Integrated Parallelization of Computation and Visualization for Large-scale Weather Applications

Preeti Malakar
Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India

I. INTRODUCTION

Critical applications like cyclone tracking and earthquake modeling require simultaneous high-performance simulations and online visualization for timely analysis. These simulations involve large-scale computations and generate large amount of data. Faster simulations and simultaneous visualization enable scientists provide real-time guidance to decision makers. However, resource constraints like limited storage and slow networks can limit the effectiveness of *on-the-fly* remote visualization. In this work, we have developed an integrated user-driven and automated steering framework INST that simultaneously performs numerical simulations and efficient online remote visualization of critical weather applications in resource-constrained environments. Such a simulation-visualization model can enable geographically distributed scientists to collaboratively analyze the visualization and provide expert opinion on the occurrence of critical events.

INST considers application dynamics like the criticality of the application and resource dynamics like the storage space and network bandwidth to adapt various application and resource parameters like simulation resolution and the frequency of visualization. The framework also provides the user control over various application parameters like region of interest and simulation resolution. We have devised an adaptive algorithm to reduce the lag between the simulation and visualization times. We propose a strategy for concurrent execution of multiple high-resolution nested simulations, which improves the overall performance. We describe our work in more detail in the following sections and provide few experimental results.

II. SIMULTANEOUS SIMULATION AND ONLINE VISUALIZATION

High simulation rates on modern-day processors combined with high I/O bandwidth can lead to rapid accumulation of data at the simulation site. We have shown in our work that it is important to consider these resource constraints for continuous simulation and smooth visualization [1]. Maximum temporal resolution can be achieved by increasing the frequency of output of data, but this can decrease the simulation rate due to increase in number of writes to the disk. This can lead to rapid consumption of storage and eventually stalling of the simulations. On the other hand, decreasing the output frequency can increase the simulation rate, but will result in visualization of fewer frames. Thus we fulfil these contradictory objectives of maximum simulation rate and maximum temporal resolution by formulating our problem as an optimization problem. Since we want the best possible throughput of the simulation in spite of the resource constraints, we express the objective of our optimization problem as $minimize\ t$ where t is the execution time to solve a time step and the constraints are related to limited disk space available, I/O bandwidth, network bandwidth, continuous simulation and visualization and minimum progress rate of simulation. The solution to this constrained linear program gives the number of processors for simulation and the frequency of output for visualization. This approach results in 30% higher simulation rate and 25-50% lesser storage consumption than a naïve greedy approach.

III. A STEERING FRAMEWORK FOR WEATHER SIMULATIONS

Online visualization can allow the user to give feedback to the ongoing simulation. Such a computational steering framework for high-performance simulations of critical applications needs to take into account both the steering inputs of the scientists and the criticality needs of the application. We have shown in [2] how our adaptive steering framework INST analyzes the combined effect of user-driven steering with automatic tuning of application parameters based on resource constraints and the criticality needs of the application like minimum progress rate of simulation to determine the final parameters for the simulations. For example, if the user-specified values of output interval (OI) and minimum progress rate of simulation cannot be simultaneously satisfied due to resource constraints, INST overrides the user-specified OI.

IV. EFFICIENT ONLINE VISUALIZATION

Another challenge in online remote visualization of critical weather applications is to minimize the *lag* between the time when the simulation produces an output frame and the time when the frame is visualized. It is important to reduce the lag so that the scientists can get *on-the-fly* view of the simulation and timely measures can be taken for critical applications. This lag depends on the output frequency determined by INST and the network bandwidth. We have developed algorithms to minimize the simulation-visualization lag and concurrently visualize important events in the simulation [3]. Our algorithms adapt to the available resource parameters and the number of pending frames to be sent to the visualization site by transferring a significant subset of frames. The *auto-clustering* frame selection algorithm clusters the pending frames based on modified k-means algorithm for temporal clustering. Additionally, the *adaptive* algorithm dynamically decides the information content in the frames to be transferred, depending on the user-specified lag bound. Using experiments with different network configurations, we find that our adaptive algorithm strikes a good balance in reducing the lag and visualizing most representative frames, with up to 37% larger representativeness than a naïve greedy approach.

