Extend CKK with Parallel

Big Data Parallel

Hanyi Zhang

Shenzhen University

January 16, 2021

Table of contents

Algorithms Classification

Algorithms combinatorial optimization problems can be divided into 3 classes

- 1. find an optimal solution eventually, but that run in exponential time
- 2. polynomial-time algorithms that only find approximate solutions.
- 3. between 1. and 2., generally find better solutions the longer they are allowed to run.

By paralleling to extend algorithm 3., we can get better solution in the same time.

Problem Formulation

Partition Problem: Given a set of integer numbers $S = \{s_1, s_2, ..., s_n\}$, and we want to find a function F to partition S into two subsets S1, S2 such that the difference between the sum of S1, S2 is minimized.

$$S_1, S_2 = \underset{S_1, S_2 = F(S)}{\arg\min} |\sum_{n=1}^{k} S_1 - \sum_{n=1}^{n-k} S_2|. \quad (1)$$

Some Examples

1. Given S = {3, 1, 1, 2, 2, 1}, a valid solution to the partition problem is the two sets S1 = {1, 1, 1, 2} and S2 = {2, 3}

2. Given S = {8, 7, 6, 5, 4}, a valid solution to the partition problem is the two sets S1 = {8, 7} and S2 = {6, 5, 4}

3. Given S = {13, 12, 11, 10, 9, 8, 7}, a valid solution to the partition problem is the two sets S1 = {13, 12, 10} and S2 = {11, 9, 8, 7}

Related Algorithms List

Related algorithms: consider polynomial-time approximation algorithms, and then optimal algorithms for large problem instances

- Dynamic programming
- Greedy heuristic
- **Kurmarkur-Kurp heuristic(KK)**
- Complete Greedy Algorithm (CGA)
- **Complete Karmarkar-Karp (CKK)**

KK — From KK to CKK

KK[2](Kurmarkur-Kurp heuristic) in (8, 7, 6, 5, 4):

