International Collaboration Framework for Extreme Scale Experiments (ICEE)

Annual Project Report – 2nd year

October 1, 2014

Project period of September 1, 2013 through August 31, 2014

Investigators

John Wu[1], CS Chang[2], Scott Klasky[3],

Alex Sim[1], Jong Youl Choi[3], Michael Churchill[2]

Project Website: **http://sdm.lbl.gov/icee**

TABLE OF CONTENTS

[1] Lawrence Berkeley National Laboratory

[2] Princeton Plasma Physics Laboratory

[3] Oak Ridge National Laboratory

1 Summary

Large-scale scientific exploration in domains such as high-energy physics, fusion, and climate are based on international collaborations. As these collaborations produce more and more data at faster and faster rate, the existing workflow management systems are hard pressed to keep pace. A necessary solution is to process, analyze, summarize and reduce the data before it reaches the relatively slow disk storage system, a process known as *in-transit* processing (or in-flight analysis). The ICEE project aims to dramatically increase the data handling capability of collaborative workflow systems by leveraging the popular *in transit* processing system known as ADIOS, and integrating this with FastBit to provide effective selective data accesses. These new features will contribute to a new collaborative system named ICEE that aims at significantly improving the data flow management for distributed workflows. The improved data processing capability will enable large international projects to make near real-time collaborative decisions.

In the previous year, we have develop an initial prototype workflow based on ADIOS and have exercised system within a single data center. In the second year of this project, we continued our development of the prototype ICEE system and demonstrated its functionality between KSTAR and ORNL and between LBNL and ORNL. We have demonstrated the functionality of the workflow system by two different workflows, one involving the Electron Cyclotron Emission Imaging (ECEI) data from the KSTAR project, and the other involving synthetic diagnostic data from the XGC modeling of a fusion tokamak at PPPL. The ORNL project team has made a trip to KSTAR to demonstrate our prototype and collaboration. To ensure that the wide-area network connections are properly configured for this *in transient* work, we have worked with ESnet and NISN engineers to diagnose network connections between KSTAR and ESnet. These activities contributed to more advanced wide-area workflow system development. In the remainder of this report, we provide a brief outline of the key activities in the past year.

2 Wide-Area Workflow Development

2.1 Overview of ICEE system

Figure 1 shows the overview of the ICEE system design with our key components of ADIOS and FastBit. After data acquisition is completed from the diagnostic sensors, the raw experimental data is managed by ICEE server, and pulled into ICEE system for common activities such as building indexing of variables and recording provenance information. ICEE server, then, provides data to the remote users for analysis through wide-area networks. A key aspect of ICEE system is that, in order to minimize disk I/O overheads, operations are performed in memory as much as possible throughout the end-to-end data processing workflow. Data in memory is separated by logical time steps (or called as shots in fusion experiments) but shares similar structures between time steps. Remote clients will access the data with direct connections to an ICEE Server or via a data hub. Data communications are based on an event transport middleware, called EVPath, and integrated into ADIOS for wide-area data exchanges. Figure 2 is a demonstrated workflow based on the ECEI stream data processing during the NGNS PI meeting with a portable job monitoring tool.

