An Unstructured Mesh Partitioning Scheme for Efficiently Parallelizing an Explicit Time Domain Volume Integral Equation Solver

Ahmed Al-Jarro and Hakan Bağcı

Division of Computer, Electrical and Mathematical Sciences and Engineering
King Abdullah University of Science and Technology, Thuwal, 23955-6900, Saudi Arabia
ahmed.aljarro@kaust.edu.sa, hakan.bagci@kaust.edu.sa

Abstract: A highly scalable parallelization scheme for accelerating the explicit marching-on-in-time (MOT) based solution of the time domain volume integral equation (TDVIE) is proposed. The parallelization scheme ensures even distribution of (i) source fields and their temporal history, (ii) computation of the discretized spatio-temporal convolutions, and (iii) computation of finite differences approximating the spatial derivatives among the processors. All data exchange required in between processors is carefully designed to avoid global (collective) communications. Numerical results demonstrate that the proposed parallelized MOT-TDVIE solver scales linearly on the IBM Blue Gene/P platform for problems discretized using one million elements on up to 32,768 cores.

Keywords: Time Domain Volume Integral Equation, Explicit Marching-on-in-time Scheme, Distributed-Memory Parallelization, Unstructured Grid Partitioning

1. Introduction

It is well known that high computational costs associated with marching-on-in-time (MOT)-based time domain integral equation (TDIE) solvers have been limiting their widespread use in practical real-life electromagnetics problems [1-4]. Parallelization is one way of drastically reducing long computation times needed for the characterization of wave interactions on electrically large structures using MOT-TDIE solvers. In this work, a highly efficient and scalable parallelization scheme for accelerating the explicit MOT-based solution of the time-domain volume integral equation (TDVIE) [4] is presented.

Unlike classical MOT-TDIE solvers [1-3], the TDVIE solver [4] parallelized in this work expands unknown fields using zeroth order basis functions in space and first order interpolation functions in time, and uses finite differences to approximate the spatial derivatives. The TDVIE is tested using point matching in space and time. The stability of the time marching is maintained using a predictor-corrector algorithm. The simplicity of the solver, allows for a memory efficient implementation: No interaction matrices are stored and the discretized spatio-temporal convolutions are fully computed on the fly during time marching. The only storage is the temporal history of the source fields. Therefore, the parallelization scheme presented here mostly focuses on efficient distribution of the computational work and aims at achieving high scalability in computation times.

The proposed parallelization scheme is built upon the work detailed in [5, 6]. It ensures even distribution of (i) source fields and their temporal history, (ii) computation of the discretized spatio-temporal convolutions, and (iii) computation of finite differences approximating the spatial derivatives among the processors. It is noted here that the efficient inter process communication (IPC) required for the computation of the discretized convolutions, (ii), is achieved using a simple one-way pipeline communication strategy, so called “rotating tiles” paradigm as described in [6]. This scheme eliminates the need for global communications. The even distribution of the finite difference computations, (iii), without adversely affecting the efficiency of the discretized convolution computations is the focus of this

work and described in detail in this paper. It is further added that, (i), the even distribution of the source fields and their temporal history is an expected outcome of (ii) and (iii).

The parallelization scheme described in [5, 6] distributes equally the source fields and their temporal history among the processors without taking into account the physical locations of the source field samples. Then, the rotating tiles paradigm is applied to compute the tested fields efficiently without the need for global communications [6]. Even though this approach results in almost ideal scaling for the computation of discretized spatio-temporal convolutions, it does not optimize the halo-type communications needed for the computation of the spatial finite differences; and hence jeopardizes the overall scalability of the solver.

To this end, in this work, an alternative scheme to distribute the source/test points among the processors is proposed. The proposed scheme makes use of a graph-based partitioning method [7] to ensure that the total number of source/test points assigned to each processor is equal and the resulting number of test points that reside on the boundary of any two partitions is minimized. Numerical results demonstrate that parallelized MOT-TDVIE solver scales linearly on the IBM Blue Gene/P platform for problems discretized using one million elements on up to 32,768 cores.

