

## ВИКОРИСТАННЯ РОЗПОДІЛЕНИХ ОБЧИСЛЮВАЛЬНИХ ПОТУЖНОСТЕЙ ДЛЯ АНАЛІЗУ ДАНИХ АСТРОНОМІЧНИХ СПОСТЕРЕЖЕНЬ

**Віктор Сисуєв** – студент 3 курсу СВО бакалавра НПУ ім. М.П. Драгоманова

*Проблема додаткового залучення обчислювальних потужностей для аналізу даних астрономічних спостережень є актуальною на сьогоднішній день. Використання розподільних обчислень є перспективною для обробки даних. Ця технологія вже успішно використовувалась у проекті SETI. Запропонована оптимізація використання розподільних обчислень використовуючи технологію Блокчейн.*

**Ключові слова:** обробка даних астрономічних спостережень, розподільні обчислення, технологія БлокЧейн, Blockchain, SETI

Дана тема актуальна для вивчення на сьогоднішній день, оскільки кількість даних астрономічних спостережень отриманих від різних типів спостережної апаратури на сьогодні перевищує 150 терафлоп. Очікуваний час на обробку вже отриманої інформації 5-7 років. Вважаю доцільним залучення додаткових обчислювальних потужностей, в тому числі використовуючи розподільні.

Прикладами вірності такої точки зору є стрімкий ріст використання розподільних обчислювань у різних сферах науки і техніки.

### **Проекти, які використовують розподільні обчислення:**

*Астрономія та астрофізика*

SETI@Home — проект з обробки сигналів радіотелескопу, для пошуку радіосигналів позаземних цивілізацій;

Einstein@Home — проект з перевірки гіпотези Ейнштейна про гравітаційні хвилі за допомогою аналізу гравітаційних полів пульсарів чи нейтронних зірок;

MilkyWay@home — створення трьохмірної моделі галактики Чумацький шлях;

Cosmology@home — пошук моделі, яка найкращим чином описує наш Всесвіт, а також визначення діапазону моделей, які узгоджуються із сучасними астрономічними та фізичними даними;

Asteroids@home — метою проекту є визначення форми і характеру обертання значної частини астероїдів.

**Проект SETI.** Особливу увагу хочу приділити проекту SETI@Home — проект з обробки сигналів радіотелескопу, для пошуку радіосигналів позаземних цивілізацій. Як один з підходів пошуку позаземних цивілізацій SETI Searches використовує радіотелескопи для пошуку вузькосмугових радіосигналів з космосу. Ймовірно, позаземна цивілізація буде використовувати радіозв'язок (земні радіостанції можна зловити з найближчих зоряних систем на високочутливий радіоприймач). Якщо в радіосигналі будуть періодично-повторювані елементи, їх буде нескладно виявити, розрахувавши для запису з радіоприймача перетворення Фур'є. Ці повторювані сигнали приблизно повинні мати штучну природу і, відповідно, їх виявлення побічно підтвердить присутність позаземної технології. Сигнали, отримані радіотелескопом, переважно складаються з шуму, виробленого небесними об'єктами, радіоелектронікою, супутниками, телевізійними вежами і радарми. Сучасні проекти з пошуку позаземних цивілізацій, такі як SETI, в радіодіапазоні використовують цифрові технології для аналізу даних. Для Radio SETI потрібна неймовірна обчислювальна потужність, тому що розрахунок перетворення Фур'є — вкрай ресурсомістка задача, і в цьому випадку вона помножена на величезну кількість інформації, що надходить з радіотелескопів.

**Історія проекту.** У попередніх проектах SETI Radio Searches використовувалися спеціалізовані суперкомп'ютери, встановлені на радіотелескопах, що аналізували величезний обсяг інформації. У 1994 році Девід Геді, працюючи в програмі SERENDIP Каліфорнійського Університету в Берклі, запропонував використовувати

віртуальний суперкомп'ютер, що складається з великого числа комп'ютерів, що мають доступ до Інтернету, та організував проект SETI@home для перевірки цієї ідеї. Науковий план, який розробили Девід Геді і Крейг Каснофф з Сіетлу, був представлений на п'ятій міжнародній конференції з біоастрономії у липні 1996 року.

Проект SETI@home стартував 17 травня 1999 року. Ця версія, іменована надалі SETI@HomeClassic, проіснувала до 15 грудня 2005 року. Далі проект продовжується тільки з використанням платформи BOINC. З 3 травня 2006 року використовується клієнтське програмне забезпечення SETI@homeEnhanced.

**Фінансування проекту.** Фінансування проекту здійснюється в основному Planetary Society [en] — некомерційною організацією, діяльність якої присвячена дослідженню Сонячної системи й пошуку позаземного розуму. Planetary Society є основним спонсором SETI@home. Також великий вклад складають пожертвування від учасників проекту й безкоштовна передача обладнання від спонсорів. Крім того, є фінансові надходження від продажу товарів з атрибутикою проекту.

