SGI Altix 4700 Memory System Benchmarking
CS213 Final Project

Danhua Guo
dguo@cs.ucr.edu

1 Introduction

The SGI Altix 4700 is a CC-NUMA shared memory system. In this project, the memory performance in terms of load/copy bandwidth at different level in the system is evaluated with Extended Copy Transfer Characterization benchmark. The evaluation focuses on two aspects of the memory hierarchy, its behavior with temporal and spatial locality. The rest of the paper is organized as follows: section 2 briefly introduces the benchmark; section 3 lies out the details of experiment and performance evaluation metrics; section 4 presents the results from the experiment with discussion and analysis; section 5 concludes the project.

2 Benchmark

The Extended Copy Transfer (ECT) characterization is a method to characterize the performance of memory systems (ISCA95 and HPCA97). The benchmarking model was developed by Swiss Institute of Technology (ETHZ). It captures two aspects of the memory hierarchy, its behavior with temporal locality by varying the working set size and the spatial locality by varying the access pattern (in our case we work with strides). Furthermore users can use the same chart for characterization of local and remote transfers in other words accesses from computation and from communication.

ECT memperf measures the memory bandwidth in a 2 dimensional way. First it varies the block size which provides information of the throughput in different memory system hierarchies (different cache levels). Secondly it varies the access pattern from contiguous blocks to different strided accesses. 4 different tests are provided:

load sum: The load sum test measures the memory load performance for all the block sizes and access patterns. It accumulates the values in order to prevent the optimizing compiler to suppress the interesting part of the test.

const store: The const store test does the reverse operation of the load test. It measures the store bandwidth for all the block sizes and access patterns.

load copy: The load copy does a strided load test and stores the result in a contiguous way. It simulates a matrix transpose. It is performed for all the block sizes and access patterns.

copy store: The copy store test is the opposite of the load copy test. It performs a contiguous load and stores the data in strides. So the result of the operation is the same as in the load copy test. Again, all the block sizes and access patterns are tested.

3 Experiment and Evaluation Metrics

All the system output is evaluated in terms of memory access bandwidth. It could be either load or store depending on different applications.

The first stage of experiments uses the *Load sum* application and compares the result between different levels of caches in the memory hierarchy. The general pattern of memory bandwidth to the size of strides is also provided. The second experiment focuses on local access with *Load copy* application which works with strided load and contiguous store. The number of processors is varied here in order to get a clearer picture of the pattern. The last experiment tries contiguous loads and strided stores, which works like remote accesses to a centralized server.

4 Result

4.1 Mode 0: Load Sum

The SGI Altix 4700 Bandwidth System has 16KB Primary Cache, 1MB Secondary Cache, 12MB L3 Cache per core. The 32 cores share a memory of 128GB. However, for regular users, the accessible size of memory is limited to 8GB. Therefore for each core, the actual working main memory size is 8GB/64 = 125MB.

