

2. Запропонований новий спосіб за своєю простотою дозволяє швидко і багаторазово оцінювати потенційний виробіток на невеликому обмеженому майданчику при проектуванні чи побудові промислової ВЕС.

1. *Petersen E.L.* Estimation of Wind Resources / Erik Lundtang Petersen and Ib Troen // Wind Energy in Denmark. Research and technological development, 1990. – Copenhagen: Ministry of Energy. Danish Energy Agency. – P.64. – p.30-37.
2. *Рыхлов А.Б.* Ветроэнергетический потенциал на различных высотах приземного слоя атмосферы на юго-востоке европейской территории России / А.Б. Рыхлов // Известия Саратовского университета, 2014. – Т. 14. Сер. Науки о Земле, вып. 1. – С.30-37.
3. *Васько П.Ф.* Определение технических показателей эффективности использования ветроэнергетических агрегатов в Украине / П. Ф. Васько, А. А. Брыль, П. П. Пекур // Энергетика и электрификация, 1995. – № 2. – С.48-51.
4. *Кравчишин В. С.* Моделювання енергетичного потенціалу вітрової електричної станції / В. С. Кравчишин, М. О. Медиковський, М. О. Галушак // Вісник Національного університету «Львівська політехніка». Сер.: Інформаційні системи та мережі. – 2016. – № 854. – С.80-87.
5. *Справочник по специальным функциям* / под ред. М. Абрамовица и М. Стиган // М.: Наука, 1979. – 830 с.
6. Патент на корисну модель №135302 Укр., F03 1/00 GO1P 5/00. Спосіб оцінювання виробітку вітряної електричної турбіни / Подгуренко В.С., Терехов В.С., Гетманець О.М., Черепанов А.І., опубл. 25.06.2019р., Бюл. №12.
7. *Барановський Д.І.* Біометрія в програмному середовищі MS Excel: Навчальний посібник / Д.І. Барановський, О.М. Гетманець, А.М. Хохлов. – Х.: СПД Ф.О. Бровін О.В., 2017. – 90 с.
8. *Гершанік В.І.* Числові методи вирішення задач енергетики. Навчальний посібник / В.І. Гершанік – Миколаїв: УДМТУ, 2003. – 96 с.

<http://doi.org/10.5281/zenodo.3859675>

Поступила 9.09.2019р.

УДК 004.02

О.О. Сігарьов, м. Київ

ДОСЛІДЖЕННЯ ЗАЛЕЖНОСТІ ЕФЕКТИВНОСТІ ОБЧИСЛЕНЬ ВІД СТУПЕНЯ РОЗПАРАЛЕЛЮВАННЯ СИСТЕМ РІВНЯНЬ ПРИ ВИКОРИСТАННІ МЕХАНІЗМУ ПОВІДОМЛЕНЬ ДЛЯ ОБМІНУ

Abstract. The article presents the results of a study of the dependence of the efficiency of high-performance computing systems on the change in the degree of parallelization of the system of equations in the case of using a message transfer mechanism for exchange.

Вступ

Крупноблочне розпаралелювання є основним видом розпаралелювання для супер-ЕОМ. При цьому глибина розпаралелювання алгоритму задовольняє закладені в ньому можливості розпаралелювання лише на 1%. Це істотний недолік всіх суперкомп'ютерів. З цієї причини є завдання, яким бракує продуктивності сучасних суперкомп'ютерів для їх реалізації [1 – 3].

Залежність ефективності від зміни ступеня розпаралелювання систем рівнянь

В роботі [4] був проведений аналіз можливості отримання лінійного росту прискорення реалізації паралельного алгоритму з ідеальним паралелізмом в обчислювальних системах з масовим паралелізмом з ростом числа процесорів в них. Таким чином остаточною формула для прискорення набуває вигляду:

$$S(p) = \frac{n(2n-1)}{\frac{n(2n-1)}{p} + (2\alpha(\frac{n}{p} + \tau) + L)(p-1)} . \quad (1)$$