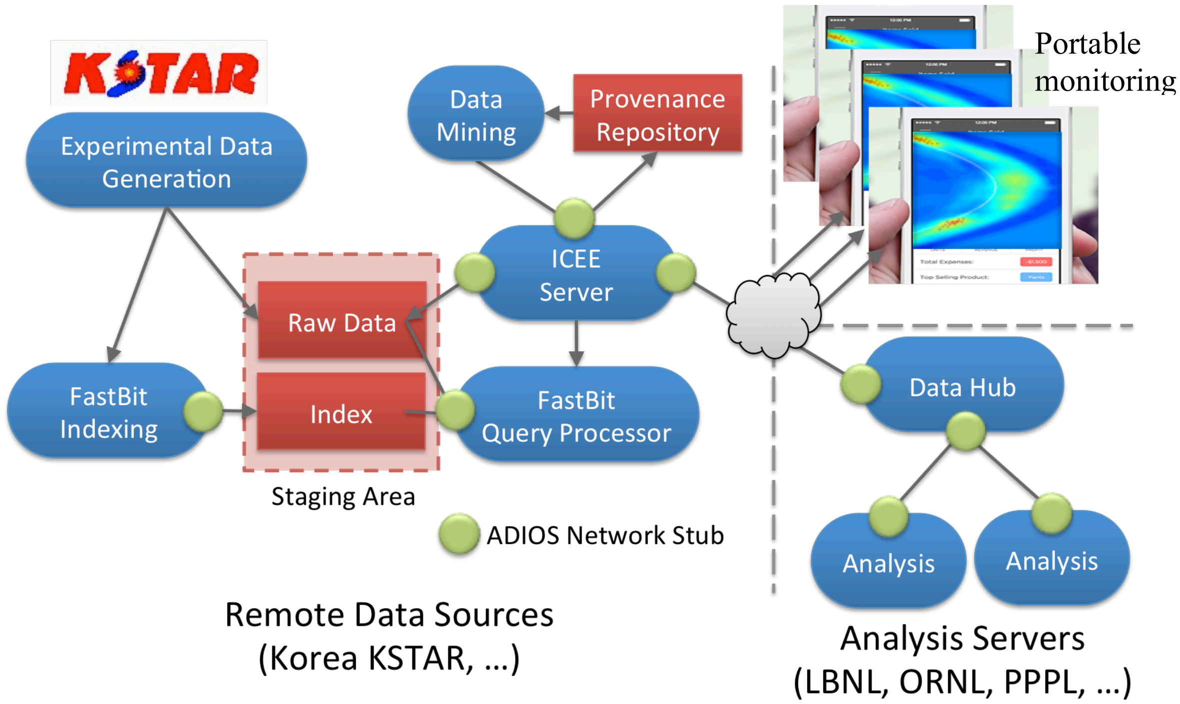

Figure 1. Overview of ICEE system.

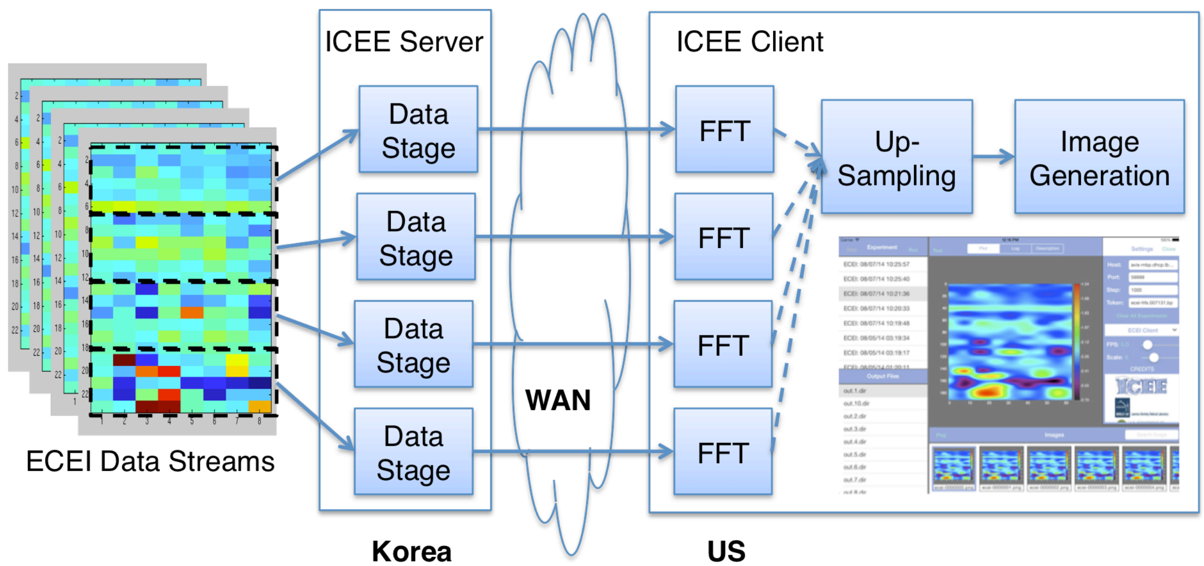

Figure 2. ECEI stream data processing workflow and ICEE system

2.2 Blobs detection workflow

Blobs, shown in Figure 3, are intermittent bursts of plasma particle and energy in the edge of tokamak driven by nonlinear turbulence. Blobs could have deleterious effect on fusion performance because they transport heat and particles away from the confined plasma and lead to increased levels of impurities, bypassing control mechanisms. The blob physics needs to be understood to predict ITER performance.

Experimental diagnostics measure blobs on ~μs timescales with significant radial resolution. Typical blob lifetimes are measured to be 10's μs. The diagnostics measuring blobs generate large amounts of data in short periods of time. Therefore, blob detection is a good example of large scale data analysis that require extremely high performance.