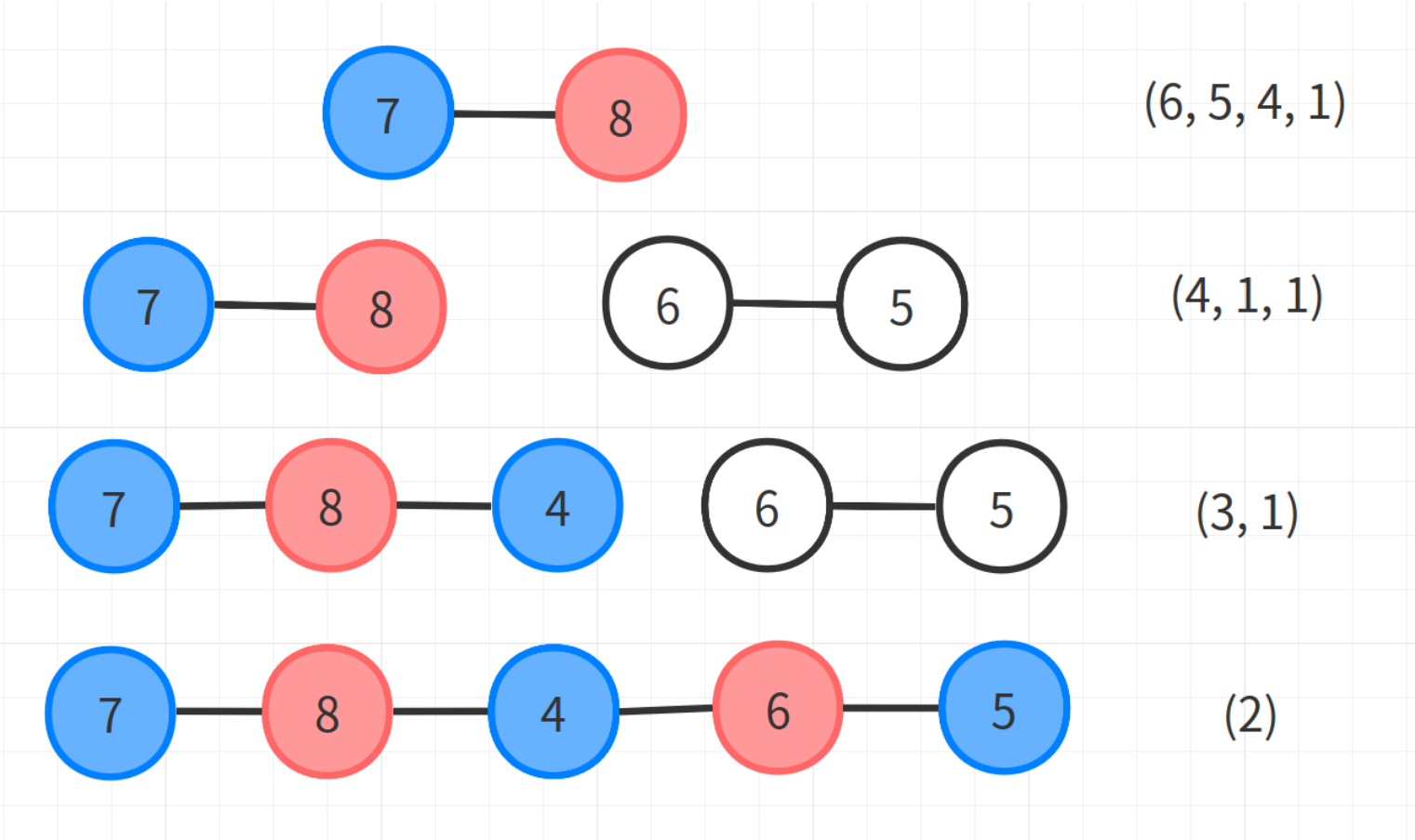


Figure: Tree from KK partitioning of (4, 5, 6, 7, 8)

CKK — From KK to CKK

CKK[1](Kurmarkur-Kurp heuristic) in (8, 7, 6, 5, 4):

