



Contribution ID: 237

Type: **Talk**

Resource Scheduling and Allocation in HPC Cyberinfrastructure

Wednesday, 3 December 2025 11:20 (20 minutes)

Abstract

High-Performance Computing (HPC) systems play a pivotal role in modern scientific research, enabling complex simulations, data analysis, and large-scale modelling across disciplines such as climate science, genomics, physics, and engineering. As these systems grow in scale and sophistication, the efficient scheduling and allocation of computational resources become crucial for ensuring optimal system performance, maximising resource utilisation, and meeting the needs of diverse user communities. In HPC environments, resource scheduling and allocation determine how tasks are assigned to hardware resources such as CPUs, GPUs, memory, storage, and network bandwidth. Effective scheduling strategies are critical for maintaining fairness among users, optimising job throughput, reducing waiting times, and enhancing energy efficiency. Cyberinfrastructure, the integration of advanced computing platforms with large-scale data storage and high-speed networks, adds layers of complexity to resource management. The heterogeneity of hardware, dynamic workload demands, and multi-user environments require advanced resource scheduling algorithms. Traditional approaches like First-Come, First-Served (FCFS), Shortest Job First (SJF), and Backfilling have evolved to meet these challenges, while more advanced strategies like Priority Scheduling, Gang Scheduling, and Hybrid Scheduling offer increased flexibility and efficiency. Energy-aware scheduling has also gained importance, given the significant portion of operational costs that energy consumption in large HPC systems can account for.

Presenting Author

Email

Student or Postdoc?

CHPC User

CHPC Research Programme

Workshop Duration

Primary author: Dr OGUNLEYE*, Olalekan Samuel (University of Mpumalanga)

Presenter: Dr OGUNLEYE*, Olalekan Samuel (University of Mpumalanga)

Session Classification: DIRISA