**Програмне забезпечення.** Клієнтське програмне забезпечення побудоване на базі платформи з відкритим програмним кодом (GNU General Public License), де кожен охочий учасник проекту може внести свій вклад не тільки в розрахунки, але і в розробку й тестування програмного забезпечення. Тому клієнтське забезпечення доступне для більшості популярних операційних систем і типів центральних процесорів.

**Як працює проект.** Проект полягає в обробці даних радіотелескопу обсерваторії Аресібона предмет пошуку сигналів, які можна інтерпретувати як штучні. Інформація, отримана з опромінювача радіотелескопу, записується з високою щільністю на магнітну стрічку (заповнюючи приблизно одну 35-гігабайтну DLT плівку в день). При обробці дані з кожної стрічки розбиваються на 33000 блоків по 1049600 байт, що становить 1,7 с часу запису з телескопу. Потім 48 блоків конвертуються в 256 завдань на розрахунок, які розсилаються не менше ніж на 1024 комп'ютери учасників проекту. Після обробки результати передаються комп'ютером учасника проекту в Space Sciences Laboratory (SSL) Каліфорнійського університету, Берклі (США) за допомогою програмного забезпечення BOINC.

Кожен користувач персонального комп'ютера, що має доступ до Інтернету, може підключитися до проекту (такий підхід дає безпрецедентну обчислювальну потужність, зумовлену великою кількістю комп'ютерів, що беруть участь в обробці даних).

**Розвиток проекту.** На 25 березня 2012 року, проект SETI@home є найбільш популярним на платформі BOINC — загальне число учасників проекту становить понад 1,2 млн. За обсягом обчислень у день, за станом на 25 березня 2012 року, проект займав п'яту позицію з результатом 1,6 петафлопс, поступаючись проектам Folding@home, PrimeGrid, DistRTGen і MilkyWay@home. Результати використовуються також і для дослідження інших астрономічних об'єктів. Подальше продовження й доповнення до проекту SETI@Home — проект AstroPulse (Beta) (астрономічні дослідження).

Для AstroPulse (Beta) існують клієнти для GNU/Linux (в тому числі і для 64-розрядних версій) і Microsoft Windows. 27 січня 2009 року, було оголошено про створення нового відкритого проекту — setiQuest. Як очікується, в його основу ляжуть вихідні коди SETI@Home, які повинні бути передані спільноті під відкритою ліцензією в другому кварталі 2010 року.

На сьогодні проект, що працював протягом 21 років (проект пошуку позаземного життя SETI @ Home) оголосив про своє закриття в кінці березня. Для своєї роботи він використовував комп'ютери звичайних користувачів. Це зробило його найвідомішим із серії проектів Search for Extra-Terrestrial Intelligence пошуку розумного життя на інших планетах.

Творці проекту називають причиною його заморожки критичну малу ефективність. За час його роботи було опрацьовано величезну кількість інформації, проте це не

дозволило прийти до якого-небудь фінального висновку. Вчені говорять, що необхідно зосередитися на аналізі вже отриманих результатів. Університет Берклі продовжить використовувати обчислювальні потужності користувачів в рамках своєї платформи BOINC.

### **Використання розподільних обчислень в астрономії**

Спостережна астрономія – галузь астрономії, пов'язана з отриманням спостережних даних про небесні об'єкти із застосуванням телескопів та інших астрономічних приладів.

Як наука астрономія практично позбавлена можливості проведення експериментів з об'єктами Всесвіту, що дещо компенсується можливістю спостерігати й досліджувати величезну кількість астрономічних явищ. Подібні спостереження дозволяють, наприклад, простежувати деякі закономірності властивостей, що виявляються об'єктами. Результати вивчення близьких об'єктів, які виявляють певні властивості (наприклад, змінні зорі) можна поширити на віддаленіші об'єкти з подібними властивостями: так, залежно від періоду пульсації цефеїд можна оцінювати їх світність та відстані до інших галактик.

Галілео Галілей застосовував телескоп для спостереження небесних об'єктів і записував результати спостережень. Відтоді спостережна астрономія істотно розвинулася, удосконалювалася техніка створення телескопів.

**Розподілені обчислення** (розподілена обробка даних) — спосіб розв'язання трудомістких обчислювальних завдань з використанням двох і більше комп'ютерів, об'єднаних в мережу.

Розподілені обчислення є окремим випадком паралельних обчислень, тобто одночасного розв'язання різних частин одного обчислювального завдання декількома процесорами одного або кількох комп'ютерів. Тому необхідно, щоб завдання, що розв'язується було сегментоване — розділене на підзадачі, що можуть обчислюватися паралельно. При цьому для розподілених обчислень доводиться також враховувати можливу відмінність в обчислювальних ресурсах, які будуть доступні для розрахунку різних підзадач. Проте, не кожне завдання можна «розпаралелити» і прискорити його розв'язання за допомогою розподілених обчислень.