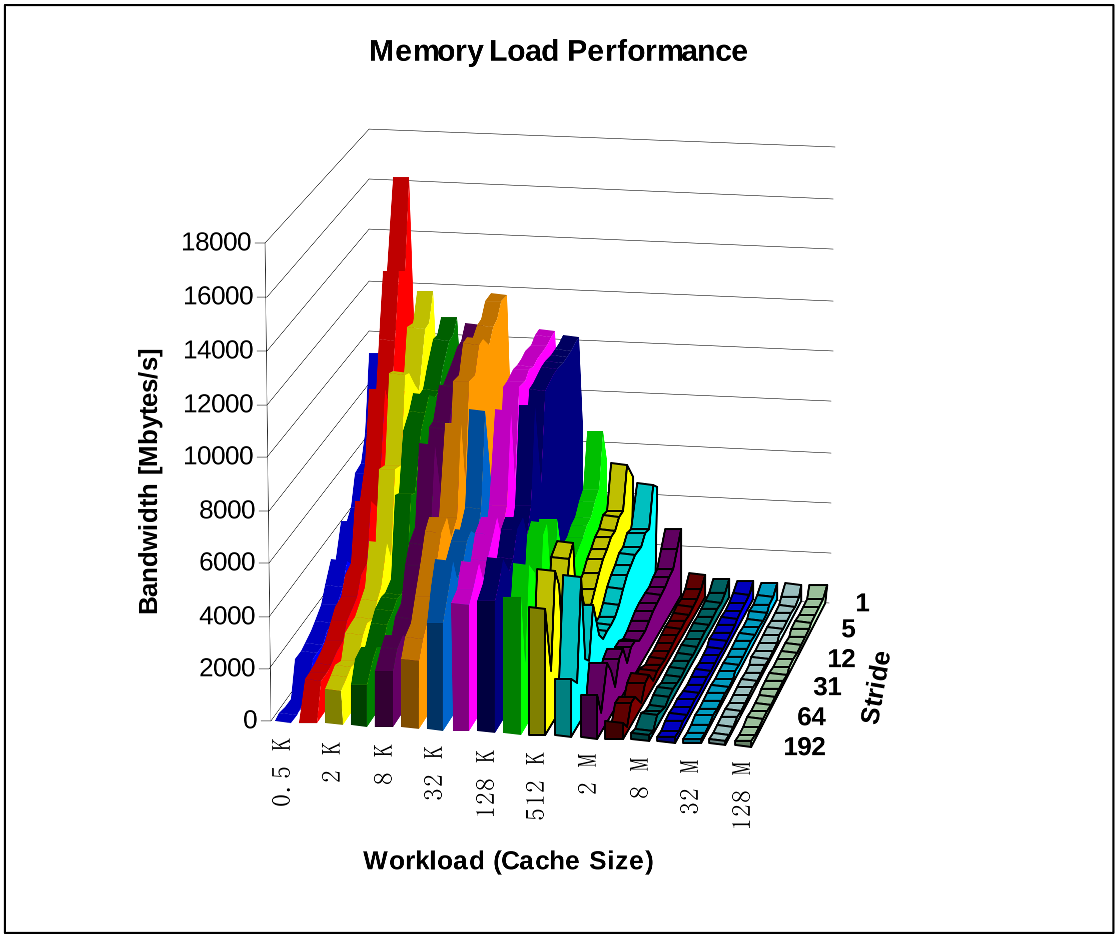

Figure 1

In figure 1, the X-Axis varies the working set parameter. This shows how the memory system hierarchy supports temporal locality, i.e. the effect of cache hits through reuse of recently accessed data.

The Y-Axis varies the access pattern for contiguous blocks and strides (even and uneven). The stride parameter shows how well caches and external stream logic help with read ahead and other means of improving bandwidth for accesses with spatial locality.
A slope of increasing performance marks the end of the access pattern axis toward lower strides. Its steepness indicates improved bandwidth for loads with contiguous accesses and accesses with small strides. A selection of even, odd, and prime strides permits to detect performance gains and losses due to a banked memory system.

The first cluster (1~16K) with the highest bandwidth is performance of L1 cache. It works much better and accesses 8 double word per clock cycle. The red bar indicates the performance at 18000MB/s. Also the L2 cache (16K~1M) works at a remarkable speed with 12000MB/s. The L3 cache reduced its bandwidth to about 4000MB/s. The load bandwidth of the main memory is nearly 100% faster for contiguous blocks but slower for strided access, which can be explained by the larger cache lines of the MIPS processor.