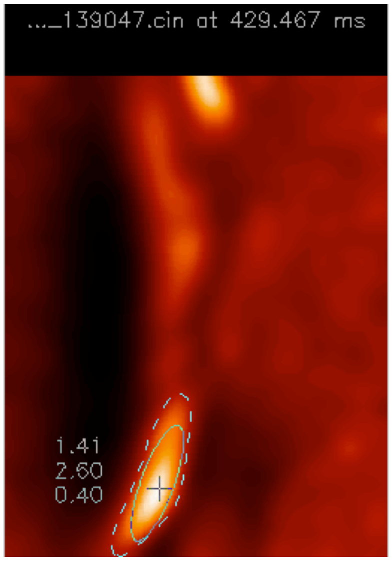

Figure 3. An example of a blob

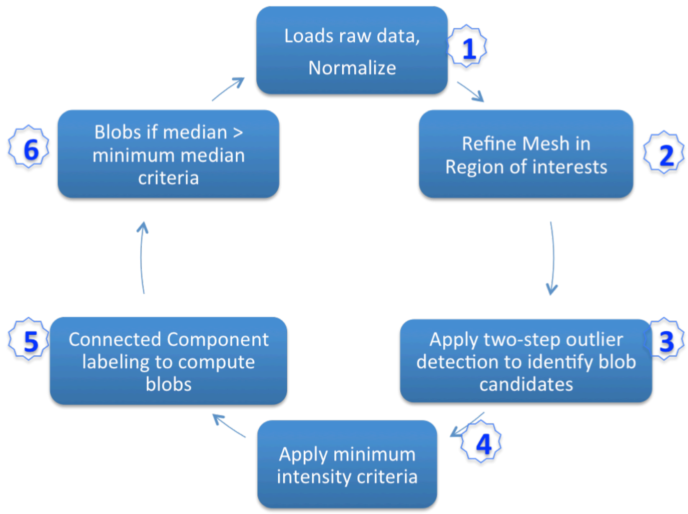

Figure 4. Blob detection algorithm

To detect blobs efficiently, we have developed a parallel outlier detection algorithm, shown in Figure 4. In addition to the performance enhanced compared to the earlier implementation in matlab, this new method is also able to capture blobs at the edge of the region of interests, which were frequently missed by the earlier detection algorithm. This is a first research work presenting a high performance blob detection approach that completes in a few milliseconds. An example of the blob detection algorithm results is shown in Figure 5. We have examined the parallel performance of the blob detection algorithm. The results shown in Figure 6 indicate that our approach scale well and effectively take advantage of large parallel machines to improve the blob detection performance.

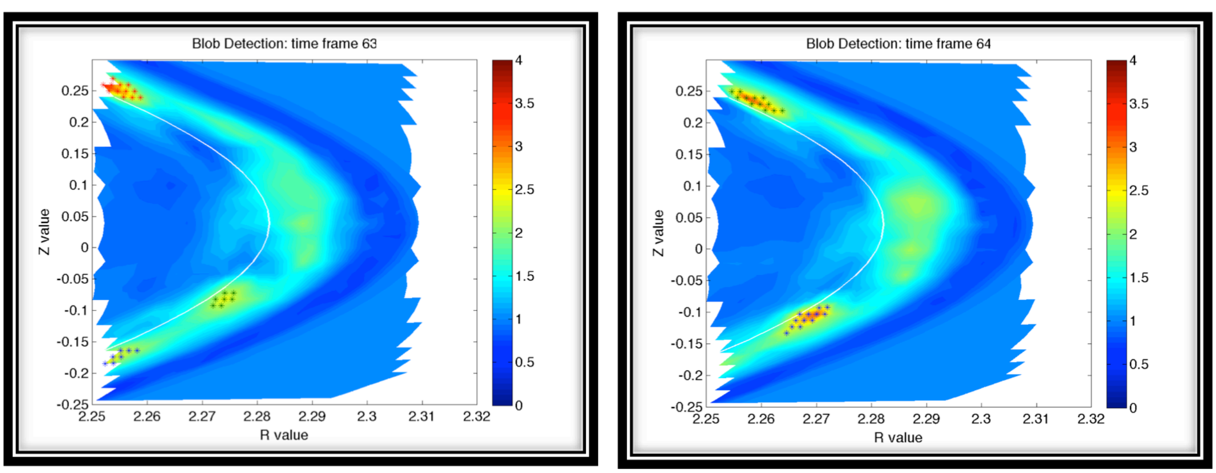

Figure 5. An example of the blob detection in two continuous time frames on synthetic image data. Dot clusters mark the grid points in regions where blobs were detected by the algorithm

