Reducing latency to volume visualization on PC cluster

PAULO ALEXANDRE BRESSAN*, MARCELO KNÖRICH ZUFFO
Laboratory of Integrated Systems (LSI) – Polytechnic School of University of São Paulo (EPUSP) – São Paulo – Brazil
{pbressan, mkzuffo}@lsi.usp.br

Abstract
Volume Visualization is an important tool in many scientific applications, requiring intensive processing and dealing with large amounts of data. Therefore the size of these data always exceeds the processing and visualization capacities of the conventional workstations. In this technical poster we present our working in progress about techniques of data prefetching and data prediction to volume visualization on low cost PC clusters. Our approach is based on the Distributed Shared Memory (DSM) paradigm [3], which one can greatly facilitate the parallel programming efforts on volume visualization applications.

Keywords: parallel volume visualization, PC cluster, DSM, data prefetching, data prediction.

1. Introduction
A wide range of volume visualization techniques and algorithms are available today, since volume visualization is embarrassedly parallel. Typically the parallel programming efforts in volume visualization focus on the partitioning of big volumetric datasets among processors, foreseeing the memory hierarchy in order to avoid data swapping, data redundancy and the need of intense synchronization among processors.

Moreover, there are many evidences that software DSM has the potential to provide a simple programming model for parallel computing on distributed memory systems without hardware support for shared memory. Besides, in clusters, inherent communications between processors, data sharing resulting in the parallel methods, is the main reason of inefficient of memory hierarchy.

This project is been development on a 6 computers PC cluster, named Polux, at LSI-EPUSP. Each computer is a Pentium III Xeon 1GHz, 2MB cache, 1GB RAM, Wildcat 4210 and Volume-Pro. The cluster network is Gigabit Ethernet.

In this technical poster, we discuss techniques of data prefetching and data prediction to reduce latency on parallel volume visualization.

2. Data Prefetching
In data prefetching [2], processors request in advance data that they will use later. This triggers a fetch to the memory system that brings the data to the local caches. Previous work by many researchers has conclusively shown that prefetching is highly effective.

However, data prefetching has several disadvantages when compared to data forwarding. First, prefetching requests may be issued too late to hide the access latency completely. Second, if multiple processors need the data, each one must generate a prefetch request, possibly resulting in a large instruction overhead in software-controlled data prefetching. Finally, prefetching increases network traffic and resource utilization.

3. Data Forwarding
Forwarding can hide the latency of communication-induced misses by having producer processors send data to the caches of potential consumer processors in advance[1]. Forwarding can hide the latency effectively, has low instruction overhead and uses few machine resources.

Data forwarding is appealing for three reasons. The first one is that the producer can forward the data as soon as it generates it, therefore maximizing the chances of hiding the communication latency. Secondly, if we forward when the data is produced, we only need to generate the address of the data once, possibly saving instructions over prefetching.

Of course, data forwarding has shortcomings. Unlike prefetching, forwarding is only effective for communication-induced cache misses. In addition, the compiler and hardware support for forwarding may be more complicated.

4. Conclusion
The project expects to increase performance of volume visualization applications on PC cluster, contributing to achieve interactivity and efficient.

5. References

* Project supported by Fapesp #99/12693-1 (Brazilian Research Foundation).