Ця формула враховує найважливіші чинники гальмування обчислень. Інші фактори, в тому числі і закон Амдала, не порівняні з ними. Функція (1) має максимум при $p = p_{\max}$, де

$$p_{\max} = \sqrt{\frac{n(2n-1) - 2\alpha n}{2\alpha\tau + L}} \quad (2)$$

Ефективність системи визначається за відомою формулою:

$$E(p) = \frac{S(p)}{p} . \quad (3)$$

Відомо, що СЛАР з щільними матрицями використовуються в ряді прикладних задач і їх порядок може досягати кілька тисяч рядків і стовпців. Наявність максимуму пред'являє жорсткі вимоги до розпаралелювання СЛАР. Звернемося до конкретних числових прикладів. Для більшої наочності будемо вимагати, щоб порядок системи задовольняв умові $n = 2^i$.

Проаналізуємо за допомогою формул (2), (3) як зростає прискорення і поводить ся такий найважливіший системний показник як ефективність зі збільшенням ступеня розпаралелювання СЛАР при використанні для обміну механізму повідомлень. Звернемося спочатку до СЛАР середніх розмірів (табл. 1, рис. 1) порядку $n = 512$.

α – відношення частоти процесора до частоти шини обміну (шини передачі даних)

L – латентність системи передачі даних (системи комутації)

Розглянемо три випадки:

1. $\alpha = 1, L = 0$ ($\phi 1$),
2. $\alpha = 4, L = 0$ ($\phi 2$),
3. $\alpha = 4, L = 2^{12}$ ($\phi 3$).

Де частота процесора – 2^{12} – це співвідношення між наносекундами та МГц – відносна одиниця.

Таблиця 1

| n=512 | | Прискорення | | | Ефективність | | |
|-------|-----|-------------|------|------|--------------|------|------|
| № | p | ф1 | ф2 | ф3 | ф1 | ф2 | ф3 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 2 | 2 | 1,98 | 1,95 | 1 | 0,99 | 0,98 |
| 3 | 4 | 4 | 3,91 | 3,58 | 0,99 | 0,98 | 0,89 |
| 4 | 8 | 7,9 | 7,56 | 5,35 | 0,99 | 0,95 | 0,67 |
| 5 | 16 | 15 | 14,1 | 5,32 | 0,97 | 0,88 | 0,33 |
| 6 | 32 | 30 | 24,6 | 3,53 | 0,93 | 0,77 | 0,11 |
| 7 | 64 | 54 | 36,8 | 1,92 | 0,84 | 0,58 | 0,03 |
| 8 | 128 | 86 | 42,9 | 0,98 | 0,67 | 0,33 | 0,01 |
| 9 | 256 | 103 | 36,7 | 0,49 | 0,4 | 0,14 | 0 |
| 10 | 512 | 85 | 24,4 | 0,25 | 0,17 | 0,05 | 0 |

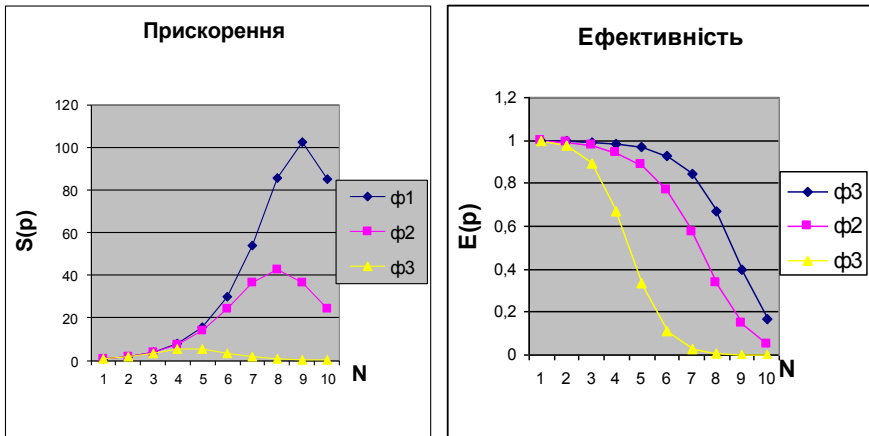


Рис. 1.

