



## Developing A Real Time Processing System for HERA

Paul La Plante<sup>\*(1)</sup>, Peter K. G. Williams<sup>(2)</sup>, Joshua S. Dillon<sup>(1)</sup>, the HERA Collaboration

(1) Department of Astronomy, University of California, Berkeley, CA

(2) Harvard-Smithsonian Center for Astrophysics, Cambridge, MA

### 1 Extended Abstract

The Hydrogen Epoch of Reionization Array (HERA) is an experiment endeavoring to measure the 21 cm signal from Cosmic Dawn and the Epoch of Reionization, the portion of the universe's history when the first luminous sources formed nearly 13 Gyr ago. When fully constructed, HERA is projected to generate more than 50 TB of data each night of observation. Storing these raw data indefinitely on site is infeasible due to the significant storage requirements, and transporting data is made difficult by limited bandwidth. Instead, calibration, analysis, and the production of science-grade data products will be performed in real time before discarding the raw data. Developing the infrastructure required for this processing provides a unique set of challenges, and monitoring the processing to quickly identify unexpected behavior is critical for ensuring successful scientific operations of the array.

We present here details about the Real Time Processing (RTP) system of the HERA experiment. The RTP system is responsible for the calibration and analysis routines performed after data is generated by the correlator. It also includes the data transfer management subsystem, which transports output products to other clusters for post-processing analysis and long-term storage. In addition to these primary purposes, the RTP system also provides valuable insight into the general health of the array. For example, the calibration step of the analysis pipeline can help identify array elements that have undergone mechanical or electronic failures, so that the construction team on the ground can address issues in a timely fashion. In addition, identifying radio frequency interference (RFI) events and relaying it back to the site team helps ensure that future data meets required levels of data quality. Accordingly, the RTP system is tied into the successful operation of the array to a great extent. The infrastructure underpinning the RTP system may be of general use to other astronomy experiments that rely on automated pipelines for analysis and data management, and future radio astronomy experiments such as the SKA.