

Low Latency Workstation Cluster Communications Using Sender-Based Protocols ¹

Mark R. Swanson

Leigh B. Stoller

E-mail: {swanson,stoller}@cs.utah.edu

UUCS-96-001

Department of Computer Science

University of Utah

Salt Lake City, UT 84112, USA

January 24, 1996

Abstract

The use of workstations on a local area network to form scalable multicomputers has become quite common. A serious performance bottleneck in such “carpet clusters” is the communication protocol that is used to send data between nodes. We report on the design and implementation of a class of communication protocols, known as sender-based, in which the sender specifies the locations at which messages are placed in the receiver’s address space. The protocols are shown to deliver near-link latency and near-link bandwidth using Medusa FDDI controllers, within the BSD 4.3 and HP-UX 9.01 operating systems. The protocols are also shown to be flexible and powerful enough to support common distributed programming models, including but not limited to RPC, while maintaining expected standards of system and application security and integrity.

¹This work was supported by a grant from Hewlett-Packard, and by the Space and Naval Warfare Systems Command (SPAWAR) and Advanced Research Projects Agency (ARPA), Communication and Memory Architectures for Scalable Parallel Computing, ARPA order #B990 under SPAWAR contract #N00039-95-C-0018