

Дослідження декомпозиції задач в хмарних обчисленнях

Виконав
студент групи ДА-61м
Крещук Максим

Декомпозиція

- Тривіальна

Різні потоки виконуються незалежно один від одного

- Функціональна

Існує залежність деяких потоків від результатів роботи попередніх потоків

- Декомпозиція даних

Один код потоку, різні набори даних

Хмарні сервіси

- Amazon EC2



- Google App Engine



- Microsoft Windows Azure



Windows Azure™

Порівняння хмарних сервісів

	Amazon EC2	Google App Engine	Microsoft Windows Azure
Категорія	IaaS	PaaS	PaaS
Підтримка паралельних розрахунків	Об'єднання Spot Instances		Windows Azure HPC Scheduler
Ціна базового плану	0,0008\$/ годину	0\$/годину	0,12\$/годину
Опис базового плану	1.7ГБ RAM, 160ГБ сховища 1 EC2 Обчислювальний юніт.	500 МБ дискового простору , 28 процесорних годин на добу безкоштовно	1.6 ГГц CPU, 1.75 ГБ RAM, 225 ГБ сховища, низька швидкість I/O операції

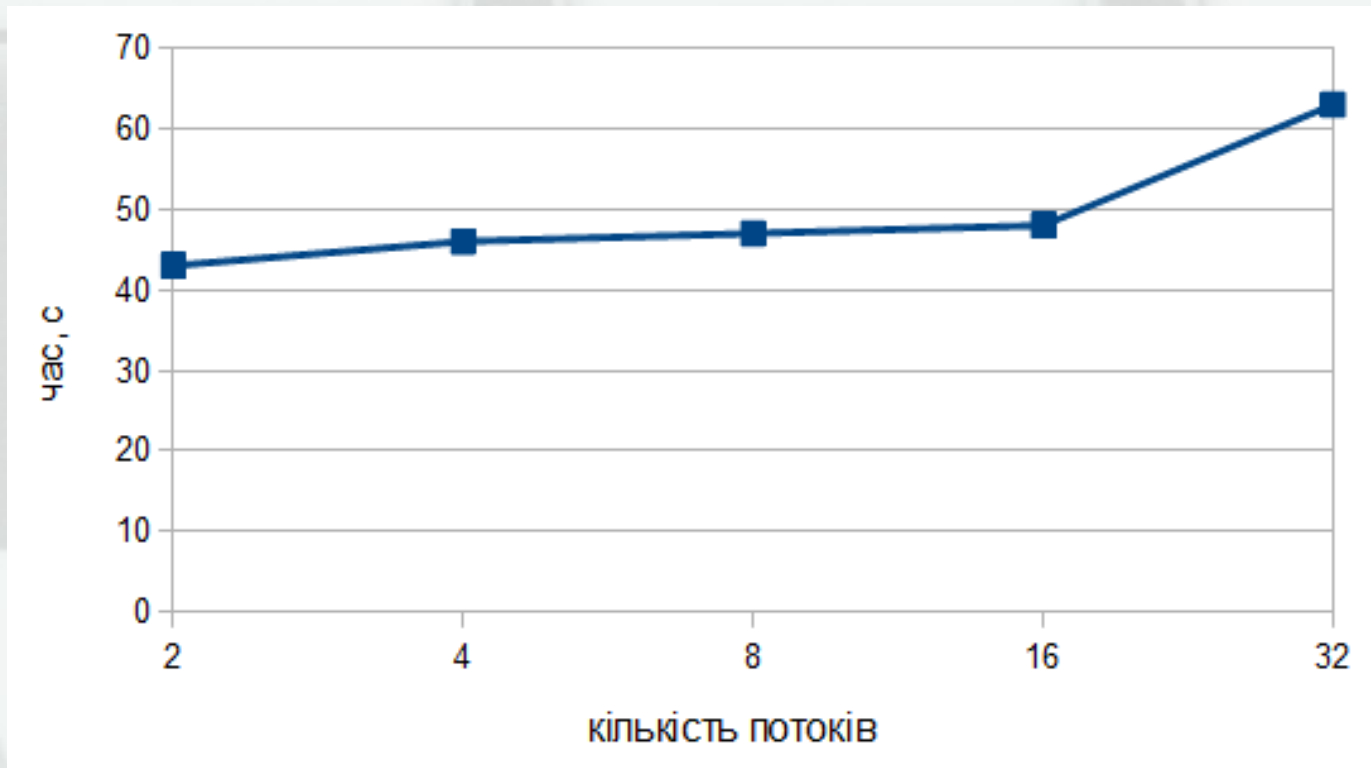
MapReduce

Складається з двох кроків:

