

Algorytmy online dla problemów konfiguracyjnych

Rozważmy typową sytuację występującą na co dzień w ośrodku obliczeniowym. W takim miejscu znajduje się pewna liczba połączonych ze sobą komputerów, a na każdym z nich można uruchomić określoną liczbę maszyn wirtualnych. Maszyny wirtualne tworzą sieć jednostek obliczeniowych i muszą się ze sobą komunikować. Podczas gdy komunikacja między dwiema maszynami wirtualnymi działającymi na jednym komputerze jest w zasadzie darmowa, jeśli działają one na różnych komputerach, wiąże się ona z pewnym kosztem przesyłania danych przez sieć.

Klasycznym zadaniem optymalizacyjnym jest obliczenie takiego rozmieszczenia maszyn wirtualnych (takiej *konfiguracji*), aby możliwości obliczeniowe komputerów nie zostały przekroczone, a całkowity koszt komunikacji został zminimalizowany. Okazuje się, że obliczenie dobrej konfiguracji jest trudne obliczeniowo, nawet jeśli wiadomo, ile danych będą przesyłać konkretne maszyny.

W rzeczywistości żądania komunikacji są generowane przez obliczenia wykonywane na poszczególnych maszynach wirtualnych i dlatego nie są znane z góry. W efekcie każde statyczne rozmieszczenie maszyn wirtualnych może szybko okazać się nieoptymalne. Nowoczesne rozwiązania wirtualizacyjne umożliwiają rozwiązanie tego problemu na poziomie technicznym: możemy przenosić maszyny wirtualne pomiędzy komputerami. Takie działanie wiąże się z pewnym kosztem związanym z uspieniem, migracją i wznowieniem pracy. Na poziomie algorytmicznym jest to jednak poważna zmiana paradygmatu: żądania komunikacji pojawiają się w *sposób online*, jedno po drugim, a w odpowiedzi na nie algorytm musi zmieniać konfigurację (przypisanie maszyn do komputerów). Te algorytmiczne decyzje muszą być podejmowane wyłącznie na podstawie przeszłych żądań komunikacyjnych i bez wiedzy o przyszłych.

Opisany powyżej problem dynamicznego przypisania maszyn wirtualnych do komputerów jest przykładem *problemu optymalizacyjnego online*, którego celem jest tania obsługa ciągu nieprzewidywalnych żądań. Aby zminimalizować koszt, możliwa jest *dynamiczna modyfikacja konfiguracji*, tak żeby była ona jak najlepiej dopasowana do żądań, ale takie zmiany muszą być dokonywane bez wiedzy o przyszłych żądaniach. Problemy pasujące do powyższego opisu, nazywamy *problemami konfiguracyjnymi*. Występują one jako kluczowe elementy w wielu zastosowaniach, począwszy od sieci komputerowych, przez organizację ośrodków obliczeniowych, logistykę, aż po metody tworzenia wydajnych struktur danych.

Rozważanie całej klasy problemów konfiguracyjnych zamiast konkretnego zagadnienia rozmieszczenia maszyn pozwala nam abstrahować od szczególnego charakteru danego zagadnienia i konstruować uniwersalne metody, mające zastosowanie dla całej gamy problemów konfiguracyjnych. Niemniej jednak ogólne rozwiązania muszą zostać później dostrojone do geometrii przestrzeni konfiguracji i typów żądań. Celem tego projektu jest *konstrukcja i rygorystyczna analiza algorytmów dla problemów konfiguracyjnych w warunkach nieprzewidywalnych żądań*.