

Projet d'optimisation métaheuristique

Workflow scheduling on cloud computing

Ruben Aloukou, Célia Benmessaouda, Chléo Chevrier et Thomas Fiou

CY Tech, CY Cergy Paris Université

Troisième année du cycle ingénieur option Intelligence artificielle

Lundi 06 février 2022

1. Modélisation du cloud computing

Nous avons choisi de modéliser le cloud computing par une infrastructure IaaS, cette infrastructure est un modèle qui est composée d'un ensemble de machines virtuelles distinctes (abrégées en VM). On note alors cet ensemble de machines $U = \{ VM_1, \dots, VM_m \}$ qui sont connectées entre elles, où m désigne le nombre de machines virtuelles. Pour $j \in [1, m]$, chaque instance de machine virtuelle VM_j est définie par :

- sa vitesse de calcul notée $MIPS_j$ correspondant au nombre d'instructions traitables
- son coût monétaire d'utilisation par unité de temps noté $prixU_j$
- une fiabilité notée λ_j

Le comportement de l'infrastructure IaaS peut être modélisé sous forme d'un graphe pondéré non orienté. Les nœuds du graphe représentent les machines virtuelles et les arêtes correspondent aux connexions entre elles. Les poids des arêtes représentent les vitesses de transferts des connexions entre les VMs.

2. Modélisation du workflow

Nous modélisons une application de workflow sous forme d'un graphe orienté acyclique noté $G(V,E)$ et on note l'ensemble de nœuds $V = \{ T_i \mid i \in [1, n] \}$ qui représente l'ensemble des tâches du workflow, où n désigne le nombre de tâches.

L'ensemble des arcs E désigne les contraintes de précédence entre les tâches. Pour $(i,j) \in [1, n]^2$, un arc est sous la forme $d_{ij} = (T_i, T_j)$, où T_i est appelée la tâche mère de T_j , T_j est la tâche fille de T_i et d_{ij} sont les données produites par T_i et consommées par T_j . Nous supposons que la tâche fille ne peut pas être exécutée avant que toutes ses tâches mères soient achevées.

3. Définitions des fonctions objectifs

a. Temps de complétion

L'une des mesures les plus utilisées dans l'évaluation des algorithmes d'ordonnancement de workflows est le temps de complétion ou makespan. Le makespan est la différence entre la date de soumission du workflow et la date de réception des résultats. On le définit de la manière suivante :

$$Makespan = \max \{ DF(T_i) \mid T_i \in V \}$$

- $DF(T_i)$ est la date d'exécution de la tâche T_i
- V est l'ensemble des tâches du workflow

b. Coût

Suite à la caractéristique orientée marché des services actuels, la plupart des fournisseurs de cloud ont fixé des prix pour l'utilisation de leurs services. Ils ont fixé le prix du transfert d'une unité de données (par exemple, par MB) entre deux services et le prix pour le traitement par unité de temps (par exemple, par heure).

$$Cost = C_{ex} + C_{tr} = \sum_{i=1}^n DF(T_i) \times prixU_j + \sum_{i=1}^n \sum_{j=1}^n D_{ij} \times TRC_{ij}$$

- C_{ex} est le coût d'exécution d'un workflow
- C_{tr} est le coût de transferts des données
- $DF(T_i)$ est la date de fin d'exécution de la tâche T_i
- $prixU_j$ est le prix unitaire de la machine
- D_{ij} est quantité de données à transférer de la tâche T_i à la tâche T_j
- TRC_{ij} est coût de communication entre la machine i et j

c. Disponibilité

La disponibilité est utilisée pour déterminer dans quelle mesure l'ensemble de machines virtuelles louées sont utilisées.

$$Availability = \frac{1}{m} \times \left(1 - \sum_{j=1}^m \frac{TimeStart_j - TimeEnd_j}{Makespan} \right)$$

- $TimeStart_j$ est la date où la VM_j est lancée
- $TimeEnd_j$ est la date où la VM_j est arrêtée.
- m est le nombre de VMs

4. Problème d'optimisation multi-objectif

Le problème d'ordonnancement de workflow revient à construire un mapping M des tâches vers les VMs qui minimise à la fois le makespan, le coût et la disponibilité.

$$\begin{cases} \min & \text{Makespan}(M) \\ \min & \text{Cost}(M) \\ \min & \text{Availability}(M) \end{cases} \Leftrightarrow \begin{cases} \min & \frac{\text{Makespan}(M)}{\max \text{Makespan}(M)} \\ \min & \frac{\text{Cost}(M)}{\max \text{Cost}(M)} \\ \min & \frac{\text{Availability}(M)}{\max \text{Availability}(M)} \end{cases}$$

Contraintes possibles (à imposer) :

- Makespan (Deadline à respecter)
- Coût (Budget à respecter)
- Disponibilité
- *Nombre de VMs utilisées*

Métaheuristique(s) retenue(s) :

- NSGA II
- PSO (*optionnel*)