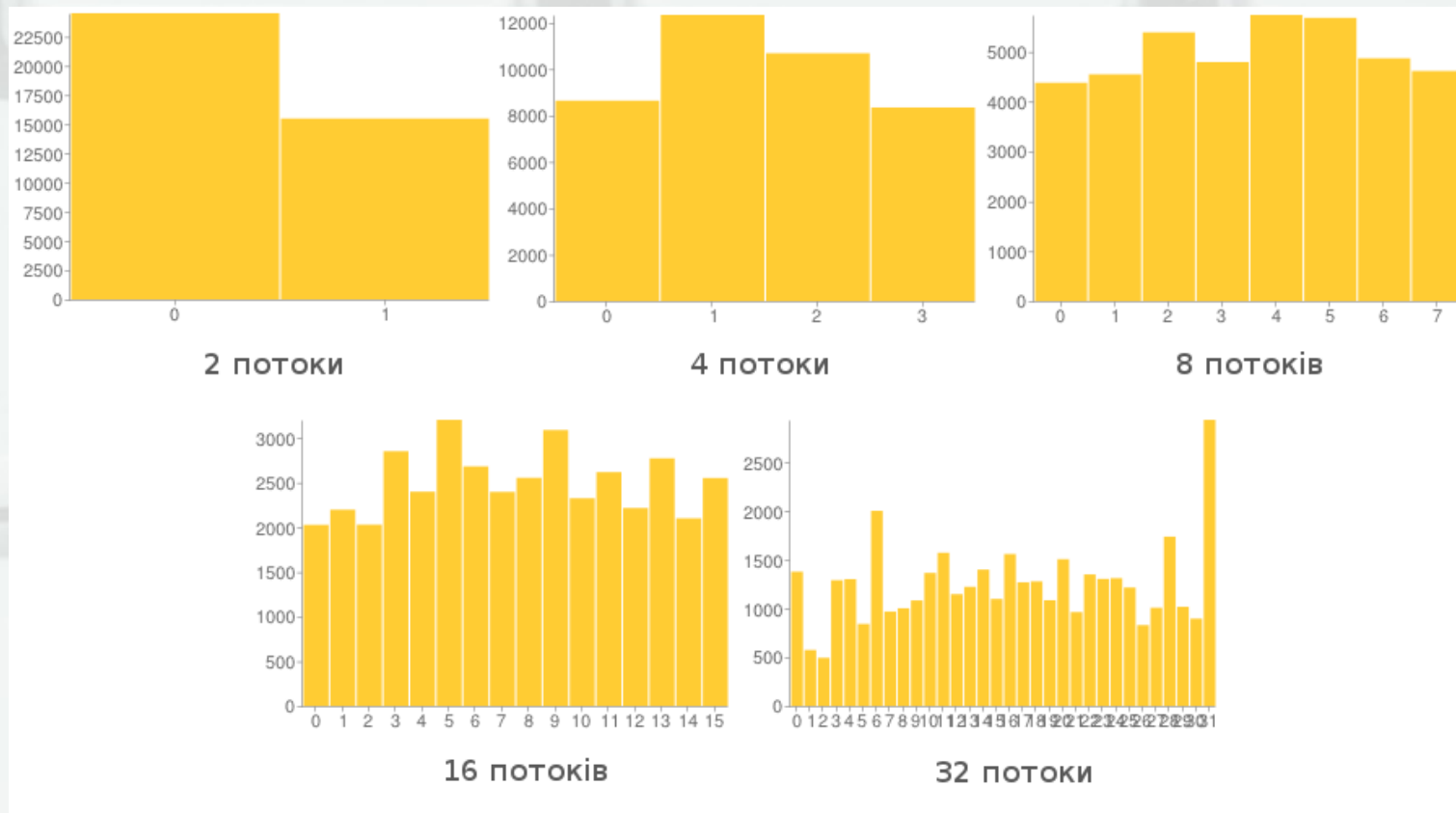
- Map(попередня обробка вхідних даних)
- Reduce(згортка попередньо оброблених даних)