У всіх випадках розглядається лише два найбільш істотних фактора гальмування обчислень і падіння ефективності: α і L .

Значення факторів ϕ_1 вказують на те, що міжпроцесорний обмін здійснюється на частоті процесора i , крім того, відсутня латентність. Така ситуація може мати місце при обміні повідомленнями в багатоядерних структурах. Ненульовий параметр α з ϕ_2 , показує вплив на прискорення і ефективність відношення частоти процесора до частоти передачі даних, значення якого відображені на осі абсцис (номери N точок графіка). Шкала осі абсцис не є лінійною, тому що аргумент функції $p = 2^{(N-1)}$.

Крива ϕ_1 графіка прискорень СЛАР показує, що максимально можливе прискорення, може бути досягнуте в даній системі в ідеальному випадку, $S(256) = 103$ при ефективності $E(256) = 0,4$ – тобто 40%. Це означає, що в точці максимуму в даній системі 60% корисного часу буде витратитися на міжпроцесорний обмін. Це реальні втрати, що вносяться програмно-апаратним механізмом передачі повідомлень.

Крива ϕ_2 графіка прискорень демонструє, що зменшення частоти передачі даних в чотири рази ($\alpha = 4$) призводить до зменшення прискорення в точці максимуму більш ніж в два рази при ефективності 0,33.

Крива ϕ_3 , де значення факторів гальмування обчислень близькі до реальних показників, показує: максимальне прискорення $S(8) = 5,35$ і ефективність $E(8) = 0,67$. Слід звернути увагу, що при $p > 64$ прискорення приймає значення менші одиниці.

Таблиця 2

| n=16384 | | Прискорення | | | Ефективність | | |
|---------|-------|-------------|----------|----------|--------------|----------|----------|
| № | p | ϕ_1 | ϕ_2 | ϕ_3 | ϕ_1 | ϕ_2 | ϕ_3 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 |
| 4 | 8 | 8 | 7,99 | 7,98 | 1 | 1 | 1 |
| 5 | 16 | 16 | 15,9 | 15,9 | 1 | 1 | 0,99 |
| 6 | 32 | 31,9 | 31,8 | 31,5 | 1 | 0,99 | 0,99 |
| 7 | 64 | 63,8 | 63 | 61,2 | 1 | 0,98 | 0,96 |
| 8 | 128 | 127 | 124 | 111 | 0,99 | 0,97 | 0,87 |
| 9 | 256 | 252 | 240 | 164 | 0,98 | 0,94 | 0,64 |
| 10 | 512 | 495 | 449 | 163 | 0,97 | 0,88 | 0,32 |
| 11 | 1024 | 950 | 780 | 110 | 0,93 | 0,76 | 0,11 |
| 12 | 2048 | 1725 | 1171 | 60,7 | 0,84 | 0,57 | 0,03 |
| 13 | 4096 | 2731 | 1366 | 31,3 | 0,67 | 0,33 | 0,01 |
| 14 | 8192 | 3277 | 1170 | 15,8 | 0,4 | 0,14 | 0 |
| 15 | 16384 | 2731 | 780 | 7,92 | 0,17 | 0,05 | 0 |

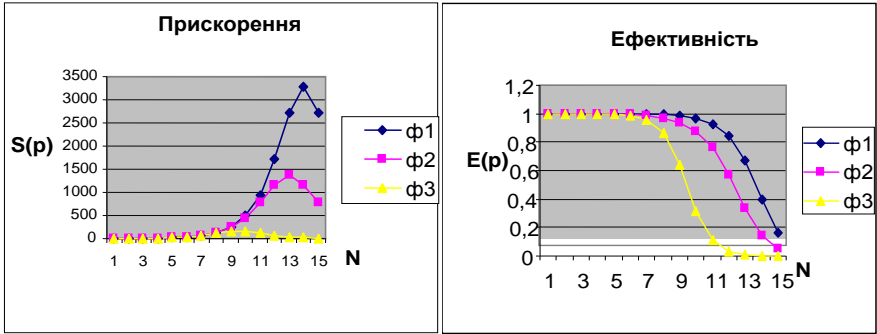


Рис. 2.

