

Practical applications

January - March 2023

1. Algorithmic application

Consider a data file containing requests for some virtual machines (VMs) corresponding to specific virtual network functions of services. These VMs are to be run onto servers in DataCenters. Each VM is characterized by (name, nb vCPU, Memory (GB), DiskSpace (GB), class of service (S1, S2, S3)), whereas each server is characterized by (name, nb vCPU available, Memory (GB) available, DiskSpace available (GB)).

What would be the optimal packing of VMs into servers so as to minimize the number of servers needed ?

Propose a mathematical formulation and algorithms providing a feasible solution. Explain, and give some ways to derive upper and lower bounds on the optimal solution for each of the following VMs placement variants:

- (a) nominal case
- (b) anti-affinity rules between some set of VMs, some VMs are responsible for services that can't be shared with some others.
- (c) all servers are partially loaded vs totally empty and all with the same characteristics
- (d) VMs could be splitted over several servers
- (e) Consider VMs families, each family is given a criticality level between 1 to 3 (Class 1 can't share physical infra with VMs of class 3).
- (f) Are these algorithms adaptable to online cases ?

Synthetic data will be generated to test the algorithms.