Contribution ID: **59**   Type: **not specified**

ESFRI Project Presentation - SKA

Tuesday, 13 December 2016 09:00 (20 minutes)

The Square Kilometre Array (SKA) has a very demanding data management, storage and processing challenge. In this talk I will concentrate on the latter stages of the analysis pipeline which will be managed by the Science Data Processor (SKA-SDP) element of the SKA (part of the observatory infrastructure) and the SKA regional centres. The tiered model adopted by the SKA is similar to that used by CERN but has additional challenges not least of which is the data volume. The SDP element ingests data at up to 1.5 TBytes/s and averaged over a period of days must process these data into science-ready data products. The first SKA-SDP challenge is that some analysis must be performed as quickly as possible with strict latency requirements with further iterative computationally expensive processing requiring a net aggregate I/O bandwidth of order 10 TBytes/s. Data management for the SKA-SDP is a major challenge and in this talk I will discuss the architecture that the SKA-SDP is currently considering which includes a data-driven execution framework to help optimise data placement and movement between memory and several layers of persistent storage. The SKA-SDP processing stage will produce about 1 PByte of data products per day. These will then be distributed to SKA Regional Centres where further processing to produce secondary data products and other science extraction on the data will occur. At this stage interaction with science products from other observatories is essential. I will briefly discuss some of the work flows required, likely requirements for transfer to and preservation at the regional centres and how this will interact with the observatory.

Presenter: Dr ALEXANDER, Paul (Head of Astrophysics, Cavendish Laboratory, University of Cambridge)