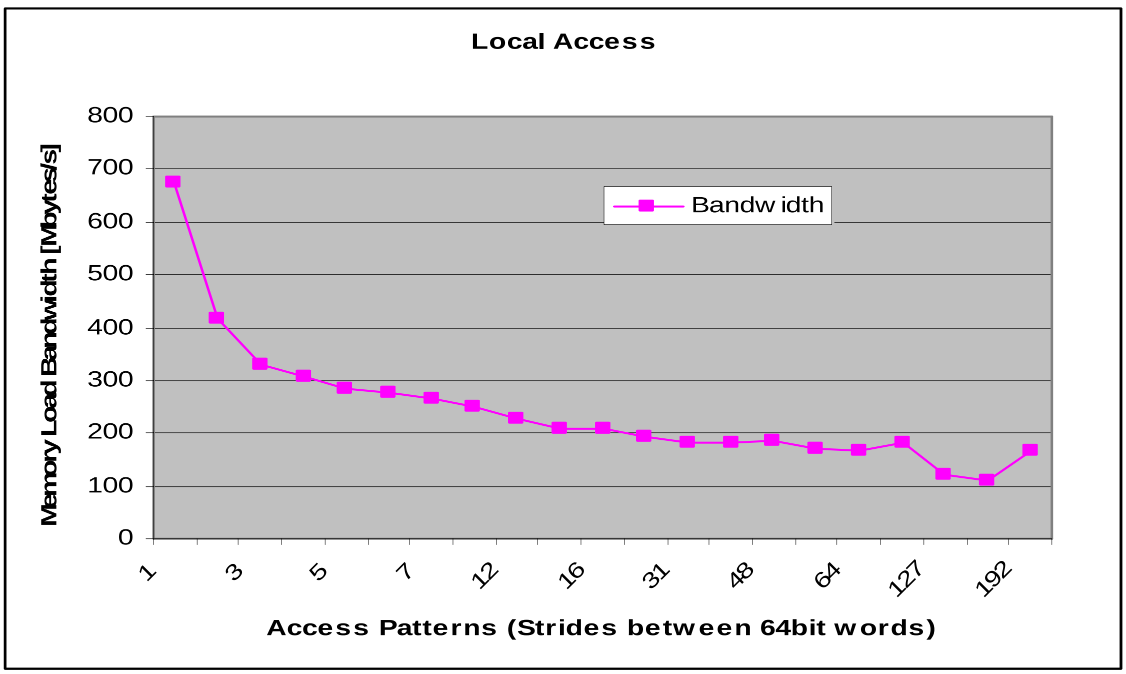

Figure 2

Figure 2 shows local access (main memory in the previous full graph) information. It is interesting that the SGI systems perform much better for contiguous loads. This can partly be explained by the large cache lines of the cache hierarchy which introduces a lot of overhead.

4.2 Mode 2: Load Copy (strided load and contiguous store)

To really reveal the pattern I separately compare the copy bandwidth in a multiprocessing scenario where either 1, 2, 4 or 8 processors copy data in the memory. For small working sets in the caches, the performance remains the same, as measurements prove. More interesting are the difference for large working sets in main memory. I not only use a simple copy as e.g. MCCalpins 'Stream Benchmark' but measure a gather copy stream where the processors read strided data and store it contiguously.