2. Parallel Explicit MOT-TDVIE Solver

A. Time-Domain Volume Integral Equation (TDVIE)

The electric field, $\mathbf{E}(\mathbf{r},t)$, in the presence of a non-magnetic dielectric scaterrer residing in free space satisfies the TDVIE [4]:

$$\mathbf{E}(\mathbf{r},t)=\mathbf{E}_0(\mathbf{r},t)+\left[\nabla\nabla\cdot-\frac{\partial_t^2}{v^2}\right]\int_V d\mathbf{r}'\frac{\left[\varepsilon(\mathbf{r}')-\varepsilon_b\right]\mathbf{E}(\mathbf{r}',t')}{4\pi\varepsilon_b\left|\mathbf{r}-\mathbf{r}'\right|},\ \mathbf{r}\in V. \tag{1}$$

In (1), $\mathbf{E}_0(\mathbf{r},t)$ is the band-limited incident electric field, v is the speed of light in free space, $\varepsilon(\mathbf{r})$ and ε_b are permittivities of the scatterer and free space, respectively, V represents the volume of the scaterrer, $R=\left|\mathbf{r}-\mathbf{r}'\right|$ is the distance between the observation point and source points, $\mathbf{r}$ and $\mathbf{r}'$. To numerically solve the TDVIE (1), V is discretized using cubic elements, $\mathbf{E}(\mathbf{r}',t')$ is expanded using zeroth order basis functions in space and first order interpolation functions in time. The resulting equation is tested in space and time using point matching. Spatial derivative operator "$\nabla\nabla\cdot$" is approximated using finite differences. The differentiation in time, "∂_t^2", is taken into account through the use of a predictor-corrector algorithm. At the predictor step, spatio-temporal samples of $\mathbf{E}(\mathbf{r},t)$ (i.e., tested field samples) are computed using the samples of the discretized spatio-temporal convolution and $\mathbf{E}_0(\mathbf{r},t)$. At the corrector step, tested field samples from the predictor step, where a less accurate approximation has to be used for evaluating "∂_t^2" are "corrected" using a more accurate approximation of "∂_t^2". At the end of each time step, tested field samples are stored as source field samples to be used in the computation of the spatio-temporal convolution in the next time step. Note that in this algorithm, spatial finite difference evaluations and corrector updates are spatially local operations while computation of the spatio-temporal convolution is global. The predictor-corrector algorithm is demonstrated to increase the accuracy and stability of the solution while maintaining the explicitness the MOT scheme [4].

B. Parallelization

The computational cost of the explicit MOT-TDVIE solver briefly described in Section 2.A is high due to the need to compute the discretized spatio-temporal convolutions. To enable the use of this solver in characterizing transient wave interactions on electrically large structures, several parallelization schemes have been proposed in [5, 6]. In this work, the parallelization scheme presented in [6], which efficiently parallelizes the computation of the discretized convolutions via the use of a one-way pipeline communication strategy termed "rotating tiles", is further enhanced. This technique eliminates the need for the otherwise collective communications that rely on globally executed routines using summation

operations [5]. It is further added that, in [5] and [6], the partitioning is carried out without taking into account the physical locations of source points, i.e. using a structured grid for their distribution (partitioning). Each processor stores the time history of the source field samples that belong to the partition assigned to it. The test point partitioning is initially (at the beginning of the rotation) is the same as the source point partitioning, however, the test point partitions are rotated among the processors. When a processor receives a test point partition, it adds the contribution from its source field samples to the tested fields sampled at test points of the received partition and passes the partition to its neighboring processor. At the end of a full rotation all contributions to the tested fields are computed at a given time step without the need for global collective communications.

Even though this scheme results in almost ideal scaling of the solver, it does not automatically optimize the distribution of the source field samples; and therefore the computations of the discretized spatio-temporal convolutions, or the communications required for the computation of the spatial finite differences. This unoptimized distribution of the source field samples can hinder the scalability of the solver, especially when the number of processors is high and the number of tested field points in each partition is relatively low. In this work, the rotating-tiles scheme is enhanced using an unstructured graph partitioning technique [7] to enforce the even distribution of the source field samples, as well as to minimize the halo-type communications, as described next.