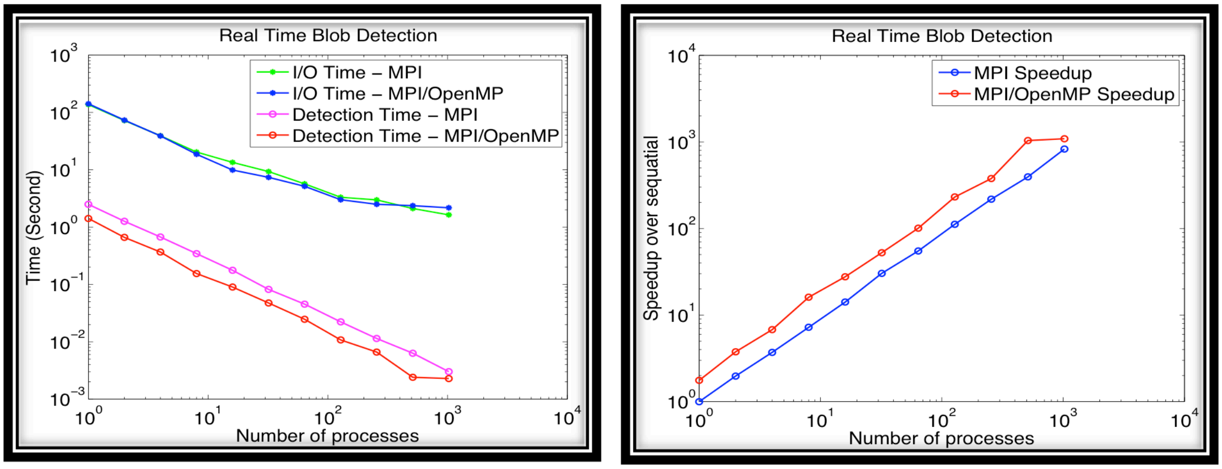

Figure 6. Blob detection time, I/O time and speedup varying number of processes

3 Network performance diagnostics

We have measured the network performance between our Korean experimental site and the US to accommodate our network performance needs over wide-area network. National Institute of Supercomputing and Networking (NISN) and ESnet manage the national networks in Korea and the US respectively, and we have worked with network engineers in both countries to diagnose the network performance.

Recent data throughput testing from Korea to NERSC in the US over 10Gbps shared network link was achieved about 450MB/sec (3.6 Gbps), shown in Figure 7, transferring 260GB. Current experimental data reading rate of 4GB every 10 seconds (3.2 Gbps) from the KSTAR processing system can be accommodated within the current 3.6 Gbps throughput performance from Korea to the US.

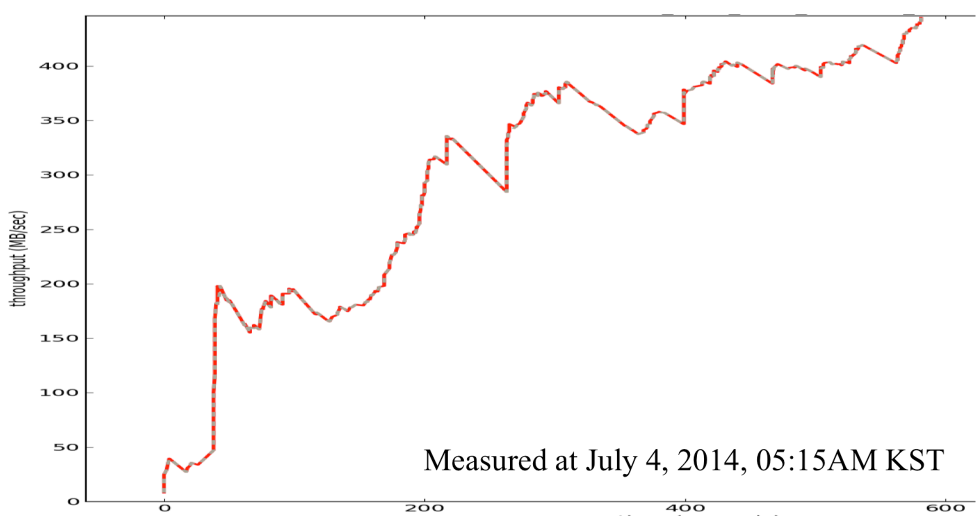

Figure 7. Data throughput from Korea to the US

4 KSTAR Collaboration

KSTAR (Korean Superconducting Tokamak Advanced Research) device is an advanced fusion device located at National Fusion Research Institute (NFRI) of Korea to study long-pulse aspects of magnetic confinement fusion as in ITER. Large-scale data will be produced from the KSTAR experiment. It has a large number of international participants that could significantly benefit from the real-time accesses to the diagnostic data collected during the operation of the device. ICEE (International Collaboration Framework for Extreme-scale Experiments) is a project motivated by this need for large-scale data collaboration. Now that we have demonstrated the capability of our *in transient* workflow system, we plan the following tasks for further collaborations with KSTAR.

1. Work with the network engineers at NFRI, NISN and ESnet to connect the in transient workflow system with the archived data at KSTAR. This type of post-processing workflow is similar to the workflows we have been exercising recently, but is a necessary step to prepare for the main task described next.
2. Demonstrate the stream processing of data in the wide-area network with ADIOS on KSTAR's ECEI data.
3. Explore the option of developing a second workflow based on another diagnostics.

5 Plans for Year 3

- ICEE system supporting workflows that allows dynamic additions and modifications
- Supporting multiple data sources
- Optimize data movements
- Continue collaborate with the KSTAR data management team, NISN and ESnet networking team
- Connect the monitoring tool with distributed workflow through ADIOS API