network channels.

3 Experiment

3.1 Bandwidth Intensive Job Models

Our experiment is designed to measure the communication bandwidth of two unrelated applications of different bandwidth and to examine how the two different applications interact with each other. In our model, two different types of application programs run simultaneously on the Paragon; one is the bandwidth intensive application, the other is the independent non-bandwidth intensive application. Each application is sending to and receiving from the participating nodes packets of a certain size. We chose an 8X8 Paragon mesh as a test bed node allocation for the experiment.

Figure 1 (b) through (e)) shows the four different node allocation models used. Nodes on the 8X8 mesh are divided into 3 groups; BIJ (Bandwidth Intensive Job) nodes running the BIJ applications, Non-BIJ nodes running the Non-BIJ applications, and idle nodes which do nothing. The only difference between BIJ applications and Non-BIJ applications is the packet size they use. Figure 1 (a) simply shows the 32 Non-BIJ nodes (indicated by circles); their locations are fixed to get consistent responses of the same Non-BIJ nodes in our four different node allocation models. In Figure 1 (b), 4 BIJ nodes are indicated by dark circles) and 32 Non-BIJ nodes are allocated with 28 idle nodes (no circles). We also use an 8 BIJ nodes model, a 16 BIJ nodes model and a 32 BIJ nodes model with the same 32 Non-BIJ nodes as indicated in Figures 1 (c), (d) and (e), respectively.

3.2 Application program for the experiments

To create a reasonable amount of traffic in the communication channels, each application is sending and receiving packets concurrently and continuously with two main buffers, buffer A and buffer B. Two major send operations are executed in each application to finish its job. First, each node sends all its designated portions of buffer A to its designated receivers. Second, as soon as a node senses that any portion of buffer B has been filled by a designated sender of the portion of buffer B, it sends back the same portion of the buffer B to the original sender as an acknowledgment. Each node acts as a sender and a receiver at the same time. When any acknowledgment packet arrives from a designated receiver, the packet is stored into the designated space of buffer A.

Figure 2 illustrates the mechanism of sending and receiving packets of the application program. This figure shows only message passing within identical job nodes; in our experiments, BIJ nodes interact only with BIJ nodes and Non-BIJ nodes interact only with Non-BIJ nodes, but both types of nodes must share the network bandwidth. In Figure 2,

---

3The logical node numbers of the 4 BIJ nodes are 0, 15, 48 and 63 on the 8X8 mesh.

---

the numbers between two buffers of each node indicate the destinations of each packet of buffer A and the designated senders for each packet in buffer B. The first packet (shadowed portion) of Node 0's buffer A is sent to Node 1 and is stored in the shadowed portion of buffer B of Node 1; as soon as Node 1 senses that the portion is filled, it sends the packet to the original storage in buffer A of Node 0.

To avoid as much as possible traffic bottlenecks for a node at a particular time, we used the balanced round robin technique in assigning the orders of sending packets for each node. For example, the sending order of Node 0 is 1-2-3, and that of Node 2 is 3-0-1 in Figure 2. To get just the communication time resulting from the span of sends and receives, our synthetic applications avoid any bandwidth consuming operations such as reads and writes or any unnecessary steps. Each node checks the time between the start of the first send (buffers A and B are ready to receive status before this point) and the time of the confirmation that all designated packets of buffer A are replied by other participating nodes' acknowledgment sends. Note that the time measurement used in this experiment is not a strict one such as the time of actions between only two nodes, but is a comprehensive one that is the time between sending all the designated packets of buffer A (size of packets depends on the application) to many nodes and receiving all the acknowledgment packets accordingly. All measured times of each node are averaged over 10 executions on a physical 8X8 Paragon mesh. The averaged times of each node are summed up again by its type of application and