

# A Spectral Line Data Processing Pipeline for SKA using DALiuGE

Ning Ouyang, Shuoying Lu, Leping Lin

## Abstract

With the construction of the Square Kilometre Array (SKA) project, the world's largest radio telescope, a new era in the detection of neutral hydrogen spectral lines is about to be ushered in, and the SKA spectral lines present tremendous opportunities and unprecedented data processing challenges. In this paper, we propose a fast processing pipeline for spectral data based on the Data Activated Liu Graph Engine (DALiuGE) and compare and analyze it with the Message Passing Interface (MPI). The proposed pipeline is proved to be more stable and efficient after several comparison experiments. The scalability and execution efficiency performance of SKA spectral line data fast processing pipeline using DALiuGE for data volume changes have been verified, and it is believed that DALiuGE can be a solution for more scientific tasks facing massive data processing in SKA Science Regional Center.

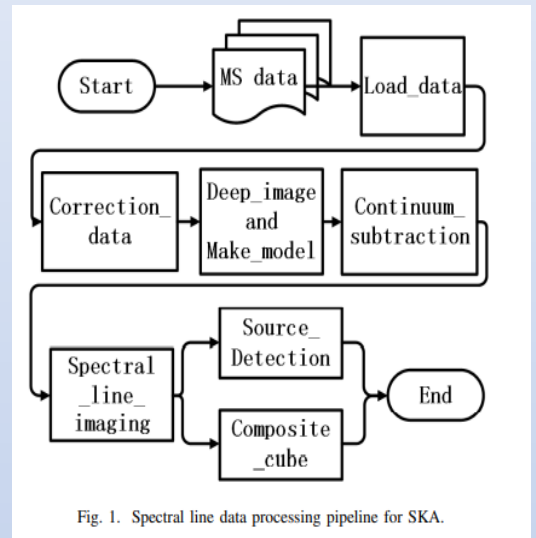


Fig. 1. Spectral line data processing pipeline for SKA.

## Pipeline design and parallel implementation

Referring to the existing mainstream spectral line data processing pipeline, this paper proposes a SKA spectral line data processing pipeline based on a distributed execution framework. It consists of imaging, source search, and composite cube image, as shown in Fig 1.

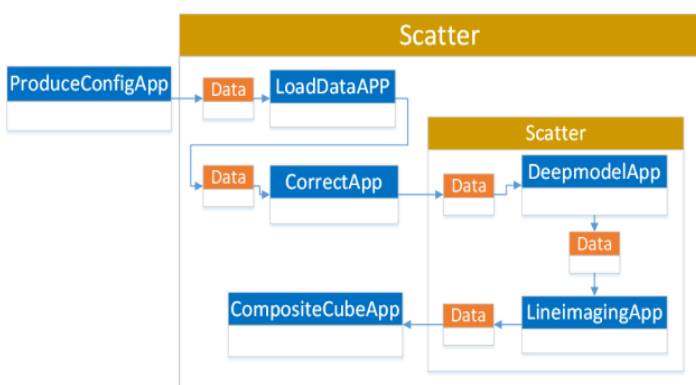


Fig. 2. Logic diagram of SKA spectral line data processing pipeline

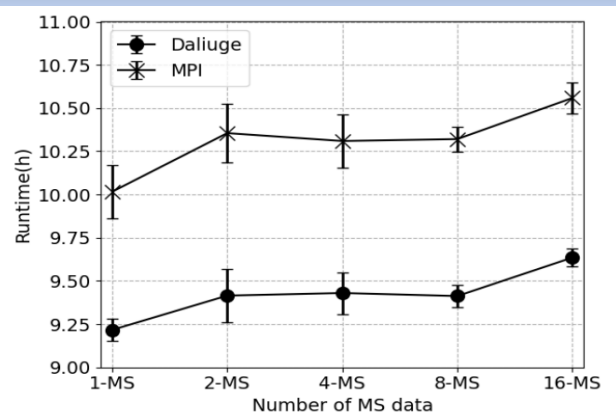


Fig. 3. Time comparison between DALiuGE version and MPI parallel version of ska 21cm spectral line data processing pipeline

## Conclusion

We applied the DALiuGE distributed execution framework to the SKA 21cm spectral line data processing pipeline, and optimized the entire pipeline for high-performance distributed computing, improving the efficiency of high-performance computing clusters and shortening the working time. The experiment of SKA 21CM spectral line data processing pipeline based on distributed execution framework provides a suitable development direction for DALiuGE.