Leveraging an Innovative Fast Parallel Sorting Algorithm for Efficient Seismic Imaging

Mufeed AlMatar, Saleh Aldossary

Saudi Aramco

Abstract

Data sorting is an integral and essential part of seismic processing. The ability to sort seismic data traces in a specific domain (e.g. Common Image Gathers) is desirable for efficient imaging algorithms. The rapid increase in size of 3D seismic data volumes due to higher channel counts and greater azimuthal coverage, presents a significant challenge in data sorting. Increasingly, conventional sequential sorting algorithms are seen as inadequate and adversely impact overall seismic processing efficiency.

Depth imaging is increasingly used to better understand complex hydrocarbon traps and structures, and to display the subsurface in greater detail. Advances in computing have made depth imaging technology more affordable and it is therefore becoming the rule rather than the exception. In addition, a practical fast sorting algorithm capable of handling large amounts of data is needed in order to alleviate the problem further.

In this paper, we present an innovative parallel sorting algorithm that can sort terabytes of data in hours instead of days.

The algorithm utilizes only physical memory and minimizes disk usage. Unlike existing serial sorting algorithms that must utilize limited resources (memory and disk) on a single node, parallel sorting algorithms can take advantage of memory on all nodes and avoid use of disks entirely. Therefore, parallel sorting achieves far higher sorting speeds and fully utilizes computation resources. If the required total memory is smaller than the data, the algorithm is able to stream and sort the data in multiple passes.

This fast sorting parallel algorithm can sort 65TB of shot gathers into CMP gathers in less than 10 hours using 4 nodes of 256GB of memory, which is 3-4 times faster than conventional sorting algorithms. The performance of the algorithm depends primarily on the number of nodes and the available memory on each node. The effect of intranode communication is negligible on the sorting performance. Leveraging this innovative sorting algorithm has resulted in a more efficient seismic imaging workflow of large 3D data sets and better utilization of seismic HPC resources.