*Щоб знизити витрати часу, пов'язані з латентністю, потрібно:*

- розробляти алгоритми, що вимагають менше пересилань даних, так як це є мірою складності системи, а також групувати запити і відповіді;
- використовувати інформацію, розташовану «близько» в гіпермережі;
- кешувати, запитувати заздалегідь і дублювати інформацію (при цьому не варто забувати, що дані мають властивість застарівати);
- переміщати дані на ЕОМ, де виконуються обчислення;
- виконувати обчислення там, де зберігаються дані.

Це вимагає вирішення виникаючих питань, пов'язаних з безпекою та використанням приватних ресурсів та сервісів.

Повніший перелік практично всіх існуючих і завершених проектів розподілених обчислень в інтернеті можна переглянути на сайті DistributedComputing. Учасники розподілених обчислень в Україні працюють як самостійно, так і в складі команд, найчисельнішою та найпотужнішою з яких є — DistributedComputingTeamUkraine. Статус команди — Національна, офіційний сайт.

**Деякі проблеми, які можливо вирішити за допомогою розподільних обчислень:** спостереження періодичних комет, спостереження зміни яскравості зорі таке інше. Для отримання найкращих результатів я пропоную використовувати вже існуючі технології блокчейну. **Блокчейн**, тобто ланцюжок блоків транзакцій (англ. Blockchain, Blockchain від block — блок, chain — ланцюг) — розподілена база даних, що зберігає впорядкований ланцюжок записів (так званих блоків), що постійно довшасяє. Дані захищено від підробки та спотворення. Кожен блок містить часову позначку, геш попереднього блока та дані транзакцій, подані як геш-дерево. Таку розподілену базу даних закладено в основу крипто

валюти Bitcoin (вона була описана 2008 і реалізована 2009 року), де слугує бухгалтерською книгою для всіх операцій.

**Переваги використання:** можливість залучати для обробки інформації практично будь які сучасні електронні пристрої (комп'ютери, мобільні телефони, носиму електроніку, телевізори, таке інше). Можливість отримувати данні з фактом підтвердження. Можливість групувати отримані данні, робити певні розрахунки та твердження. Для прикладу: для знаходження періодичних комет треба лише порівняти фотографії зоряного неба зроблені через певний час. Отримані данні будуть достеменні і їх можна використовувати для подальших астрономічних спостережень.

Над створенням нових блоків одночасно працює чимало «майнерів». Новостворений блок, що відповідає певним умовам, негайно надсилається решті членів мережі і має стати наступною ланкою ланцюжка. Постійно трапляється таке, що з різних частин мережі (від різних учасників) надходять блоки, що попереднім називають той самий блок, тобто відбувається галуження. Навмисне чи ненароком можна обмежити поширення новостворених блоків (наприклад, одне з галужень ланцюжка може деякий час розвиватися в межах локальної мережі). Тоді одночасно відбувається створення кількох гілок одного ланцюжка, що суперечать одна одній.

Коли поширення блоків поновлюється, майнери розв'язують суперечність, обираючи найдовшу гілку з найбільшим рівнем складності за єдину «достовірну». За однакової складності і довжини перевага віддається гілці, кінцевий блок якої з'явився раніше.

Розподілена база даних Blockchain — це ланцюжок блоків, що постійно зростає, зберігаючи всю історію транзакцій. Копія бази даних або її частини одночасно зберігаються на безлічі комп'ютерів та синхронізуються відповідно до формальних правил побудови ланцюжка блоків. Дані блоків не шифровані і доступні у відкритому вигляді, проте захищені від змін криптографічно через геш-ланцюжок

Зазвичай умисна зміна інформації в будь-якій копії бази або навіть в багатьох копіях не буде визнана істинною, бо не відповідатиме правилам. Деякі зміни може бути прийнято, якщо їх внести в усі копії бази (наприклад, видалення кількох останніх блоків через помилку в їхньому формуванні).

Отже, проаналізував сучасний стан астрономічних спостережень було виявлено недостатню кількість обчислювальних потужностей, що використовуються для обробки даних астрономічних спостережень та зроблено огляд використання подібної технології у інших областях науки та техніки. Пропоную залучити більше потужностей, використовуючи розподільні обчислення. Для отримання достеменних даних пропоную використовувати технологію блокчейн. Вважаю, що це дасть змогу по-перше прискорити обробку інформації отриманої за допомогою різних типів спостережної апаратури, по-друге скоротить кошти на обробку цієї інформації, по-третє отримати інформацію, яка достеменно підтверджена.

## THE USE OF DISTRIBUTED COMPUTING POWER FOR THE ANALYSIS OF ASTRONOMICAL OBSERVATIONS

**Viktor Sysuev** – 3rd year student of higher education bachelor's degree NPDU

*The problem of additional attraction of computing power for the analysis of astronomical observation data is relevant today. The use of distribution computing is promising for data processing. This technology has already been successfully used in the SETI project. The proposed optimization of the use of distribution computing using Blockchain technology.*

**Key words:** processing of astronomical observation data, distribution computing, BlockChain technology, Blockchain, SETI