КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет комп’ютерних наук та кібернетики

Кафедра математичної інформатики

Звіт з з курсу “Хмарнi обчислення”

Виконав

студент 4 курсу

групи МІ-4

Тичковський Олексій

Постановка задачі

Нехай маємо масив чисел довжини n, треба дізнатися скільки з чисел [1..n] зустрічаються в даному масиві.

Алгоритм розв’язання

Будемо розв’язувати наївно. Просто перебираючи кожне з чисел від 1 до n і перевіряючи чи є воно в массиві, проходом по масиву.

Тоді весь алгоритм виконується за O(n^2) часу.

При виконані алгоритму на декількох машинах розділимо числа 1..n таким чином: віддімо кожній машині числа, що за модулем k дорівнюють i (де k – кількість машин, i – номер машини).

Реалізація

Розподіл вхідного масиву між воркерами та збір результатів:

def solve(self):  
 # print("Job started")  
 # print(f"Starting with {self.workers\_cnt} workers")  
 self.read\_input()  
 n = len(self.arr)  
  
 chunks = [[] for \_ in range(self.workers\_cnt)]  
 for i in range(n):  
 chunks[i % self.workers\_cnt].append(i)  
  
 mapped = []  
 for i in range(self.workers\_cnt):  
 mapped.append(self.workers[i].worker\_solve(chunks[i], self.arr))  
 reduced = self.reduce(mapped)  
  
 self.write\_output(reduced)

Алгоритм, що виконується на кожному воркері:

def worker\_solve(checkers, arr):  
 result = 0  
 for i in range(len(checkers)):  
 for j in range(len(arr)):  
 if arr[j] == checkers[i]:  
 result += 1  
 break  
 return [result]

Результати роботи

|  |  |  |  |
| --- | --- | --- | --- |
|  | 1 процес | 2 процеси | 4 процеси |
| 100 елементів | 0:0:1 | 0:0:1 | 0:0:1 |
| 1000 елементів | 0:0:1 | 0:0:1 | 0:0:1 |
| 10000 елементів | 0:0:5 | 0:0:3 | 0:0:2 |
| 50000 елементів | 0:2:6 | 0:1:13 | 0:0:38 |
| 100000 елементів | 0:8:6 | 0:4:9 | 0:2:37 |

Посилання на github: https://github.com/aleksistoren/parcs\_lab