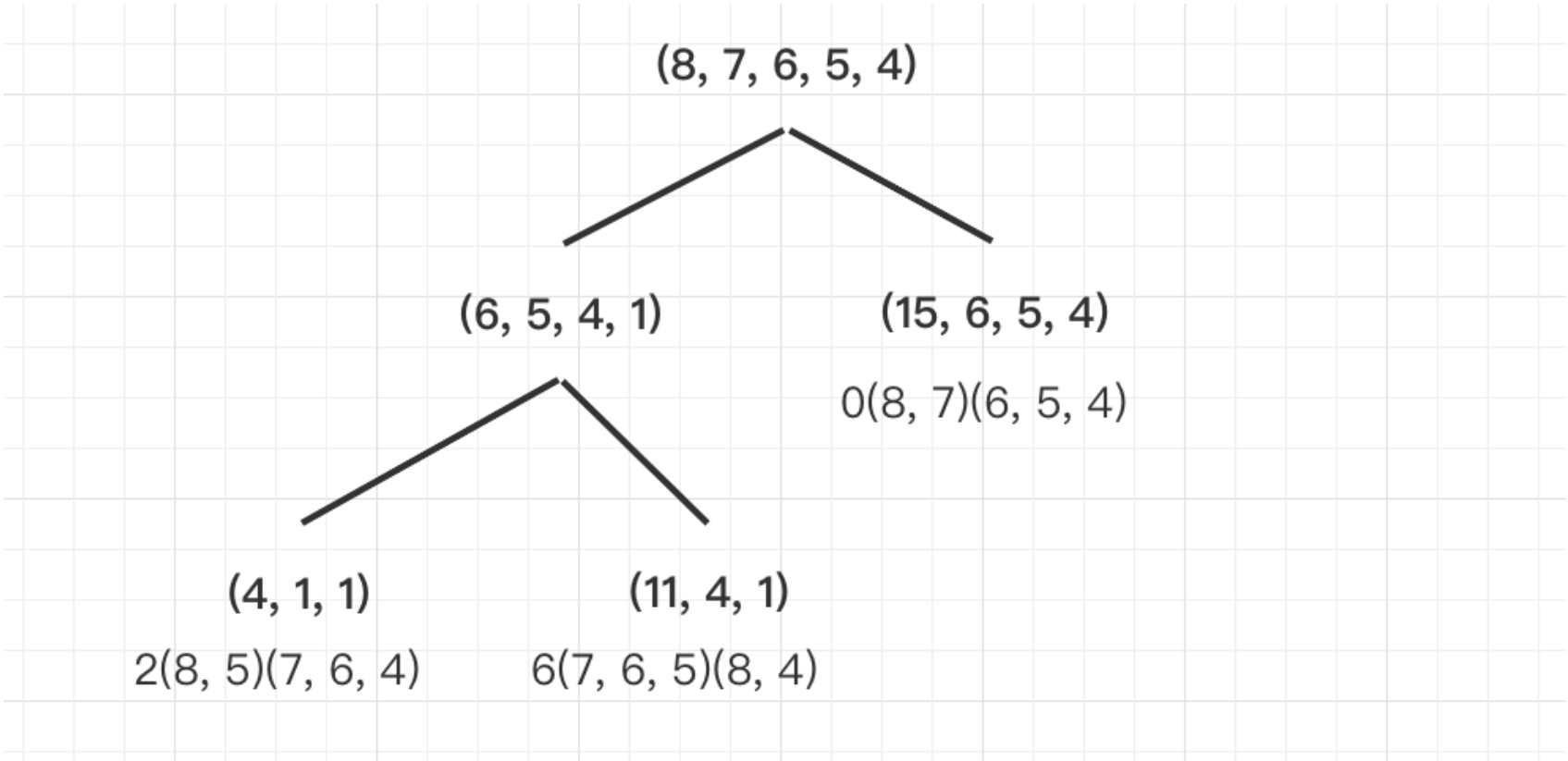


Figure: Tree generated by the CKK algorithm to partition (4, 5, 6, 7, 8)

WHY — From KK to CKK

CKK[1](complete Karmarkar and Karp)is to extend the KK heuristic to a complete algorithm.

CKK Advantages:

- 1. According to experimental results, problem instances with more than 100 numbers, CKK appears to be asymptotically faster than the best existing algorithms
- 2. the first solution found by CKK is the KK solution, and as it continues to run it finds better solutions, until an optimal solution is found.
- 3.the convenient generalization of CKK to partitioning into more than two subsets.