Для GAE створена бібліотека AppEngine-MapReduce

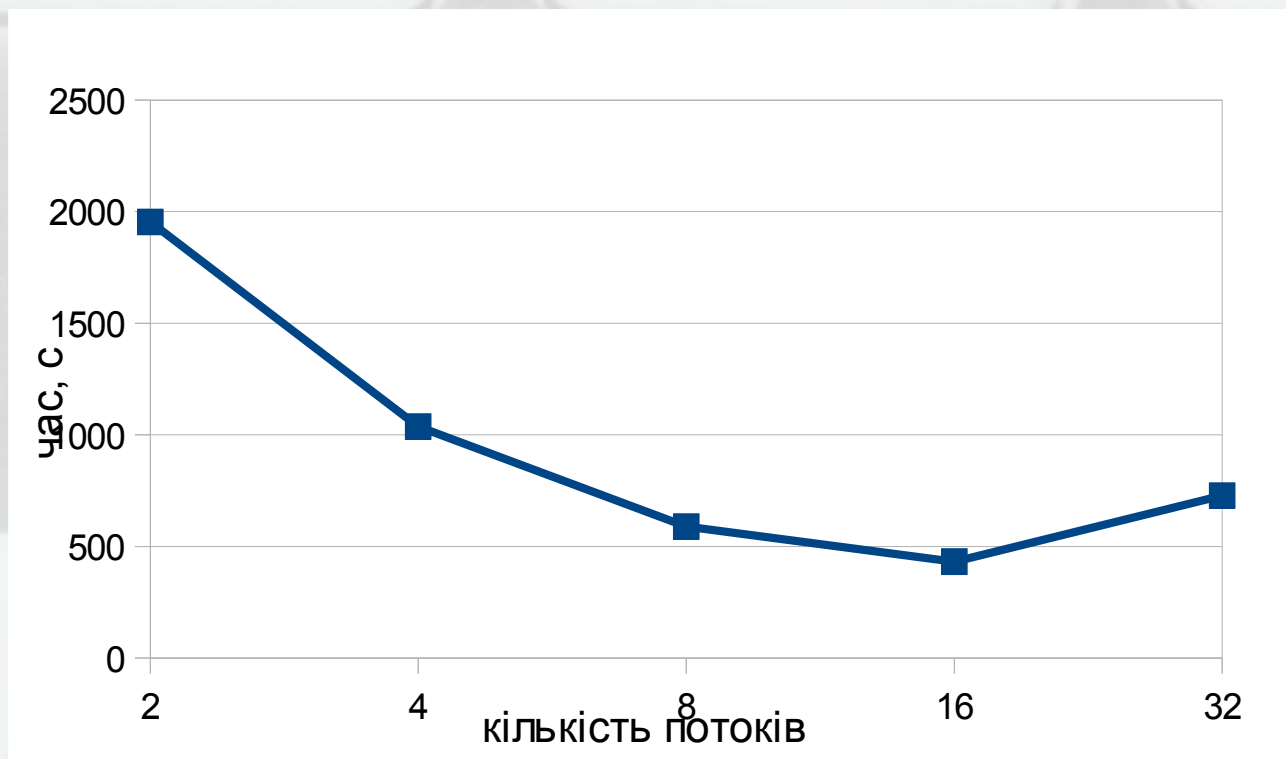
Результати роботи MapReduce на Google App Engine(1/4)