У реальній системі існують й інші фактори, що погіршують її продуктивність, тому для розглянутої СЛАР можна отримати прискорення не більше ніж у 4 рази. Очевидно, що досліджувана система не годиться для реалізації СЛАР таких порядків і її в даному випадку не можна віднести навіть до систем середнього класу. Тому, залишивши фактори гальмування (ϕ_1 , ϕ_2 , ϕ_3) незмінними, збільшимо тільки порядок СЛАР і число ОВ в системі до значення 16364, приймаючи його за найбільше, що використовується в прикладних задачах. Результати розрахунків для даного випадку наведені в табл. 2 і на відповідних графіках рис. 2.

Для аналізу введемо новий параметр $E_a = \frac{S_{\max}}{n}$ і назвемо його алгоритмічної ефективністю. Його призначення показати, якою мірою масивно-паралельна система може задовольнити, властиву алгоритму, максимальну глибину його розпаралелювання. Аналіз цих двох варіантів показує, що латентність системи комутації зовнішніх каналів обміну є найбільш потужним фактором (ϕ_3) гальмування. Поки підвищення числа ОВ дозволяє зберігати латентність системи незмінною, відносні параметри $E(p_{\max})$ і E_a не будуть залежати від порядку системи лінійних алгебраїчних рівнянь. В даному випадку для обох порядків $E_a = 0,01$. Це означає, що максимальне прискорення, яке може бути досягнуто на розглянутій багатопроцесорній системі (з урахуванням факторів ϕ_3) в 100 разів менше того прискорення, яке дозволяють отримати розглянуті СЛАР в ідеальному випадку. Однак зі зростанням числа ОВ латентність системи буде зростати. Можна стверджувати, що на фоні розглянутих факторів вплив закону Амдала мінімальний.

Висновки

Кожен з нині існуючих методів міжпроцесорного обміну описується деякою програмною функцією, реалізація якої супроводжується протоколами

обміну та управлінням з боку операційної системи. Ця функція є зовнішньою по відношенню до нерозпаралеленої програми, і час її реалізації – це чисті втрати продуктивності. Можлива компенсація цих втрат за рахунок більш раціонального кешування, не усуває природу самих втрат. Здійснення міжпроцесорного обміну за допомогою передачі повідомлень призводить до великих втрат продуктивності навіть в разі, коли всі елементи системи працюють на частоті процесорів.

1. *Lazou C.* Cray's Adaptive Supercomputing – A Paradigm Shift. March 24, 2006. <http://www.hpcwire.com/features>.
2. *Трахтенгерц Э.А.* Программное обеспечение параллельных процессов. М.: Наука, 1987 г. – 272 с.
3. *Booth, Nick.* NEC claims 10-Petaflop supercomputing breakthrough. March, 2008. <http://www.theinquirer.net/gb/inquirer/news>
4. *Сігарьов О.О., Душеба В.В.* Динаміка прискорення реалізації паралельного алгоритму в масштабованих системах з масовим паралелізмом // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ ім. Г.С. Пухова НАН України. – Вип. 79. – К.: 2017. – С.9-16.

<http://doi.org/10.5281/zenodo.3859677>

Поступила 3.10.2019р.

УДК 009.4

Б.М. Гавриш¹
Б.В. Дурняк¹
О.В. Тимченко^{1, 2}

АНАЛІЗ СУЧАСНИХ ІНСТРУМЕНТІВ DATA MINING

Abstract. The article analyzes modern Data Mining tools. A detailed description of each Data Mining tool is given. The principles of operation of these tools are considered, the main criteria for comparison are presented. The pros and cons of each Data Mining tool are listed. Conclusions are made about the effectiveness of the DMST tool for analytical projects.

Keywords: DMST, Intelligent Processing, Data Mining, Mathematical Packages, Business Analytics, Data Analysis, Mat Package.

¹ Українська академія друкарства, Львів

² University of Warmia and Mazury Olsztyn, Poland

© Б.М. Гавриш, Б.В. Дурняк, О.В. Тимченко