

Established capability for near-real time networked analysis of big KSTAR data at NERSC (PPPL, ORNL, ESnet, NERSC, KSTAR, KISTI)

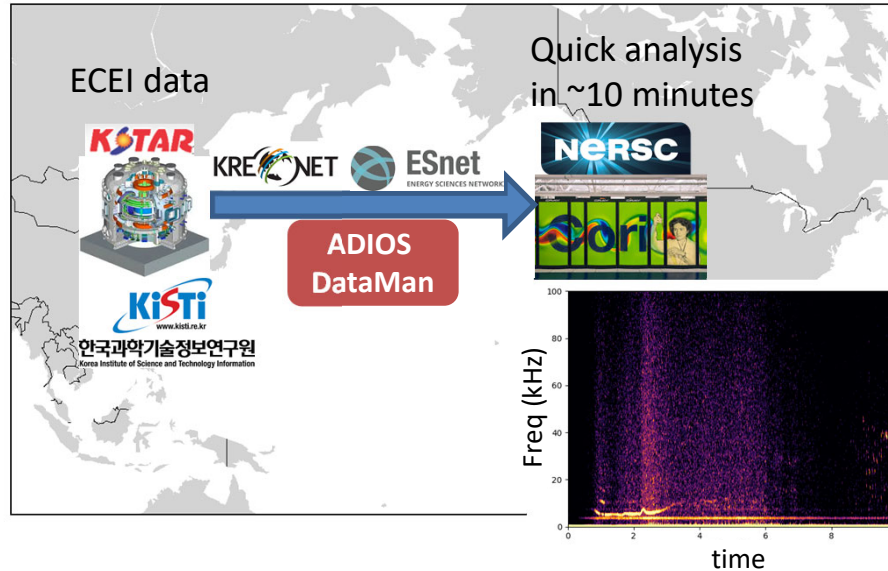
FES Highlight

Objectives

- Research and develop a streaming workflow framework, to enable near-real-time streaming analysis of KSTAR data on a US HPC
- Allow the framework to adopt ML/AI algorithms to enable adaptive near-real-time analysis on large data streams

Impact

- Created a framework to enable US fusion researchers to have broader and faster access to the KSTAR data, enabling
 - Faster analysis of data
 - Faster and autonomous utilization of ML/AI algorithms for incoming data
 - More informed steering of experiment
 - Quicker utilization of US HPC for KSTAR collaboration



Accomplishments

- Created end-to-end Python framework DELTA¹, streams data using ADIOS DataMan over WAN (at rates > 4 Gbps), asynchronously processes on multiple workers with MPI multi-threading
- Applied to KSTAR streaming data to NERSC Cori. Reduces time for an ECEi analysis from 12 hours on single-process to 10 minutes on 6 Cori nodes.
- Implemented deep convolutional neural networks for working with multi-scale fusion data, e.g. ECEi, for recognizing events of interest.²
- On-going: improve “adaptive” nature of data stream: adaptive compression at KSTAR source