Heterogeneous cluster environments are becoming an increasing popular platform for executing parallel applications. Efficient heterogeneous programs must account for the differences inherent in such an environment. We propose the HBSP\(^k\) model of computation as a framework for developing applications for heterogeneous clusters of workstations. The utility of the model is demonstrated through the design and analysis of the scatter and one-to-all broadcast algorithms. Extensive experimentation illustrates the benefits of using the model for heterogeneous program development. By hiding the non-uniformity of the underlying system, the HBSP\(^k\) model provides a framework that embraces the heterogeneity of the underlying system.

Keywords: heterogeneous distributed computing, BSP, cluster computing, collective communication, performance evaluation

1. INTRODUCTION
Heterogeneous cluster environments are becoming an increasing popular platform for executing parallel applications. Such environments consist of a collection of machines with myriad differences such as varying computational power and incompatible data formats. Performance gains in heterogeneous environments result from effectively exploiting the speeds of the underlying components. Executing standard (homogeneous) distributed applications on heterogeneous platforms leads to low-end systems becoming a bottleneck, which reduces overall system performance. Thus, a new approach is necessary for the design of efficient heterogeneous distributed applications.

The \(k\)-Heterogeneous Bulk Synchronous Parallel model (HBSP\(^k\)) is the model that we propose for the development of general-purpose heterogeneous applications (Williams, 2000). It is an extension of the BSP model of parallel computation (Valiant, 1990). The superscript \(k\) refers to the number of network layers present in the heterogeneous environment. For example, \(k = 0\) refers to a single-processor computer, \(k = 1\) models a network of workstations, and \(k = 2\) denotes a cluster of clusters environment. Here, we focus on the development of programs for a heterogeneous cluster of workstations, which consist of a single communication network. Thus, a specific instantiation of the generalised HBSP\(^k\) model, known as the HBSP\(^1\) model, is used to develop heterogeneous parallel algorithms for workstation clusters.