During the unstructured partitioning, first, the underlying mesh discretization of the volume, V, is converted into a graph. Then, this very dense graph is coarsened down from its original representation, but obviously of an identical volume with smaller number of nodes. This is achieved by successively collapsing together a maximum set of adjacent pairs of nodes. This final coarser graph is then processed with the objective to ensure that source/test points are evenly distributed among partitions while the overall number of elements that reside on the boundary of any two partitions is minimized. This objective is usually referred to as the edge-cut [7]. This process is executed iteratively by moving vertices between partitions in order to better improve the quality of the final partitioning solution [7]. It should be noted here that during this iterative process, priority is given to the portion of the graph that is in close proximity to the partitions boundaries [7].

The unstructured partitioning of the underlying mesh discretization of the volume, V, can be invoked through METIS stand-alone programs or application programming interface (API) [7]. The METIS program takes as input the element-node array of the mesh to be partitioned and computes a target size of partitions, N_p, for both its elements and its nodes. The unstructured graph partitioning step is a static task, i.e., it is not modified during the MOT loop, and therefore it is carried out only once before the marching starts. The solver then makes use of the elements partitioning, where several other pre-processing tasks are needed before marching with the aim to optimize the repeated data communications that take place during the MOT loop. These tasks involve generation of look-up tables and data maps that are used for the rotating-tile and unstructured halo-type communications.

3. Numerical Results

The scalability tests of the proposed parallelized TDVIE solver are performed on Shaheen, an IBM Blue Gene/P platform located at the King Abdullah University of Science and Technology (KAUST) Supercomputing Laboratory (KSL). Shaheen consists of 16 racks, each of which contains 1024 quad-core compute nodes. Each node running at 850MHz is equipped with 4GB of memory. Scalability of the proposed solver is tested for two problems where the scatterer is discretized using 571595 and 1097240 elements. Figs. 1 and 2 present the strong scaling of the proposed scheme and compares it to ideal scalability. The strong scaling, for a given problem size, is defined as $S_{N_p} = \log_2(T_{N_p} / T_{ref})$. Here, N_p represents the number of processors used in the simulations, and T_{ref} and T_{N_p} are the total times recorded for the simulations executed on 64 and N_p number of processors.

It is clear from Figs. 1 and 2 that the proposed scheme scales extremely well on the IBM Blue Gene/P platform. This clearly demonstrates that the unstructured partitioning technique ensures the even

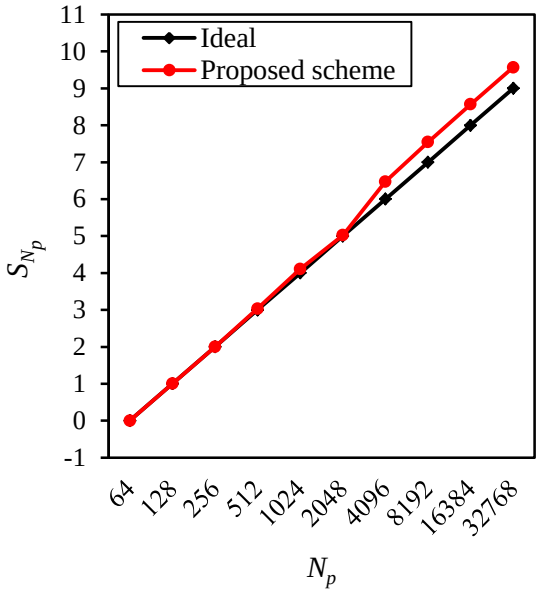

Fig. 1. Scalability of the proposed scheme for a problem with 571595 elements.