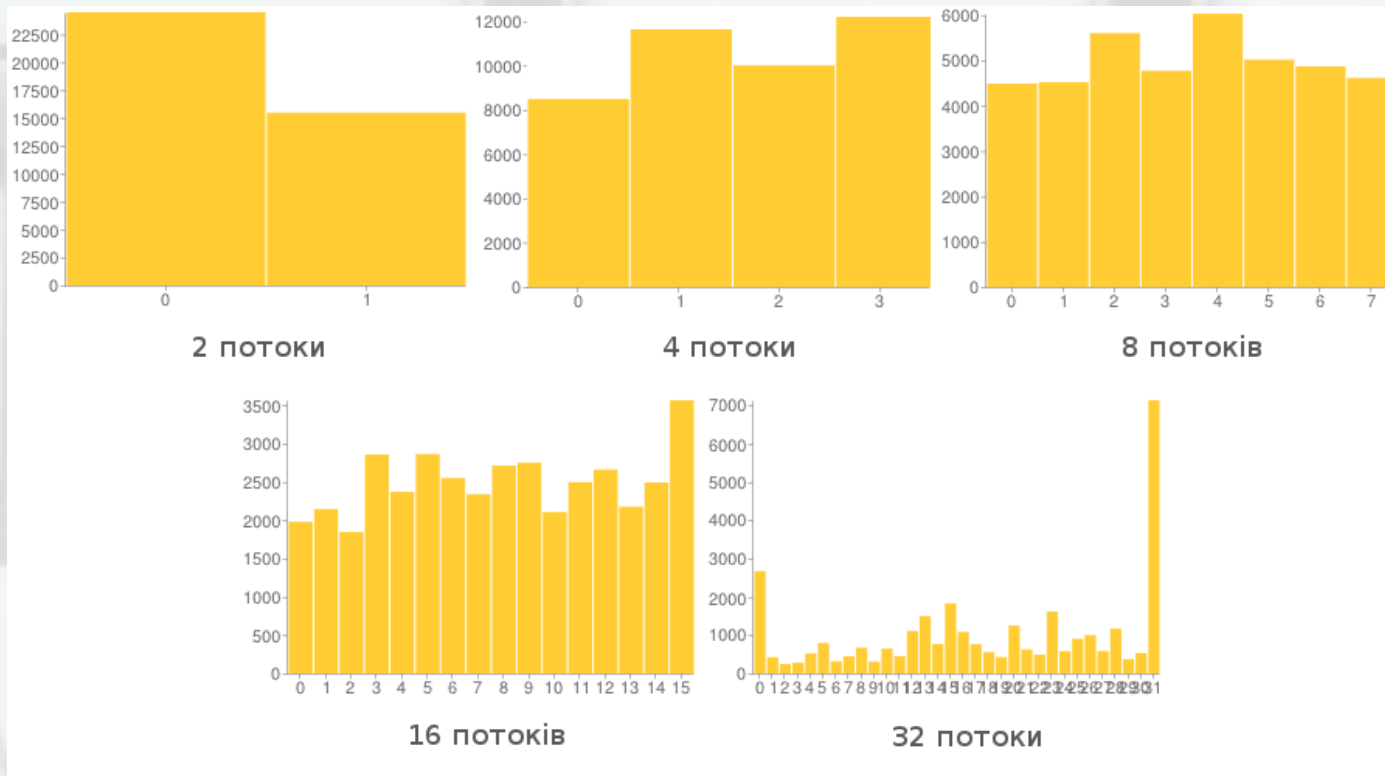
Результати роботи MapReduce на Google App Engine(2/4)



Результати роботи MapReduce на Google App Engine(3/4)



Результати роботи MapReduce на Google App Engine(4/4)