Extend CKK to Parallel

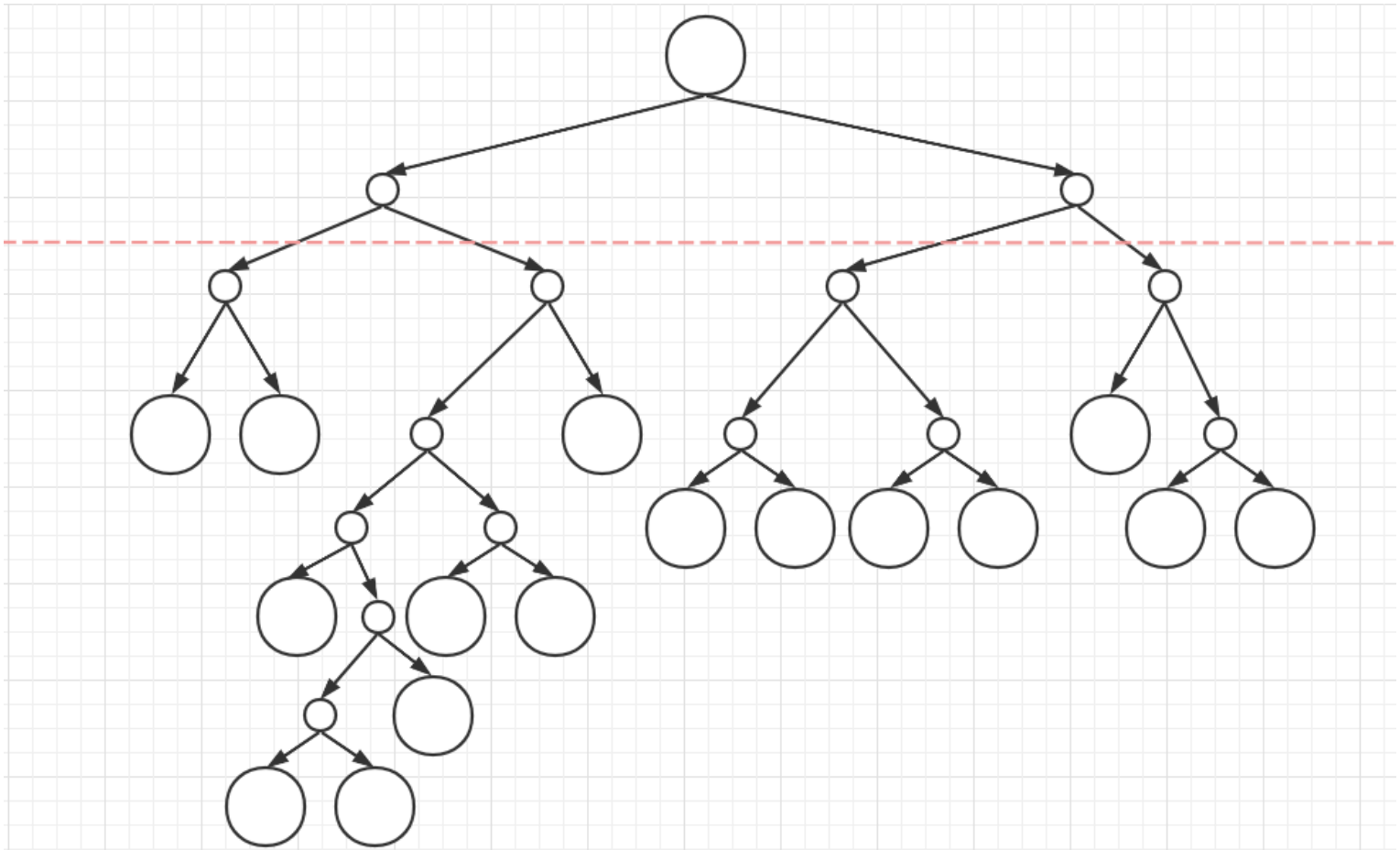
