Abstract

High-Performance Computing (HPC) clusters represent one of the most complex computational environments, requiring sophisticated operating system management to coordinate thousands of processors and nodes efficiently. This report investigates the critical role of operating systems in managing HPC clusters, with particular emphasis on three fundamental aspects: job scheduling mechanisms, resource allocation strategies, and inter-node communication protocols. Through comprehensive analysis of current HPC operating systems including Linux distributions, specialized cluster OS variants, and hybrid approaches, this study examines how these systems handle the unique challenges of parallel processing, fault tolerance, and scalability. The research methodology encompasses literature review, comparative analysis of major HPC operating systems, and examination of real-world implementations across various computational domains. Results indicate that modern HPC operating systems have evolved sophisticated multi-level scheduling algorithms, dynamic resource allocation mechanisms, and high-speed communication protocols that significantly impact cluster performance. The study concludes that the choice of operating system and its configuration plays a crucial role in determining the overall efficiency, scalability, and reliability of HPC clusters, with implications for scientific computing, artificial intelligence, and large-scale data processing applications.

Keywords: High-Performance Computing, Operating Systems, Job Scheduling, Resource Allocation, Inter-Node Communication, Cluster Computing