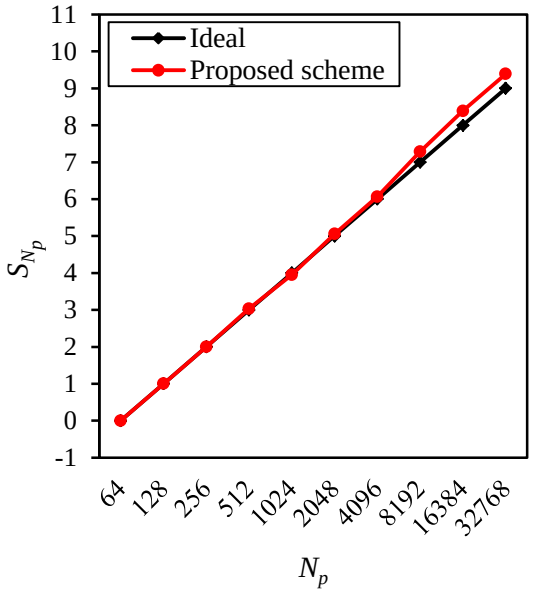

Fig. 2. Scalability of the proposed scheme for a problem with 1097240 elements.

distribution of the overall computation load among the processors. It should be noted here that as N_p is increased, for a fixed problem size, the communicated message sizes reduce linearly to the level where they are small enough to fit within the CPU cache. At this point, the average latency of data access (read/write) will be that of the cache access, which is much shorter than the latency of main memory access. As a result, and given that S_{N_p} is being measured against T_{ref} with 64 cores, for the cases where access to only small memory segments per process is needed, the super-linear scaling effect is observed. This super-linear scalability is evidently observed for $N_p \geq 2096$ and $N_p \geq 4096$ in Figs. 1 and 2, respectively.

4. Conclusions

A highly scalable parallel TDVIE solver is described. The proposed parallelization scheme makes use of an unstructured graph partitioning technique to evenly distribute the computations required for the evaluation of discretized spatio-temporal convolutions and the spatial finite differences. All data exchange required in between processors is carefully designed to avoid global (collective) communications. Numerical results demonstrate the (super-linear) scalability of the solver when executed on the IBM Blue Gene/P platform using up to $32,768$ cores.

References

[1] B. Shanker, A. A. Ergin, M. Lu, and E. Michielssen, "Fast analysis of transient electromagnetic scattering phenomena using the multilevel plane wave time domain algorithm," *IEEE Trans. Antennas Propag.*, vol. 51, no. 3, pp. 628-643, 2003.

[2] A. E. Yilmaz, J. M. Jin, and E. Michielssen, "Time domain adaptive integral method for surface integral equations," *IEEE Trans. Antennas Propag.*, vol. 52, no. 10, pp. 2692-2708, 2004.

[3] H Bagci, A. E. Yilmaz, and E. Michielssen, "An FFT-accelerated time-domain multiconductor transmission line simulator," *IEEE Trans. Electromagn. Compat.*, vol. 52, no. 1, pp. 199-214, 2010.

[4] A. Al-Jarro, M. A. Salem, H. Bagci, T. M. Benson, P. Sewell and A. Vukovic, "Explicit solution of the time domain volume integral equation using a predictor-corrector scheme", *IEEE Trans. Antennas Propag*., vol. 60, no. 11, pp. 5203-5214, 2012.
[5] A. Al-Jarro, M. Cheeseman, and H. Bagci, "Distributed-memory parallelization of an explicit time-domain volume integral equation solver on Blue Gene/P," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 27, no. 2, pp. 132-144, 2012.
[6] A. Al-Jarro, M. Cheeseman, and H. Bagci, "A distributed-memory parallelization of the explicit time-domain volume integral equation solver using a rotating tiles paradigm," in Proc. *Applied Computational Electromagnetics Society (ACES) Conference*, Ohio, USA, April, 2012.
[7] G. Karypis, and V. Kumar, "A fast and high quality multilevel scheme for partitioning irregular graphs," *SIAM J. Sci. Comput*., vol. 20, no. 1, pp. 359-392, 1998.