вхідної задачі та виконання на його основі ієрархічного декомпозиціювання обчислювального процесу на графовім рівні. Це дозволило виконати сканівну покрокову оптимізацію базового варіанту розкладу, побудованого у просторово-часових координатах без врахування обмежень на кількість обчислювальних елементів, а також: спрямований аналіз базового варіанту розподілу робіт, детерміноване декомпозиціювання вхідного графа, виділяння критичних зон, рівнів, вузлів, виробляння стратегії оптимізування плану за часом або кількістю процесорів у відповідності до функції мети.

6. На основі запропонованого структурно-семантичного аналізу графа розроблено метод попереднього локалізування зони пошуку оптимізованого варіанту розміщення інтервально-упорядкованих робіт у обчислювальному середовищі з мінімальними комунікаційними витратами та часом планування у ПОС різноманітної топології.

7. Запропоновано новий алгоритм пошуку вершин, які входять у критичні шляхи, що у поєднанні зі структурно-семантичним аналізом та декомпозицією вхідного графа, дозволили виділити у графі "резидентні", "транзитні" та "мультитранзитні" вершини, звести експоненціальну часову складність задачі великої вимірності до NP-повноти малих дискретних задач та завдяки цьому отримати прийнятний час її вирішування.

8. Розроблено інтерфейсне середовище моделювання, що дозволяє: 1). створити діалогове інтерфейсне інформаційне середовище, яке дає змогу виконувати автоматизоване конструювання плану занурювання розпаралеленої програми, заданої у вигляді ациклічного орієнтованого графа в ПОС; 2). виконувати широкомасштабні порівняльні досліджування різноманітних методів та засобів планування, що дозволило підтвердити високу ефективність запропонованих у дисертаційній роботі засобів планування.

Результати дисертації відображені у таких наукових роботах.

1. Симоненко В.П. Организация вычислительных процессов в ЭВМ, комплексах, сетях и системах, Киев, "Век +", 1997 г., 304 стр.
2. Симоненко В.П., Организация вычислительных процессов в параллельных вычислительных системах, Москва, ЦНТИ "Информсвязь", 1996 г., 150 стр.

3. V.P. Simonenko, Blanca Hernandez, K.G. Samofalov, La organizacion del proceso de calculo de los sistemas decomputacion. Cuba, Santa-Clara, 1985 г. (учебное пособие на исп. языке), 191 стр.
4. Самофалов К.Г., Симоненко В. П., Автоматизация составления расписания занятий в ВУЗе, Киев,МВ и ССО УССР ,1973 г., 45 стр.
5. Симоненко В.П., Обоснование выбора метода динамической диспетчеризации работ в распределенных ВС/ "Электронное моделирование", Том 19, № 3, 1997 г., стр.51-65.
6. Луцкий Г.М., Самофалов К.Г., Симоненко В.П., Метод предварительного анализа исходной информации и выработки стратегии решения задачи назначения в распределенной вычислительной среде, Ж., "Электронное моделирование", № 4, 1995 г., стр. 36-41.
7. Самофалов К.Г., Симоненко В.П., Пространственно-временная модель планирования работ в параллельной вычислительной системе, Ж. "Электронное моделирование", № 4, 1996 г. стр. 22-29.
8. Pham Hong Hanh, Simonenko Valery, Task Assignment for Scheduling Jobs and Resources in Parallel Distributed Systems, Journal of Computer Sciense and Cibernetics, T. 12, S., 1996,pp. 64-76.
9. V. Simonenko, Principales esferas de aplicacion de la computacion en la educacion, r. Educacion, No. 29, 1978, pp. 19-24.
10. V. Simonenko, "Objective-Oriented Algorithm for Job Scheduling in Heterogeneous Parallel Systems, Second international conference on parallel processing and applied mathematics - Zacopane, Poland, 1997. Part 2, pp.378-387.
11. Самофалов К.Г., Тартаковский Э.И., Симоненко В.П., Кулаков Ю.А., Автоматизация диспетчерской службы вуза, Ж. "Механизация и автоматизация управления, №3, К., 1972 г., стр. 21-27.
12. G. Loutsky, Samofalov K.G., V. Simonenko, A method for analyzing the initial information and elaborating the solution strategy for the sceduling problem in a parallel computing system, Engineering Simulation, Vol. 13, pp. 601-610, 1996.
13. G. Loutsky, V. Simonenko, J. Kanevski, A new approach in solving problems in a distributed computer environment during dynamic schduling, First international conference on parallel processing and applied mathematics- Poland,1994.84-91 стр.
14. Pham Hong Hang, Valery Simonenko, A new algorithm and simulation for task assignment in parallel distributed systems, Conference ESM-96, Budapest, 1996, pp. 95-99.
15. Valery Simonenko, Pham Hong Hang, Adaptation of algorithm for job-resurce assignment in heterogeneous distributed systems, Conference PDPTA "96, Sunnyvale, California USA, 1996. pp.835-846