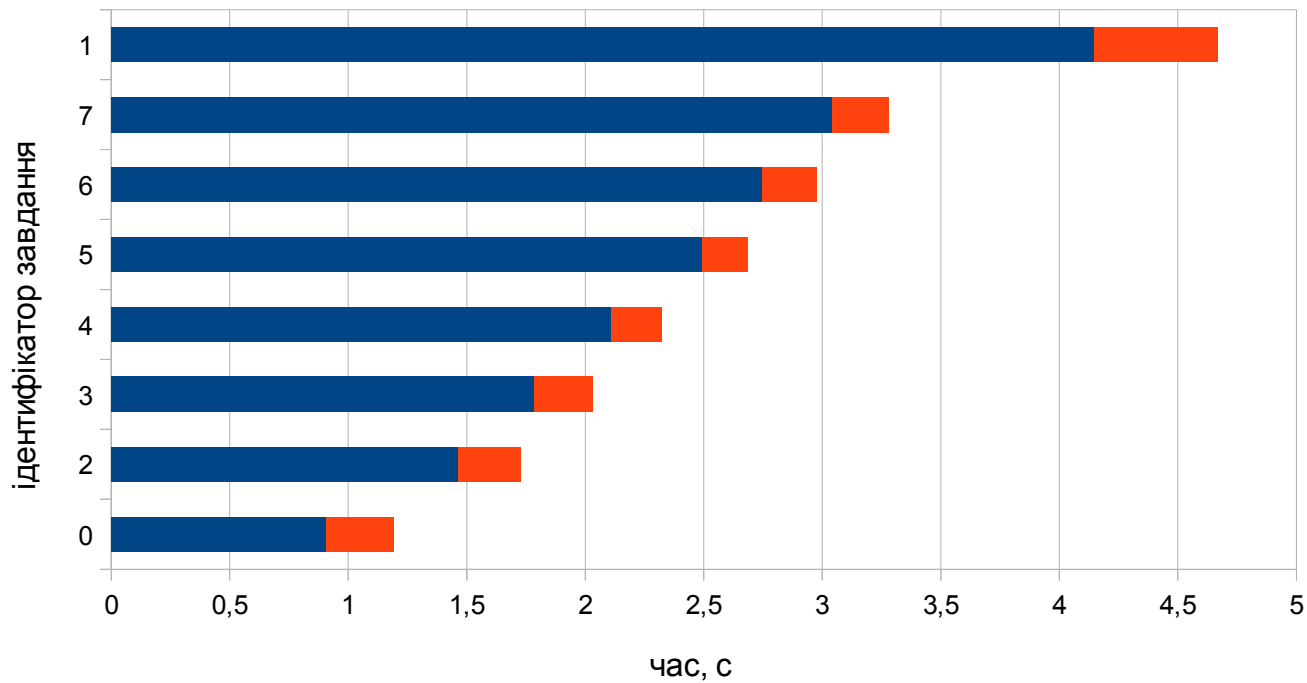
API Task Queue

- Задача розбивається на завдання
- Завдання додаються в чергу
- Завдання відправляються на обробку в залежності від завантаження системи

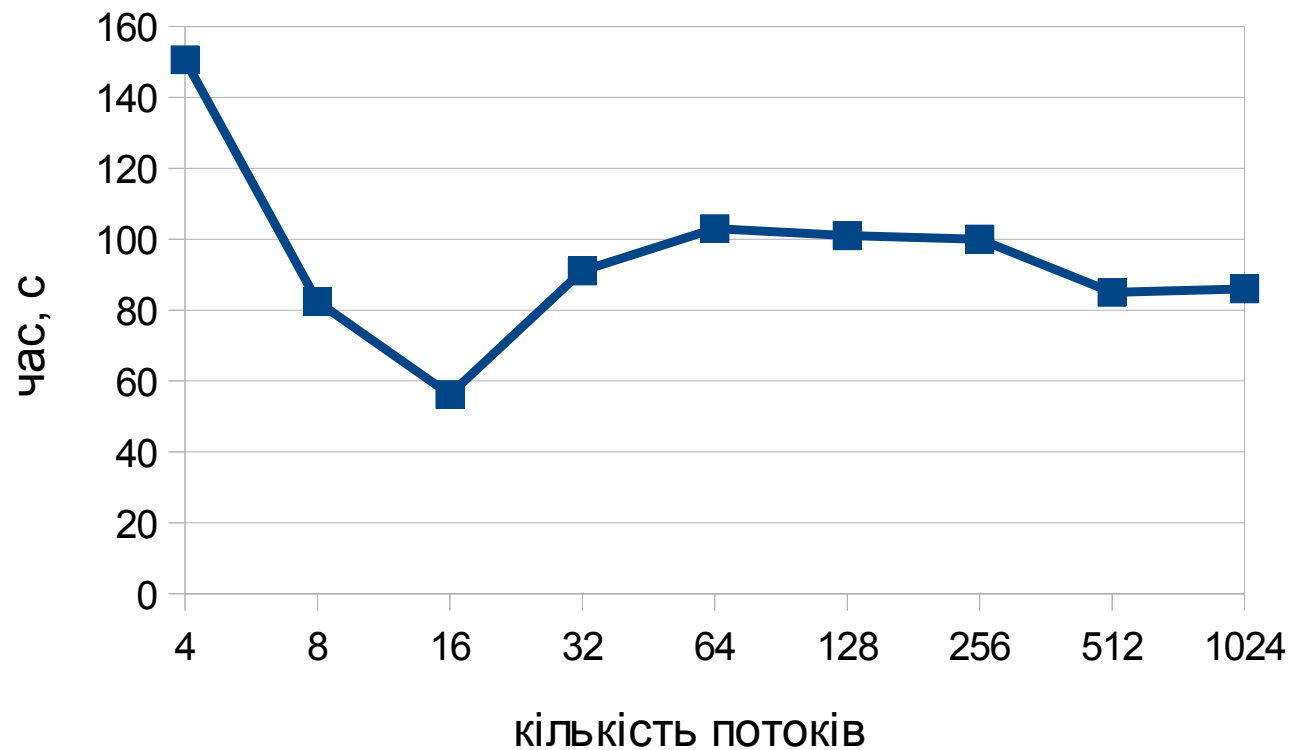
Результати роботи API Task Queue в GAE(1/4)



