Topic: F (Data management, analysis, and assimilation)

Oral presentation: Yes

A Case for Real-time Comparative Analytics

John Wu (LBNL), Scott Klasky (ORNL), CS Chang (PPPL)

Growing Data from Experiments and Simulations

With increased automation and faster data acquisition systems, large-scale scientific experiments are generating ever-increasing amount of data. At the same time, science projects, such as, LHC (Large Hadron Collider) and ITER (formerly, International Thermonuclear Experimental Reactor), involve more and more scientists from all continents of the world. In most of these projects, a team of scientists is present at the central site to monitor the progress of the on-going data collection, adjust the control settings, and prevent catastrophic events; while most others access the data remotely after an experiment has concluded. Allowing these remote users to conduct their analysis operations in real time as the data is being collected would enable more scientists to provide feedback to the execution of the shared experiments and increase scientific output. Often there is a strong desire for (near) real-time participation, however there are significant roadblocks to such a remote participation. For example, the computer network might not be fast enough to transfer a sufficient amount of data for a meaningful analysis, the analysis operations might take too long to provide timely feedback, or there might not be convenient software framework to compose the analysis procedures to conduct the necessary analyses and provide timely feedback.

There is also a vibrant community of physicists studying fusion through advanced simulations, such as the XGC1 code developed under the leadership from PPPL. XGC1 is a full-f gyrokinetic ion-electron particle code, specifically designed for edge plasmas. It is a massively parallel guiding center ion neoclassical particle code. Neutral particles, from wall-recycling and gas puffing, are simulated together with the plasma particles, providing a self-consistent solution incorporating plasma-material interactions and atomic physics effects. It uses an unstructured triangular mesh for evaluation of the physical observables. In a recent planning process, the lead scientist estimated that a 24-hour run of XGC1 code could produce more than 1 petabytes (PB) of output. Simulation programs such as XGC1 are producing quantities that are easily comparable with the experimental observations. We believe a good way to utilize the data from experiment and simulation at the same time is to perform comparative analytics. Clearly, this is not the only way to make use of the using the massive amounts of data, but it is a use case with immediate benefit to the large experiments such as ITER. In the near-term, we see a key challenge of data management research is to support real-time comparative analytics.

Toward Real-time Comparative Analytics

On a large experiment, data collection is typically centralized at one site, while the computing power is distributed worldwide at the participating research institutions. There are many obstacles to achieving real-time comparative analytics. For example, accessing the necessary data over the wide-area network is challenging. Locating the available storage and computing resources for such a real-time workflow would be also challenging because most of current high-performance computers (HPC) are running in batch mode. At the same time, developing the necessary analysis functions for the evolving variety of HPC platforms is yet another challenge. Clearly, considerable research efforts are still needed to address these research topics. However, there are causes to be optimistic that we could overcome these obstacles. Next, we briefly outline a recent research effort that took a few bold steps toward this goal of real-time comparative analytics. The name of this project is ICEE, which is shorthand for International Collaboration Framework for Extreme-scale Experiments

The work on ICEE was originally motivated by the collaborative analysis needs on the data from the Korea Superconducting Tokamak Advanced Research (KSTAR) facility. KSTAR is a magnetic confinement fusion device that operates in pulses, known as shots. Each run of an experiment includes many such shots. During a shot, many diagnostic instruments collect gigabytes of information that needs to be stored, analyzed and shared among scientists. As in most pulse-based experiments, there are three types of analyses: in-shot analysis, between-shot analysis and post-run analysis. A shot in KSTAR can last from a few seconds to 300 seconds for a full-mode production. Due to the relatively short duration of each shot, the analyses performed during a shot used to concentrate on providing safety critical functions such as monitoring the health of an experiment to make sure there is no instability occurred. If instability is observed, the experiment needs to be aborted in milliseconds to avoid damage to the experimental facility. As more computing resource become available, scientists are exploring the possibility of adjusting a number of experimental parameters during a shot to improve the stability of the fusion plasma. However, it is very challenging to move the amount of data produced in a 300-second pulse (could be a few terabytes in the future) and provide analysis results in seconds. Between two consecutive shots, there is a longer time period for between-shot analysis. Even in this case, the analysis time is in minutes, which is still challenging for a widely distributed workflow.

A significant number of analysis algorithms have been developed over the years and most of them are currently used in post-run analyses. Thus, there is a large pool of analysis algorithms suitable for studying the data from the KSTAR diagnostic devices. In addition, the computer network for transmitting the data to and from South Korea has been upgraded to 10Gpbs lines, which is sufficient to transfer a large portion of the data recorded by the KSTAR diagnostic devices in seconds. The ICEE software framework is designed to connect these analysis algorithms and data management capabilities to provide an efficient workflow composition and execution engine.

The ICEE software framework leverages the state-of-art networking and data management technologies, and enables users to design and construct near real-time distributed data-processing workflows over wide-area networks. A key technology in this system is the ability to transfer a large amount of data from memory to memory without writing any data records to disk. This technique significantly reduces the data access latency. At SC14 conference, four different teams successfully applied this framework on four distinct applications: fusion blob detection, Electron Cyclotron Emission Imaging (ECEI) processing, microscopy image analysis, and materials design. Next, we describe the fusion blob detection workflow.

To detect the blobs in fusion plasma, the diagnostic image and simulation data are shipped from their sources to a computer center such as NERSC or OLCF for analysis. To minimize the amount of data shipped, we may limit the portion of image shipped for the analysis operations. For example, the regions of the image that obviously does not contain a blob do not need to be analyzed and therefore do not need to be shipped. Once the necessary data is transferred, the analysis program applies an outlier detection algorithm to determine the blobs. To reduce the analysis time, we parallelize the detection algorithm to take advantage of parallel computing resources available at NERSC and OLCF. In the demonstration workflow, the analysis results are shipped to an iOS based monitoring device, which could be located anywhere with network accesses.

In short, by parallelizing the analysis operations, identifying the appropriate computing resources, locating the most relevant data records for a particular analysis, and reducing the data communication cost, we have demonstrated that it was possible to complete comparative analysis in seconds, which is faster enough for many interactive use cases. Future version of comparative analytics systems will need to support high-data rates from both experiments and simulation programs. In addition, we anticipate more advanced analysis workflows that need to dynamically adjust themselves in response to the analysis output, network conditions, available computing resources and so on. Clearly, a lot more research is needed, but based on the preliminary work, we are hopeful that real-time comparative analytics is an achievable goal.