

***В. І. Павлюк**, студ. 4 курсу спеціальності  
«Міжнародна інформація» Східноєвропейського  
національного університету імені Лесі Українки*

*Науковий керівник: **С. В. Федонюк**, к.геог.н., доц. кафедри  
міжнародної інформації СНУ ім. Лесі Українки  
(Україна, м. Луцьк)*

**Розробка проекту на основі технології розподілених  
обчислень (безкоштовні додатки)**

*Роботу виконано на кафедрі міжнародної  
інформації СНУ ім. Лесі Українки*

***Pavliuk V.I. Working of the project based on the technology of distributed computing (Free Application).** Distributed computing - a way of resolving labor-intensive computing tasks using multiple computers, often combined in a parallel computing system. Performance of sequential computation in distributed systems has meaning within many tasks at once, such as distributed control systems. The feature of distributed multiprocessor systems, as opposed to local supercomputers, you can build unlimited productivity by scaling.*

Можна сказати, що родоначальниками машин для обчислень були системи, які використовували спеціальні процесори і складні комунікаційні рішення, які володіли великим енергоспоживанням і вартістю. Ця вартість обумовлювалась унікальністю встановлюваного обладнання, в першу чергу це процесорів[4].

За деякими припущеннями, початок розподіленим обчисленням було покладено приблизно два сторіччя тому. Звичайно, що ні про які комп'ютери тоді не могло бути і мови, але це був час, коли перед вченими ставили завдання великих за змістом обчислень. І зрозуміло, що ці задачі потрібно було вирішувати, використовуючи принцип розподілених обчислень. В кінці 18-го століття уряд Франції вирішив істотно поліпшити логарифмічні і тригонометричні таблиці. Робота була пов'язана з величезними на ті часи обсягами розрахунків, а тому її доручили професіоналові, барону Гаспару де Проні, який був головним обчислювачем при французькому уряді в 1790-1800 рр. У результаті з'явилася знаменита "обчислювальна мануфактура" барона де Проні[3].

Барон сміливо взяв ідею поділу праці і приніс її принципи в обчислювальний процес. Учасники проекту були розділені на три рівні. Низький рівень в системі був зайнятий звичайними людьми - обчислювачами, від яких було необхідно лише проводити точну арифметику. На другому рівні були освічені технологі ("сервери" проекту), які організовували рутинний процес, розподіляючи завдання і обробляючи отримані обчислювачами дані. Високий рівень займають чільні французькі математики ("адміністратори - програмісти" проекту), в тому числі Адрієн Лежандр і Лазар Карно. Вони готували математичне забезпечення - по суті писали "програми". В результаті, барону де Проні вдалося організувати процес таким чином, щоб зменшити дуже складні проблеми з набором рутинних операцій, завдяки суворій системі контролю та добре налагодженій системі розподілу роботи між обчислювачем [1].

Ідеї де Проні підштовхнули Чарльза Беббіджа, щоб створити перший в історії комп'ютер (обчислювальну машину), а засновані на прототипі барона "обчислювальні мануфактури" використовувалися в науково-дослідних проектах до середини 20-го століття, з їх допомогою, зокрема, розраховувалися перші ядерні бомби в США, Великобританії і Радянському Союзі.

Вперше завдання сумісного використання вільних обчислювальних ресурсів комп'ютерів була оригінальним способом вирішення на початку 1970-х, вченими з дослідницького центру PARC (Xerox, Palo Alto), які розробляли програмне забезпечення для перших комп'ютерних мереж. У 1973 році Джон Шох і Жон Хапп написали просту програму, яка запускалася вночі в локальну мережу PARC, розповзалася по працюючих комп'ютерах і змушувала їх виконувати обчислення.

Наступний стрибок у технічному рішенні спільної роботи багатьох комп'ютерів над єдиними складними задачами відбувся наприкінці 1980 - х років, коли були персональні комп'ютери та електронна пошта. У 1988 році Арьєн Ленстра і Марк Менес, який працював на компанії DEC (Palo Alto), написав програму для факторизації (розкладання на множники) довгих чисел. Щоб швидко завершити поставлене завдання програма могла запускатися на декількох не пов'язаних одна з одною машин, кожна з яких обробляє свій невеликий фрагмент числа. Нові блоки завдань відправляються на комп'ютери учасників (а підраховували результати були прийняті) через центральний серверний проект по звичайній електронній пошті. Для успішного розкладання на множники числа завдовжки в сто символів цьому співтовариству потрібно було два роки і сотні персональних комп'ютерів[2].

Тепер треба вказати на різницю між двома областями розподілених обчислень. Поняття Distributed Computing і Grid Computing часто змішують, але це неправильно.