16. Pham Hong Hanh, Valery Simonenko, Objective-Oriented Algorithm for Job Scheduling in Parallel Heterogeneous Systems, (11th International Parallel Processing Symposium (г. Женева, Швейцария, 1-5 Апреля 1997.) pp.127-145
17. Самофалов К.Г., Симоненко В.П., Применение ЭВМ для составления расписания занятий, КГУ, Сб. "Проблемы высшей школы", №6, 1971 г., стр. 147-153
18. Симоненко В.П., Кулаков Ю.А., Копычко С.Н., Об одном алгоритме определения существование систем различных представителей, ИК АН УССР, Сб. "Теория задач и способов их решений" 1975 г., стр. 78-80.
19. Кулаков Ю.А., Симоненко В.П., . Тартаковский Э.И, Алгоритм сжатия информации для составления расписания занятий, ИК АН УССР Сб. "Методы построения систем подготовки пользователей", 1974 г., стр. 85-90.
20. В.П. Симоненко, Э.И. Тартаковский, Некоторые вопросы диспетчеризации ВУЗА,Вестник киевского политехнического института, №9, 1972г., стр. 92-95.
21. Литвин У.А., Симоненко В.П., Решение одной задачи составления расписания на ЦВМ "МИР", КГУ, Сб., " Механизация учета и вычислительных работ", вып. 12 1972 г., стр. 41-52.
22. Самофалов К.Г., Симоненко В.П., Автоматизация составления и оценка качества расписания, КГУ, Сб., Проблемы выс -шей школы, вып.11, 1972г, стр. 54-60.
23. Симоненко В.П., Разработка и исследование алгоритмов, программ и технических средств для системы планирования учебного процесса ВУЗа, Автореф. на соискание ученой степени к.т.н., Киев, 1972. 16 стр.
24. Корочкин А.В., Русанова О.В., Симоненко В.П., Методические указания к выполнению курсовых работ по курсу " Системное программное обеспечение",(Учебное пособие) Киев, КПИ, 1993 г. 20 стр.
25. Луцкий Г.М., Кулаков Ю.А., Симоненко В.П., Корочкин А.В., Методические указания к лабораторному практимуму по дисциплине " Мультимикропроцессорные системы" для слушателей спецфакультета, Киев, КПИ , 1987 г. 52 стр.
26. Луцкий Г.М., Симоненко В.П., Методические указания по курсу "Организация вычислительных процессов на ЦВМ", (Учебное пособие) Киев, КПИ , 1979 г. 42 стр.
27. Луцкий Г.М., Кулаков Ю.А., Симоненко В.П., Методические указания и контрольные задания по дисциплине "Организация вычислительных процессов на ЦВМ", (Учебное пособие) Киев, КПИ , 1986 г. 32 стр.
28. Симоненко В.П.,Корочкин А.В., Ордынский В.В., Методические указания к выполнению лабораторных работ по курсу "Системное программное обеспечение", (Учебное пособие) Киев, КПИ , 1993 г. 44 стр.

29. Луцкий Г.М., Симоненко В.П., Кулаков Ю.А., Методические указания к проведению самостоятельной работы по дисциплине "Организация вычислительных процессов на ЦВМ", (Учебное пособие) Киев, КПИ , 1989 г. 52 стр.
30. Симоненко В.П., "Новый метод динамической реконфигурации вычислительных процессов в распределенных системах", Сб. трудов "Информсвязь", ВИНИТИ 15.02.96, № 20-83-СВ-96 Москва, 1996 г. , стр. 49-58.
31. Симоненко В.П., С. Лохиани, Симоненко А.В. "Алгоритм распределения работ в параллельной вычислительной системе", Деп. в ВИНИТИ 07.02.95, № 278-Ук95 Киев, 1995 г. стр. 15
32. Лахиани С., Симоненко В.П., "Сравнительный анализ вычислительной сложности решения задачи о максимальном потоке сети", Деп. в ВИНИТИ 07.02.95, № 279-Ук95 Киев, 1995 г., стр. 17.
33. Лахиани С., Симоненко В.П., "Сравнительный анализ вычислительной сложности решения задачи поиска максимального паросочетания", Деп. в ВИНИТИ 07.02.95, № 280-Ук95 Киев,1995 г., стр. 20
34. Симоненко В.П., Адаптивный алгоритм диспетчеризации работ в распределенной вычислительной системе, Материалы международного форума информатизации МФИ-96, М. 1996 г., стр.77-79.
35. Симоненко В.П., Структурно-семантический анализ графа задач для системы статического планирования распределенной ВС, Труды II межрегионального научно-технического семинара "Применение пластиковых карт и защита информации", г. Сергиев-Посад, Россия. стр. 82-83.
36. Симоненко В.П., Ю.С. Каневский, Адаптивное динамическое планирование в распределенной вычислительной системе, Матер. международной конф., " Проблемы алгебры и кибернетики", Гомель 1995., стр. 162-164
37. Симоненко В.П., Каневский Ю.С., Симоненко А.В., Алгоритм статического планирования работ в параллельной вычислительной системе, Матер. международной конф. "Проблемы алгебры и кибернетики", Гомель 1995., стр.160-162
38. Симоненко В.П., Симоненко А.В., Метод пошагового конструирования для решения задачи назначения в распределенных вычислительных системах, Матер. международной конф. " Проблемы алгебры и кибернетики", Гомель 1995., 158-160 стр. 158-159.
39. Симоненко В.П., Симоненко А.В., Элементы теории нового подхода к решению задачи поиска максимального паросочетания, Труды "Второй украинской конференции по автоматическому управлению", Львов, 1995г.стр. 93-94.
40. Симоненко В.П., Пространственно-временное планирование работ в параллельной вычислительной системе, Труды "Второй украинской конференции по автоматическому управлению", Львов, 1995 г., стр. 135 - 136.