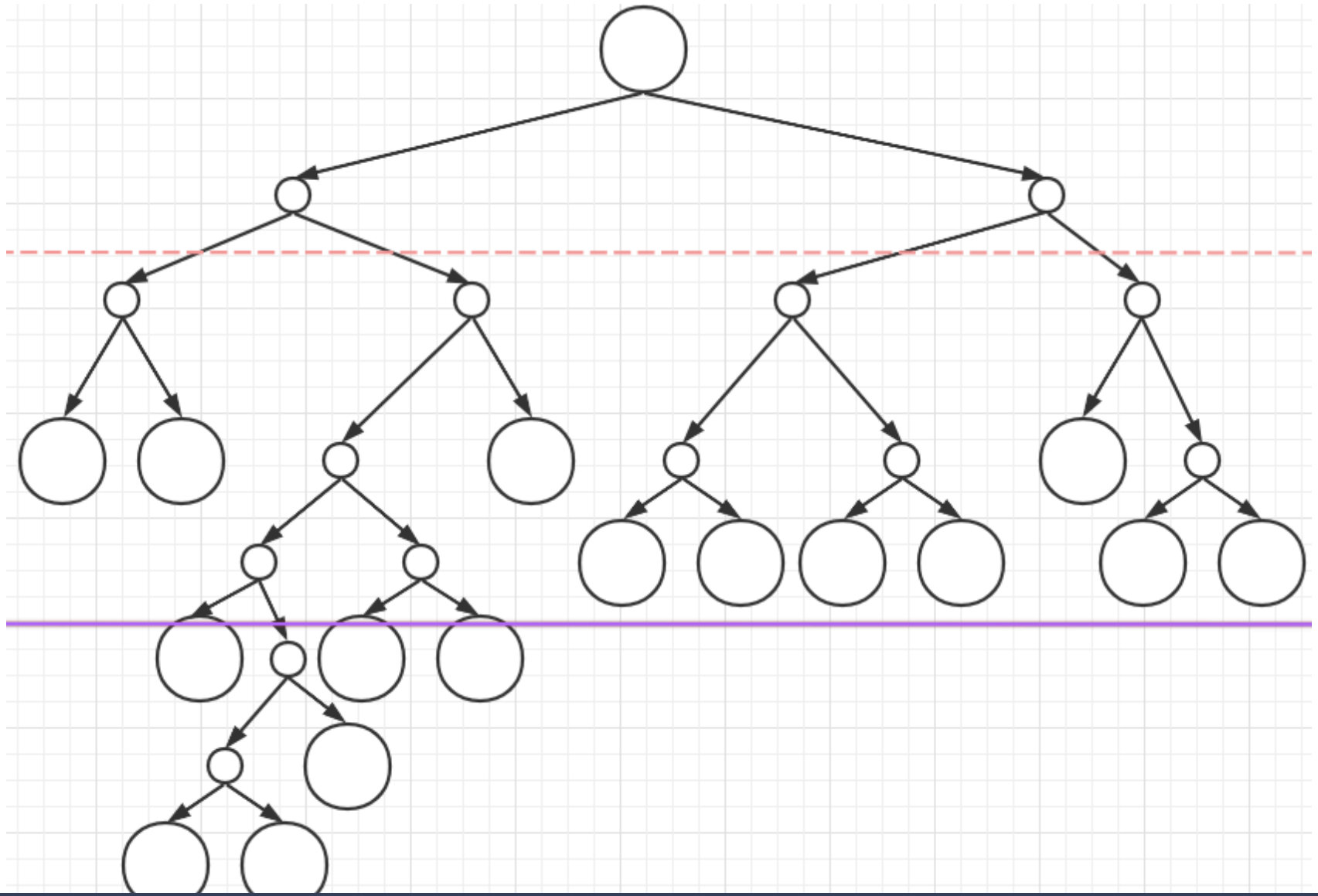