Завдання, які реалізуються за допомогою Grid Computing, вимагають активної взаємодії між комп'ютерами і синхронізації, що реалізуються з використанням спеціальних бібліотек MPI, PVM, BSPlib та інші. Ці бібліотеки зазвичай можуть бути використані тільки в кластерних системах, так як вони потребують високої швидкості мережевих взаємодій і, отже, великий обсяг безперервного мережевого трафіку. Той факт, що мережеві взаємодії здійснюються між усіма машинами, пов'язаних з розрахунками, також означає заборону на використання міжмережевих екранів або інших засобів обмеження мережевих потоків. Він також, як правило, означає, що абсолютно всі машини повинні бути доступні з самого початку до кінця вирішити проблему. Таким чином, реалізована на таких принципах обчислювальна мережа можлива найчастіше в межах деякої замкнутої організації. Через ці обмеження, завдання, які вимагають використання цих бібліотек взагалі не підходять для великомасштабних розподілених проектів.

Обчислювальні завдання, які не потребують ніякої координації або синхронізації між машинами, не вимагаючи додаткового контролю в процесі роботи, як правило, вирішується за допомогою технології Distributed Computing. Цей тип завдання дуже підходить для великих інтернет – проектів розподілених обчислень, що реалізуються за допомогою зв'язку "сервер - комп'ютер учасника", тому що фрагменти завдання можуть бути виконані в будь-якому порядку і можуть бути перерозподілені на інші машини, якщо деякі з результатів не вдалося отримати з першого разу. Звідси, до речі, виникає

необхідність для розрахунку надмірності, коли той же блок завдання надсилається кільком учасниками. Це необхідно, щоб гарантувати отримання результатів обчислення окремого блоку, це не завжди можливо, враховуючи різноманітність учасників і встановленого на їх комп'ютерах програмного забезпечення [3].

Звичайно, технологія Distributed Computing накладає жорсткі обмеження на коло задач, які вирішуються за її допомогою, але також очевидні і переваги. Деякі науково-дослідні організації, особливо некомерційні, досить часто недостатньо фінансування для покупки або оренди спеціалізованих можливостей тих же суперкомп'ютерів або кластерних обчислювальних мереж. Крім того, суперкомп'ютери і кластери вимагають постійного обслуговування, модернізації та навчання відповідального персоналу. Розподілені обчислення за допомогою армії користувачів Internet позбавлені цих недоліків і часто підключені до проекту обчислювальної потужності, співрозмірні і навіть перевершують потужність сучасних суперкомп'ютерів.

13 травня 1999 р. Девід Гиді і Крейг Кеснов запустили проект на пошук позаземних сигналів цивілізації SETI @ Home (SETI - пошук позаземного розуму), який став наймасовішим проектом розподілених обчислень всіх часів. Величезній популярності (більше 5 млн. учасників) сприяло те, що проект був вперше перекладений на рейки Distributed Computing наукове завдання, що інтригує і яка далеко від нудної факторизації або злому чергового ключа. Приголомшливий успіх проекту (з точки зору повернутих обчислювальних потужностей) закликав учених активно використовувати невикористаний потенціал звичайних користувачів комп'ютера всесвітньої павутини в багатьох областях досліджень.

1 жовтня 2000 р. під керівництвом Віждая Панде, засновника наукових установ Pande Group при Стенфордському університеті, почав проект Folding @ Home. З його чітких практичних цілей, хорошої науковій основі і реальним результатам проект отримав серйозну підтримку (більш ніж півмільйона членів за час існування) і значні обчислювальні ресурси (більше 200 тис. процесорів, які постійно або тимчасово підключені зараз до проекту) .

BOINC - програмний комплекс для швидкої організації розподілених обчислень. Складається з серверної і клієнтської частин. Спочатку розроблявся для найбільшого проекту добровольчих обчислень -SETI@home, але згодом розробники з Каліфорнійського університету в Берклі зробили платформу доступною для сторонніх проектів. На сьогоднішній день BOINC є універсальною платформою для проектів у галузі математики, молекулярної біології, медицини, астрофізики та кліматології. BOINC дає дослідникам можливість задіяти величезні обчислювальні потужності персональних комп'ютерів з усього світу[3].

Платформа працює на різних операційних системах, включаючи Microsoft Windows і варіанти юніксоподібних GNU/Linux, FreeBSD, NetBSD, OpenBSD, Solaris і Mac OS X. BOINC поширюється під ліцензією GNU Lesser General Public License, як вільне програмне забезпечення з відкритим вихідним кодом [2].

Для користувачів поняття BOINC частіше використовується в контексті поняття BOINC-клієнт - універсальний клієнт для роботи з різними (BOINC-сумісними) проектами розподілених обчислень. BOINC-клієнт дозволяє

брати участь одночасно в декількох проектах з допомогою однієї загальної програми управління (boinc або boinc.exe).

*Джерела:*

1. Аксак Н.Г. Паралельні та розподілені обчислення: підруч./ Н.Г.Аксак, О.Г. Руденко, А.М.Гуржій. – Х.: Компанія СМІТ, 2009. – 480с.

2. Кузьменко Б.В., Чайковська О.А. Технологія розподілених систем.

3. Ю. Шийка, Р. Шувар. Використання розподілених обчислень.

4. А.Е. Дорошенко, О.В. Алистратов, Ю.М. Тырчак, А.П. Розенблат, К.А. Рухлис Паралельні обчислення та розподілені системи.