Сімоненко В. П. Методи та засоби ефективного розв'язання задач статичного та дінамічного планування у розподілених обчислювальних системах.-рукопис.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальностю 05.13.13 — обчислювальні машини, системи і мережі, Національний технічний університет України "Київський політехнічний інститут", Київ, 1997.

Дисертація присвячена розробляннію практичних просторових планувальників систем розподіленого обробляння даних. Вирішена наукова проблема розробки високоефективних методів та засобів проектування адаптивних статичних та динамічних планувальників, що забезпечують в умовах просторово-часових обмежень організування обчислювального процесу адекватну стану обчислювального середовища та вхідному потоку робіт з метою підвищення ефективності використання ресурсів та скорочення додаткових витрат на організацію обчислень у гомогенних та гетерогенних розподілених системах обробляння інформації. Розроблені концепції, методи та засоби виконання процедур спрямованого пошуку розподілу робіт у СРОД з централізованим управлінням. Визначені процедури горизонтальної та вертикальної кластерізації, що є основою структурного аналізу вхідної інформації та реалізації засіба локалізації зони пошуку рішення та сканувального просторово-годинного алгоритму.

Ключові слова: розподілені обчислювальні системи, планування, динамічне планування, статичне планування.

Симоненко В.П. Методы и средства эффективного решения задач статического и динамического планирования в распределенных вычислительных системах. —Рукопись.

Диссертация на соискание ученой степени доктора технических наук по специальности 05.13.13 — вычислительные машины, системы и сети, Национальный технический университет Украины "Киевский политехнический институт", Киев, 1997.

Диссертация посвящена разработке практических пространственных планировщиков систем распределенной обработки данных. Решена научная проблема разработки высокоэффективных методов и средств проектирования адаптивных статических и динамических планировщиков, обеспечивающих в условиях пространственно-временных ограничений организацию вычислительного процесса адекватную состоянию вычислительной среды и входному потоку работ с целью повышения эффективности использования ресурсов и сокращение накладных расходов на организацию вычислений в

гомогенных и гетерогенных распределенных системах обработки информации. Разработаны концепции, методы и средства выполнения процедур направленного поиска распределения работ в СРОД с централизованным управлением. Определены процедуры горизонтальной и вертикальной кластеризации являющиеся основой структурного анализа исходной информации и реализующие метод локализации зоны поиска решения и сканирующего пространственно-временного алгоритма.

Ключевые слова: распределенные вычислительные системы, планирование, динамическое планирование, статическое планирование.

Simonenko V. P. Methods and means of the effective decision of problems of static and dynamic planning in distributed computing systems.-Manuscript.

Thesis for a doctor's degree by speciality 05.13.13 — computers, system and network, National technical university of Ukraine "Kiev politechnical institute", Kiev, 1997.

The dissertation is devoted to development of the practical space schedulers of systems of distributed data processing. A scientific problem of development of effective methods and means of designing of the adaptive static and dynamic schedulers, which provide organization of computing process in conditions of space-time restrictions, which provide adequate to a condition of computing environment and entrance flow of work with the purpose of increase of efficiency of use of resources and reduction of overhead charges on organization of calculations in homogeneous and heterogeneous distributed systems of processing of the information. The concepts, methods and means of fulfilment of procedures of directed search of distribution of work in DSDP with centralized management. Procedures horizontal and vertical clastering are developed, whice are the basis of the structural analysis of the initial information and realize method of localization of a zone of search of the decision and scan space-time algorithm.

Key words: distributed computing systems, scheduling, dynamic scheduling, static scheduling.

КОС, 1997 р.
Замов.-445 тир.-100

434083

Ав 38.748