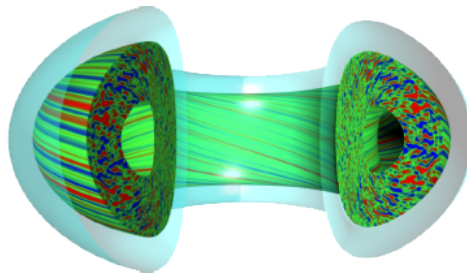


# 4th IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis



Contribution ID: 6

Type: **Regular Oral**

## Near real-time streaming analysis of big fusion data

*Friday, December 3, 2021 1:50 PM (15 minutes)*

While experiments on fusion plasmas produce high-dimensional data time series with ever increasing magnitude and velocity, turn-around times for analysis of this data have not kept up. For example, many data analysis tasks are often performed in a manual, ad-hoc manner some time after an experiment. In this article we introduce the DELTA framework that facilitates near real-time streaming analysis of big and fast fusion data. By streaming measurement data from fusion experiments to a high-performance compute center, DELTA allows computationally expensive data analysis tasks to be performed in between plasma pulses. This article describes the modular and expandable software architecture of DELTA and presents performance benchmarks of individual components as well as an example workflow. Focusing on a streaming analysis workflow where ECEi data measured at KSTAR on NERSC's supercomputer we routinely observe data transfer rates of about 500 Megabyte per second. At NERSC, a demanding turbulence analysis workflow effectively utilizes multiple nodes and graphical processing units and executes in under 5 minutes. We further discuss how DELTA uses modern database systems and container orchestration services to provide web-based real-time data visualization. For the case of ECEi data we demonstrate how data visualizations can be augmented with outputs from machine learning models. By providing session leaders and physics operators results of higher order data analysis using live visualizations may make more informed decisions on how to configure the machine for the next shot.

### Country or International Organisation

United States of America

### Affiliation

Princeton Plasma Physics Laboratory

**Primary authors:** KUBE, Ralph (Princeton Plasma Physics Laboratory); Dr CHURCHILL, Michael (Princeton Plasma Physics Laboratory); Dr CHANG, CS (Princeton Plasma Physics Laboratory); Dr CHOI, Jong (Oak Ridge National Laboratory); Dr WANG, Jason (Oak Ridge National Laboratory); Dr KLASKY, Scott (Oak Ridge National Laboratory); Dr STEPHEY, Laurie (Lawrence Berkeley National Laboratory); Dr DART, Eli (Lawrence Berkeley National Laboratory); Dr CHOI, Minjun

**Presenter:** KUBE, Ralph (Princeton Plasma Physics Laboratory)

**Session Classification:** Friday 3 Dec

**Track Classification:** Big data