

INTISARI

MANAJEMEN DAYA PADA SISTEM HPC MENGGUNAKAN METODE DEEP REINFORCEMENT LEARNING BERBASIS ADVANTAGE ACTOR-CRITIC

Oleh

FITRA RAHMANI KHASYAH

18/424186/PA/18291

Tingginya penggunaan energi merupakan salah satu permasalahan utama dalam menjalankan sistem High-Performance Computing (HPC). Pada umumnya, sistem HPC terdiri dari ratusan hingga ribuan *node* komputasi yang menggunakan banyak daya listrik meskipun sedang dalam keadaan *idle*. Salah satu cara untuk meningkatkan efisiensi manajemen daya adalah dengan menerapkan metode *Backfilling* pada *job scheduler* First Come First Serve (FCFS) (FCFS + *Backfilling*). Metode *Backfilling* memperbolehkan *job* yang datang setelah *job* pertama di antrian untuk dieksekusi terlebih dahulu jika waktu mulai *job* pertama tidak tertunda sehingga meningkatkan *throughput* dan efisiensi energi sistem. *Node* yang telah *idle* selama periode waktu tertentu juga dapat ditidurkan untuk meningkatkan efisiensi energi lebih lanjut. Namun, menidurkan *node* berdasarkan waktu *idle* dapat mengurangi efisiensi energi dan *throughput* sistem alih-alih meningkatkannya. Sebagai contoh, *job* baru dapat langsung tiba setelah *node* ditidurkan sehingga menghilangkan peluang untuk langsung mengeksekusi *job* tersebut melalui *Backfilling*.

Dalam penelitian ini, diusulkan metode berbasis Deep Reinforcement Learning (DRL) untuk memprediksi *timing* yang tepat untuk menidurkan atau menyalakan *node*. Agen DRL dilatih menggunakan algoritma Advantage Actor-Critic (A2C) untuk menentukan *node* mana yang harus ditidurkan atau dinyalakan pada *timestep* tertentu. *Reward* yang digunakan untuk melatih agen A2C-DRL terdiri dari pemborosan energi dan rata-rata *job waiting time*. Untuk evaluasi, agen A2C-DRL dibandingkan dengan metode *timeout policy* (*node* dimatikan setelah *idle* selama periode tertentu) dan *baseline* (*node* tidak dimatikan sama sekali) menggunakan simulator Batsim-py.

Hasil simulasi pada dataset *job* HPC DIKE UGM menunjukkan bahwa agen A2C-DRL mengungguli semua konfigurasi dan menghemat 58,2% energi dibanding *timeout policy* optimal (1 menit). Pada dataset *job* NASA iPSC/860, agen A2C-DRL mengungguli 11 dari 12 *timeout policy*, menghemat hingga 21,34% energi dan hanya kalah 2% dibanding *timeout policy optimal* (15 menit). Agen A2C-DRL cenderung mematikan *node* lebih sering sehingga memiliki rata-rata *job waiting time* yang ting-

gi, namun agen menghemat lebih banyak energi dibandingkan metode *timeout policy* yang sama agresifnya. Hasil ini membuktikan bahwa agen A2C-DRL mampu mematikan dan menyalakan *node* pada waktu yang tepat.

Kata-kata kunci : HPC, Manajemen Daya, Penggunaan Energi, Deep Reinforcement Learning, Advantage Actor-Critic.

ABSTRACT

An Advantage Actor-Critic Deep Reinforcement Learning Method for Power Management in HPC Systems

By

FITRA RAHMANI KHASYAH

18/424186/PA/18291

The high energy consumption is one of the major issues in running High-Performance Computing (HPC) systems. Typically, HPC systems consist of hundreds to thousands of computing nodes that consume a significant amount of electricity even when idle. One way to increase power management efficiency is by implementing the Backfilling method on the First Come First Serve (FCFS) job scheduler (FCFS + Backfilling). The Backfilling method allows jobs arriving after the first job in the queue to be executed first if the starting time of the first job is not delayed, thus increasing the throughput and energy efficiency of the system. Nodes that have been idle for a certain period of time can also be put to sleep to further increase energy efficiency. However, putting nodes to sleep based solely on idle time can actually decrease energy efficiency and system throughput instead of increasing it. For example, a new job can arrive immediately after a node is put to sleep, thereby eliminating the opportunity to immediately execute the job via Backfilling.

In this study, we propose a Deep Reinforcement Learning (DRL)-based method to predict the proper timing for putting nodes to sleep or turning them on. The DRL agent is trained using the Advantage Actor-Critic (A2C) algorithm to determine which nodes should be put to sleep or turned on at a particular timestep. The reward used to train the A2C-DRL agent consists of energy waste and average job waiting time. For evaluation, the A2C-DRL agent is compared with the timeout policy method (nodes are turned off after idle for a certain period) and baseline method (nodes are never turned off) using the Batsim-py simulator.

The simulation results on the DIKE UGM HPC job dataset show that the A2C-DRL agent outperforms all configurations and saves 58.2% of energy compared to the optimal timeout policy (1 minute). On the NASA iPSC/860 job dataset, the A2C-DRL agent outperforms 11 of the 12 timeout policies, saving up to 21.34% of energy and only losing 2% compared to the optimal timeout policy (15 minutes). The A2C-DRL agent tends to turn off nodes more often, thus having a higher average job waiting time, but the agent saves more energy than the same aggressive timeout policy method. These results demonstrate that the A2C-DRL agent can put nodes to sleep and turn them on at the right time.

Keywords : HPC, Power management, Energy consumption, Deep Reinforcement Learning, Advantage Actor-Critic.