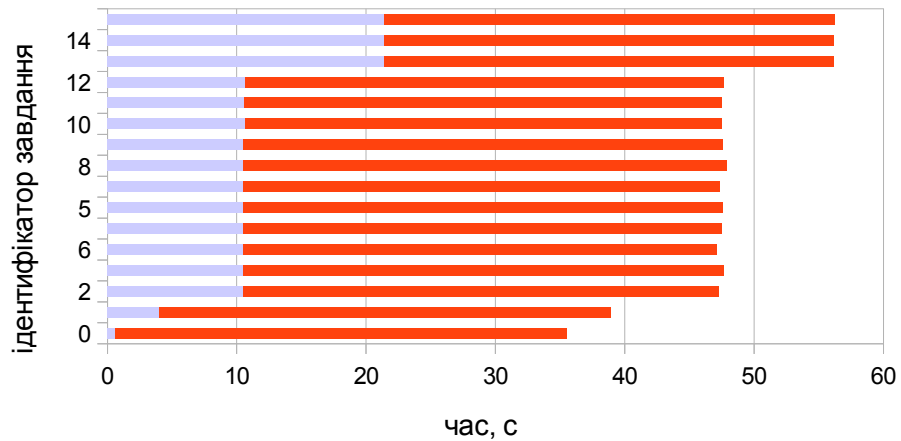
Результати роботи API Task Queue в GAE(2/4)



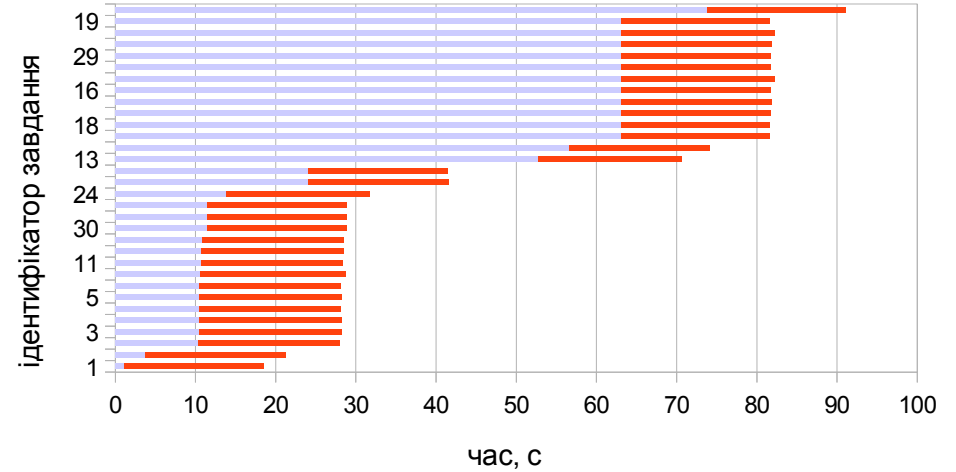
Результати роботи API Task Queue в GAE(3/4)



Результати роботи API Task Queue в GAE(4/4)



16 ПОТОКІВ



32 ПОТОКИ

Висновки(1/2)

- Декомпозиція задач можлива в хмарних обчисленнях.
- Бібліотеку AppEngine-MapReduce можна використовувати в обчислювальних задачах та для обробки дуже великих об'ємів даних.
- API Task Queue також можна використовувати в обчислювальних задачах

Висновки(2/2)

- MapReduce для декомпозиції краще використовувати для задач з інтенсивною роботою з базами даних, а Task Queue краще використовувати для декомпозиції обчислювальних задач.
- На безкоштовному акаунту в Google App Engine можливо запускати не більше 16 потоків паралельно.

Дякую за увагу