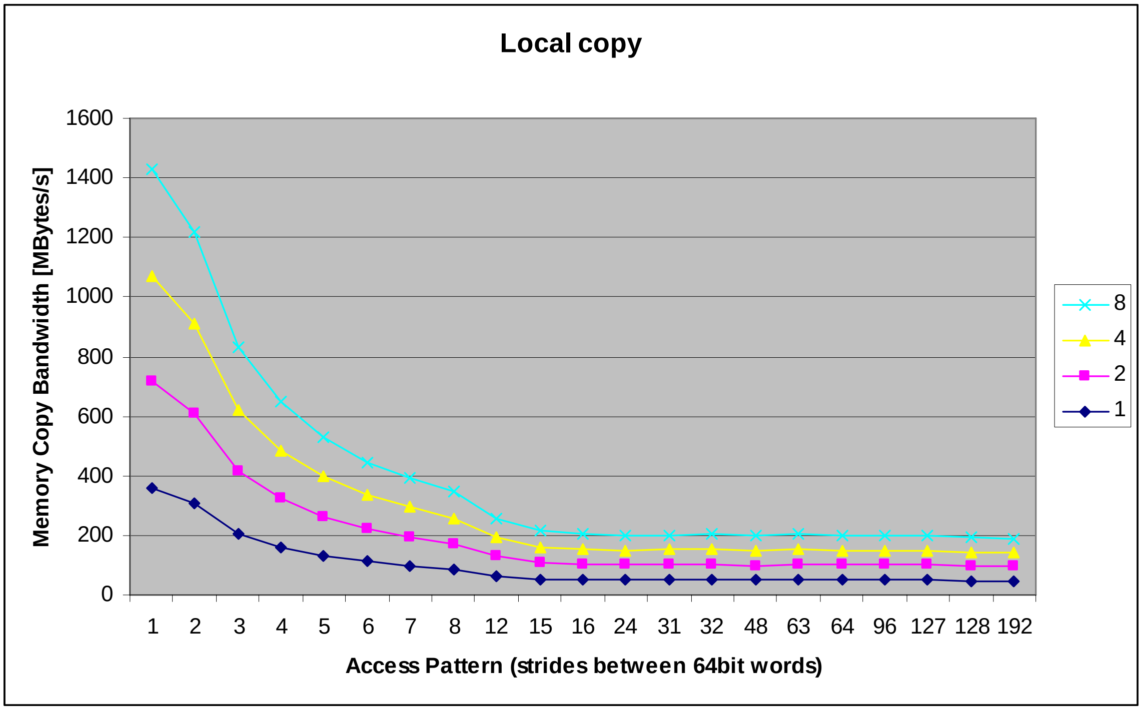

Figure 3

The CC-NUMA SGI Altix 4700 processors can copy at the same speed as one processor because each processor node has its own memory system. The performance for contiguous copies collapses on the node from 1400 MB/s to 400 MB/s. As it should be in a NUMA distributed memory machine, the performance then stays identical for all nodes.

4.3 Mode 3: Copy Store (contiguous load, strided store)

Figure 4 shows that remote in Parallel & Networking computers is fairly clear. A processor accesses data in a memory of another node over the network. In an SMP machine multiple processors access the memory over a host bus or a network. Remote means that one processor stores the data to a shared memory segment where the second processor accesses it. The CC-NUMA SGI performs at a better speed for remote memory accesses than for local accesses. NUMA in this case affects remote access bandwidth. However, as noticed from the figure, when the access are contiguous (strides = 1), local and remote access have about the same bandwidth.

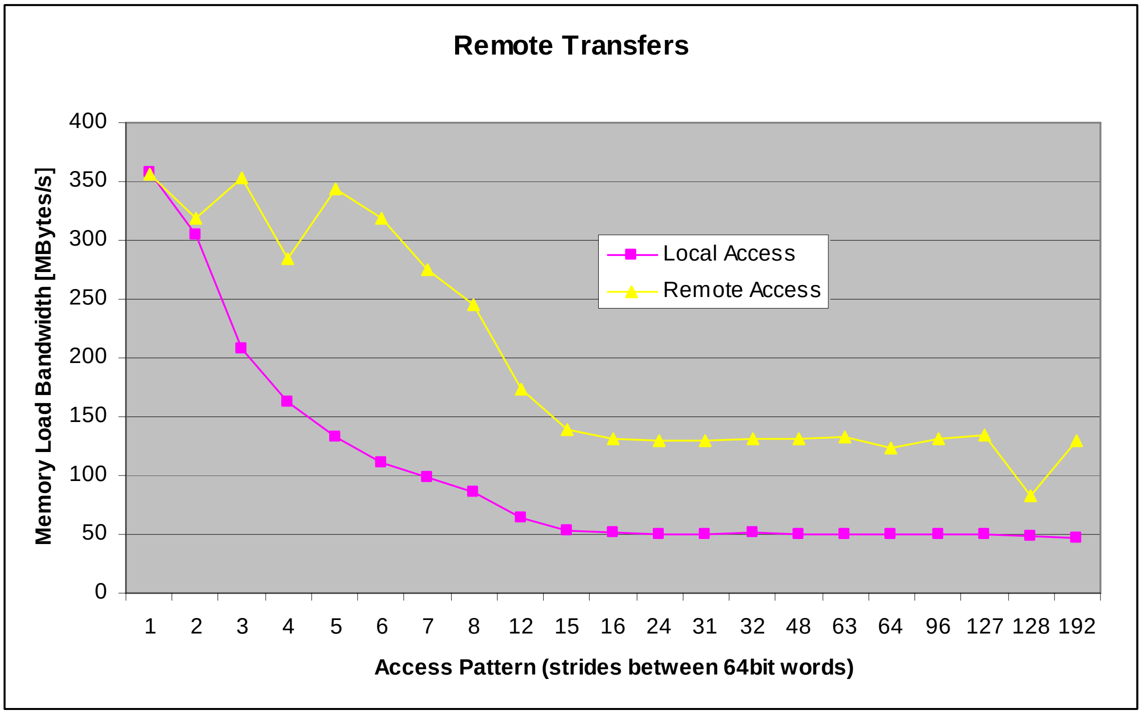

Figure 4

5 Conclusion

This project discusses about the memory system in SGI Altix 4700. As we can see from the results, the memory bandwidth decreases with the growth of stride steps. It has the best performance for contiguous memory accesses. On the other hand, the CC-NUMA SGI Altix 4700 processors can copy at the same speed as one processor because each processor node has its own memory system. The NUMA structure also determines a stable bandwidth for both local and remote memory accesses.