V. Simulation of Multiple Regions of Interest

Accurate and timely prediction of weather phenomena, such as hurricanes and flash floods, require high-fidelity compute-intensive simulations of multiple finer regions of interest within a coarse simulation domain. Current weather applications execute these nested simulations sequentially using all the available processors, which is sub-optimal due to their sub-linear scalability. In [4], we present a strategy for parallel execution of multiple nested domain simulations based on partitioning the 2-D processor grid into disjoint rectangular regions associated with each domain. We propose a novel combination of performance prediction, processor allocation methods and topology-aware mapping of the regions on torus interconnects. Experiments on IBM Blue Gene systems using WRF show that the proposed strategies result in performance improvement of up to 33% with topology-oblivious mapping and up to additional 7% with topology-aware mapping over the default sequential strategy.

VI. Experiments and Results

We elaborate few of the obtained results using weather simulation for a critical weather application, namely, tracking of cyclone *Aila*. We use Weather Research and Forecasting Model (WRF) for simulation and VisIt for visualization. With decreasing pressure at the *eye* of the cyclone, InSt refines the weather simulation for more accurate output. We show results for *intra-country* setup with bandwidth of 40 Mbps between the simulation and visualization sites. The *intra-country* simulations were performed in Centre for Development of Advance Computing (C-DAC), Bangalore, India. Visualization was carried out on a workstation in Indian Institute of Science (IISc) with a Intel(R) Pentium(R) 4 CPU 3.40 GHz and an NVIDIA graphics card GeForce 7800 GTX.

Efficient online visualization: Figure 1 illustrates the simulation-visualization lag for our frame selection algorithms. The *all* curve corresponds to the default policy of sending all the frames. The frame selection algorithms most-recent, auto-clustering and adaptive reduce the lag by 80% on average. The *adaptive* algorithm with lag bound of 15 minutes performs slightly better because it can adapt to the current lag and dynamically decide the information content in the representative frames to be sent.

Concurrent execution of nested domains: Figure 2 compares the execution times of four nest domains of sizes 394x418, 232x202, 232x256 and 313x337. When they are executed sequentially on 1024 BG/L cores, the total time taken for a time-step is 1.1 seconds, as shown by the first bar. In our parallel strategy, the four domains are allocated 432, 144, 168 and 280 cores according to our performance model and partitioning strategy. It can be seen that this concurrent strategy outperforms the default sequential strategy by 36%. This is a result of sub-linear scalability of the nested domains.

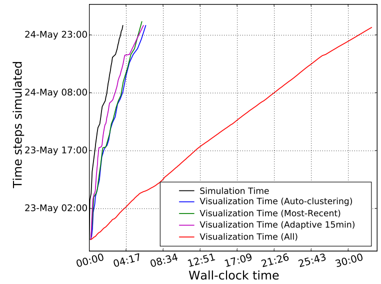

Fig. 1. for intra-country simulations

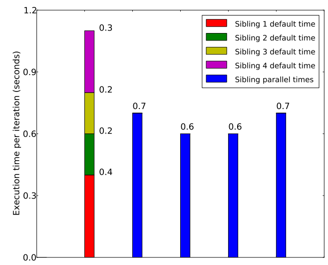

Fig. 2. Execution times of four nested domains on 1024 BG/L cores

VII. Conclusions

As we enter the exascale era, it is important to efficiently perform high-performance scientific simulations and visualization, inspite of the different growth rates in the computation speeds and the memory and network bandwidths. In this work, we presented an adaptive integrated steering framework for simulation and visualization of critical applications like cyclone tracking across various resource configurations. Our framework adapts to the resource characteristics and enables a scientist to steer simulation, who may be unaware of the resource constraints. We propose an adaptive algorithm to perform efficient online visualization. We also improve the performance of simulations with multiple high-resolution nested simulations for multiple regions of interest.

References

[1] P. Malakar, V. Natarajan, and S. Vadhiyar, "An Adaptive Framework for Simulation and Online Remote Visualization of Critical Climate Applications in Resource-constrained Environments," in *Proceedings of the ACM/IEEE Conference on Supercomputing*, 2010.

[2] ——, "InSt: An Integrated Steering Framework for Critical Weather Applications," in *ICCS 2011: Proceedings of the International Conference on Computational Science*, 2011.

[3] ——, "End-to-end Adaptive Framework for Efficient Online Visualization of Critical Weather Applications," (submitted).

[4] P. Malakar and et al., "A Divide and Conquer Strategy for Scaling Weather Simulations with Multiple Regions of Interest," in *Proceedings of the ACM/IEEE Conference on Supercomputing*, 2012, (to appear).