Figure: 1. Create parallel based on subpartitions of CKK

Extend CKK to Parallel



Serial CKK

n = 85

1 135593 2 784553 3 88829 4 2087303 5 17347515

6 4077319 7 1407140 8 362729 9 3172672 10 7905105

trial = 10

average: 并行 116s

average: 串行79s

Figure: Comparison between Serial and Parallel Algorithm

Parallel CKK

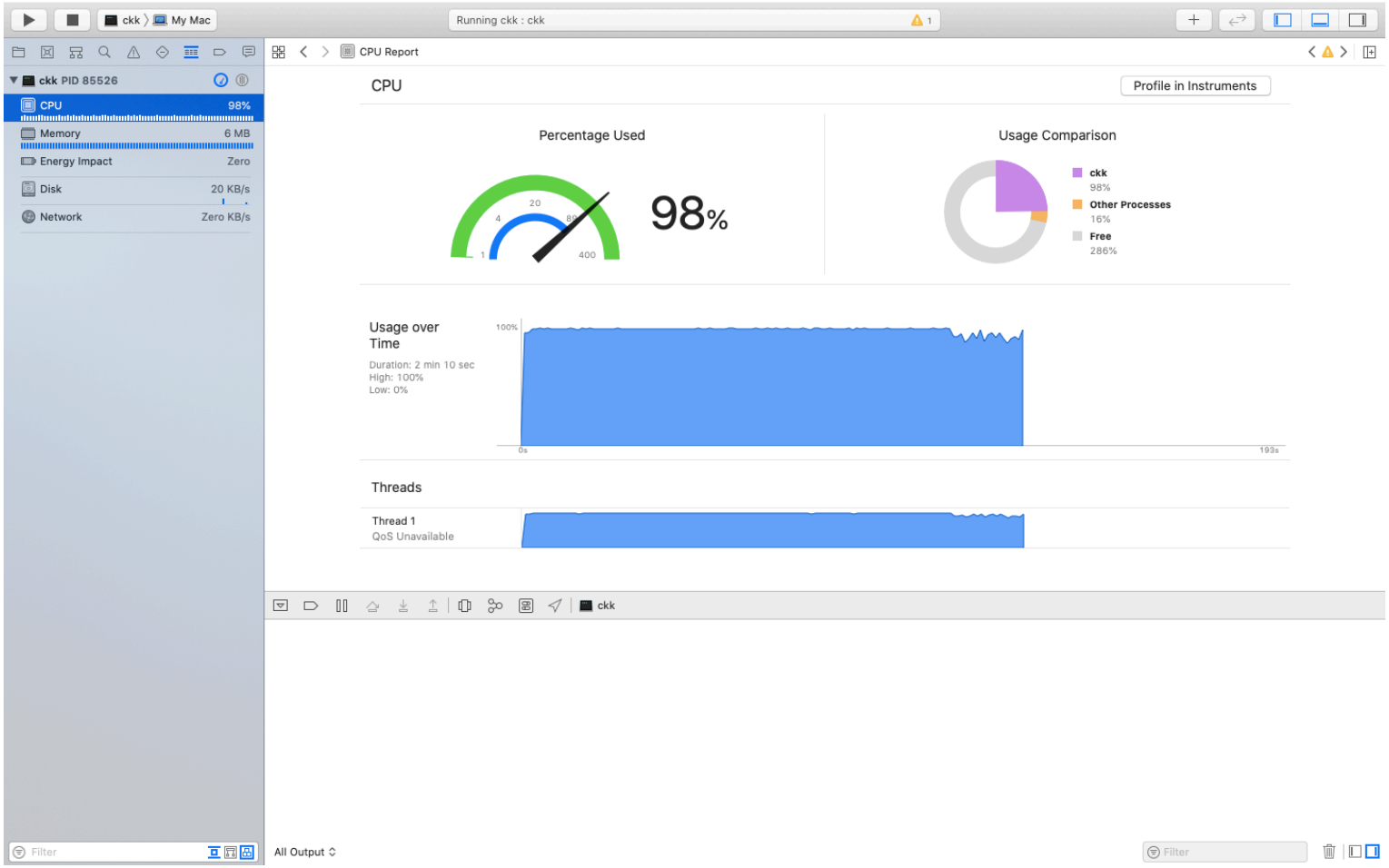


