Offre de stage Recherche M2

Title : Energy efficient learning algorithms for controlling both energy and performance in data centers

Subject :

With rapid growth of new services such as video-streaming, augmented reality, online games, requiring computational resources in data-centers, network operators face two main challenges : QoS guarantee, and resource energy consumption.

We consider a distributed service architecture executed on physical servers in tandem, hosting virtual machines. Users service requests arrive in the first server node and then go to the next node, to execute step by step each task of the service. An autonomous agent controls the number of active resources (Virtual Machines) on each physical server. We model the system by a multidimensional MDP(Markov Decision Process), whose size could be very large.

We focus on reinforcement learning techniques in order to learn the activation/deactivation policy of the resources. Indeed, such approaches allow greater flexibility and adaptability to the system. Unfortunately, as the state space size becomes very large then learning algorithms could be too slow with bad convergence time, and so consuming too much resources and therefore energy. We propose to adapt and develop reinforcement learning techniques based on decomposition approaches and *Hierarchical Reinforcement Learning* [1, 2] in order to accelerate learning algorithms. We focus on two approaches: *1-State abstraction* which consists in decomposing the large state space into regions where we define smaller MDP for determining local policy, from which the optimal strategy on the entire state space could be found from a hierarchical construction. *2-Abstract action with option* which is based on applying temporal abstraction to the problem (called *options*): decision-making should not be required at every step but instead temporally extended activities can be selected to achieve sub goals. Similarly to divide and conquer techniques, the benefit of decomposition techniques is the resolution of sub-problems on smaller state spaces (or regions), the trade off being that extra effort is required to combine the solutions of these sub problems into a solution to the original problem. Also the definition of the regions which has an impact on the quality of the global solution.

We propose in this project to tackle these problems by developing energy efficient leaning algorithms for dynamic control of resources in data centers, guaranteeing both QoS and energy consumption. The main steps of the internship are :

- State of the art about hierarchical reinforcement learning techniques

- Algorithm development based on these techniques applied on tandem servers
- Comparison of their efficiency (convergence time) with other algorithms

Candidate profile Master 2 research internship, good knowledge of probability, optimization techniques, and some knowledge of Markov Decision Process and reinforcement learning techniques is appreciated.

Contact : The CV must be sent to the three e-mails :
Erwan.Le-Pennec@polytechnique.edu, hind.castel@telecom-sudparis.eu, Emmanuel.Hyon@parisnanterre.fr.

Source of funding and place of the internship: This internship is funded by E4C (Centre sur l'energie et le climat) of Institut Polytechnique de Paris for 6 months. The subject can lead to a thesis and the supervisory team will seek funding for a thesis. The internship will take place at Palaiseau (Ecole Polytechnique and Télécom SudParis), and remote work organization is possible.

References

[1] C. Daoui, M. Abbad, and M. Tkiouat. Exact Decomposition Approaches for Markov Decision Processes: A Survey. *Advances in Operations Research*, 2010.

[2] B. Hengst. *Reinforcement Learning: State-of-the-Art*, chapter Hierarchical Approaches, page 293–323. Springer, 2012.