

Platforms, Services, Products and Competences for Data Processing Centres

PRODUCTS

ALTEC is working to develop a set of products both to manage and process space and non space data and to manage complex large infrastructure using big data technologies. ALTEC products are reusable components ready to set up new processing pipeline and software infrastructure.

- **ASDP** (ALTEC Space Data Processing) is a processing framework providing services to manage workflow and processing resources, orchestrate processing and track data products.
- **TECSEL2** is a time series data analytics and data management software implemented with big data technologies.
- **AGILA** (ALTEC Ground Infrastructure Logs Analysis) is a tool to manage and mine logs of complex infrastructure in order to analyze them jointly to find recurring patterns and predict anomalies.
- **CONSOLE LOG** is a tool to record the console activities and manage the information flows between different operators (also working as activities schedule).

PLATFORMS

ALTEC have designed and operated several infrastructures: Gaia DPCT is one of the most important. It is a very **large and reliable hardware and software infrastructure**, sized to manage up 1PB of data, with **fast and secure data access**. The main features of ALTEC data processing Infrastructures can be summarized as follows:

- Scalable at hardware and software level
- Open and vendor independent
- Based on state-of-the-art IT technologies
- Integrated with external HPC and HTC infrastructure

ALTEC is working to implement its operational **cloud infrastructure** to be used cross projects. The reference cloud paradigm is PaaS and the ALTEC products will be some bricks of the platform. Based on the ALTEC cloud infrastructure, several **data exploitation platforms** could be built up in the next months to support ALTEC client communities. In 2016 ALTEC has been certified according to the requirements of **ISO/IEC 27001:2013** for service provision related to the utilization of mission operations and data processing centers.

COMPETENCES

ALTEC has competences to design and implement **complex software system** providing functionalities to manage all data processing stages in a ground segment. The most important competences are:

- **Big Data Management**: several years of experience gathered in the huge data store management is reusable to organize and access large data store through relational database, NoSQL database and distributed and block filesystem.
- **Data Processing**: capability to develop distributed processing framework and ability to use the most spread big data processing framework.
- **Data Analytics**: multidimensional analysis, full and partial cross-correlation, machine learning, classification with R, IDL, Hadoop, Spark etc.
- **Big Data Architecture**: several big data clusters dedicated to research projects and technologies experiments
- **Traditional and Cloud Infrastructure**: traditional infrastructures designed to maintain and use one PB of data (Gaia DPCT infrastructure); cloud infrastructure based implemented and operated in research project. ALTEC uses external platforms such as the PICO big data platform and the Marconi HPC platform. both hosted at CINECA.

SERVICES

ALTEC has an important experience in the **design and operations of scientific systems**. ALTEC has been involved in the operations of the Gaia mission since the commissioning phase and it has contributed to the first Gaia Data Release. ALTEC team is **supporting scientists along all mission phases** to bring their algorithms in real running systems. ALTEC leads the Gaia DPCT Operations Team.

The **ALTEC** Ground Infrastructure Team **manages and monitors the operational infrastructures** continuously and if required the **24x7** proactive monitoring services could be activated.