Figure: Serial CKK

Parallel CKK

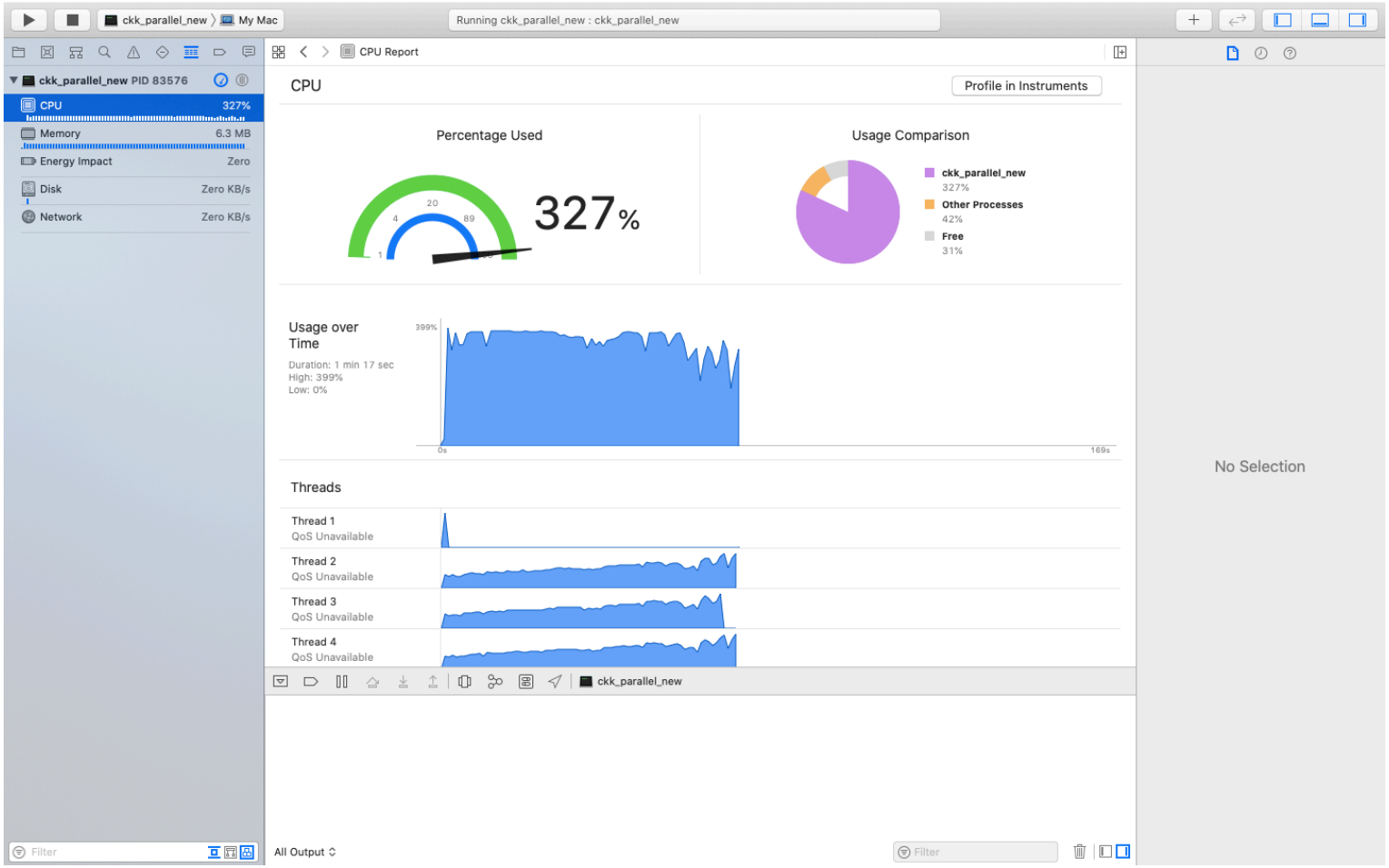


Figure: Parallel CKK

Parallel CKK

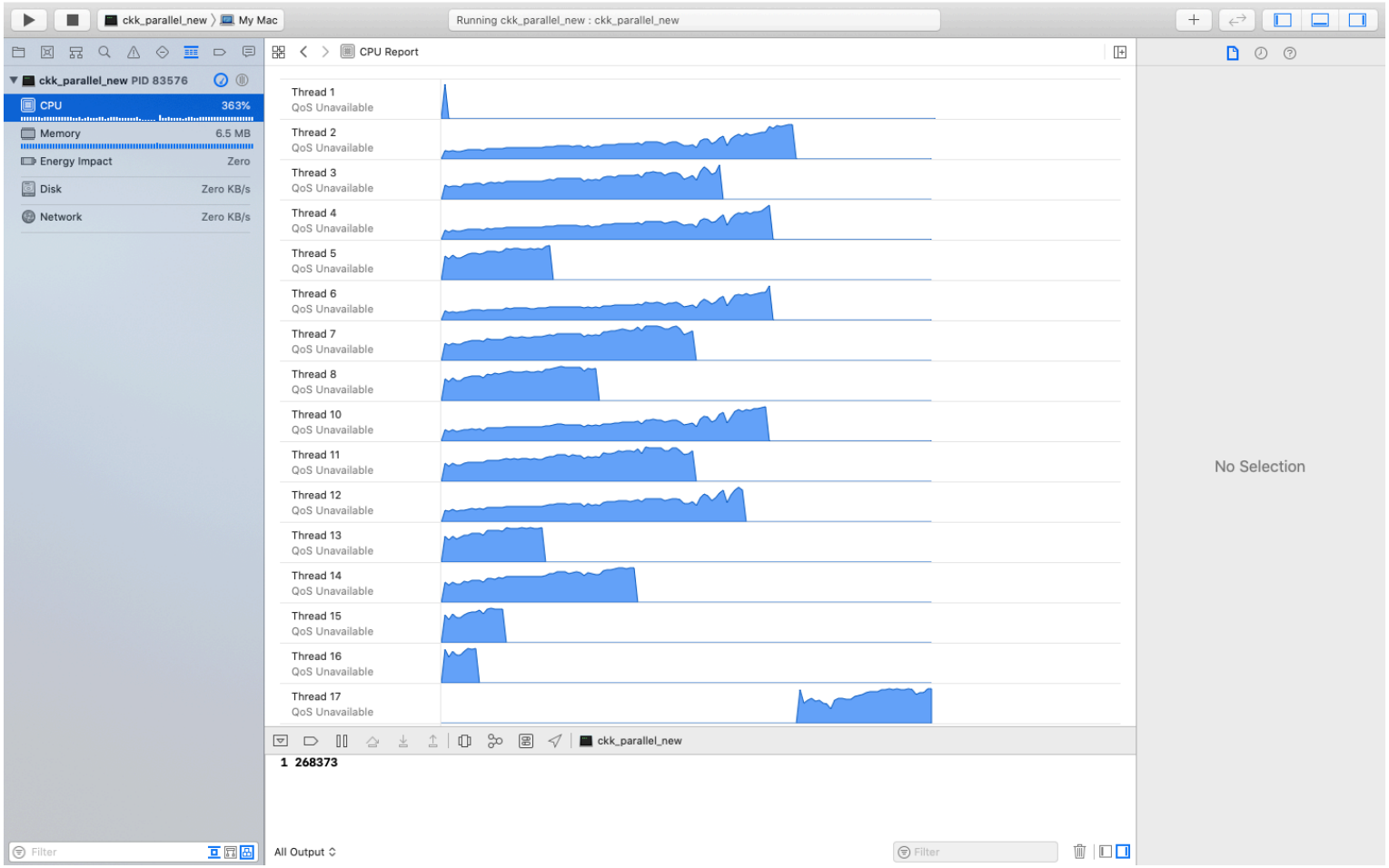


Figure: Parallel CKK

reference

[1] Korf R E. A complete anytime algorithm for number partitioning[J]. Artificial Intelligence, 1998, 106(2): 181-203.

[2] Karmarkar N, Karp R M. The differencing method of set partitioning[M]. Berkeley: Computer Science Division (EECS), University of California, 1982.

[3] Korf R E. Multi-way number partitioning[C]//IJCAI. 2009, 9: 538-543.

[4] Schreiber E L, Korf R E, Moffitt M D. Optimal multi-way number partitioning[J]. Journal of the ACM (JACM), 2018, 65(4): 1-61.