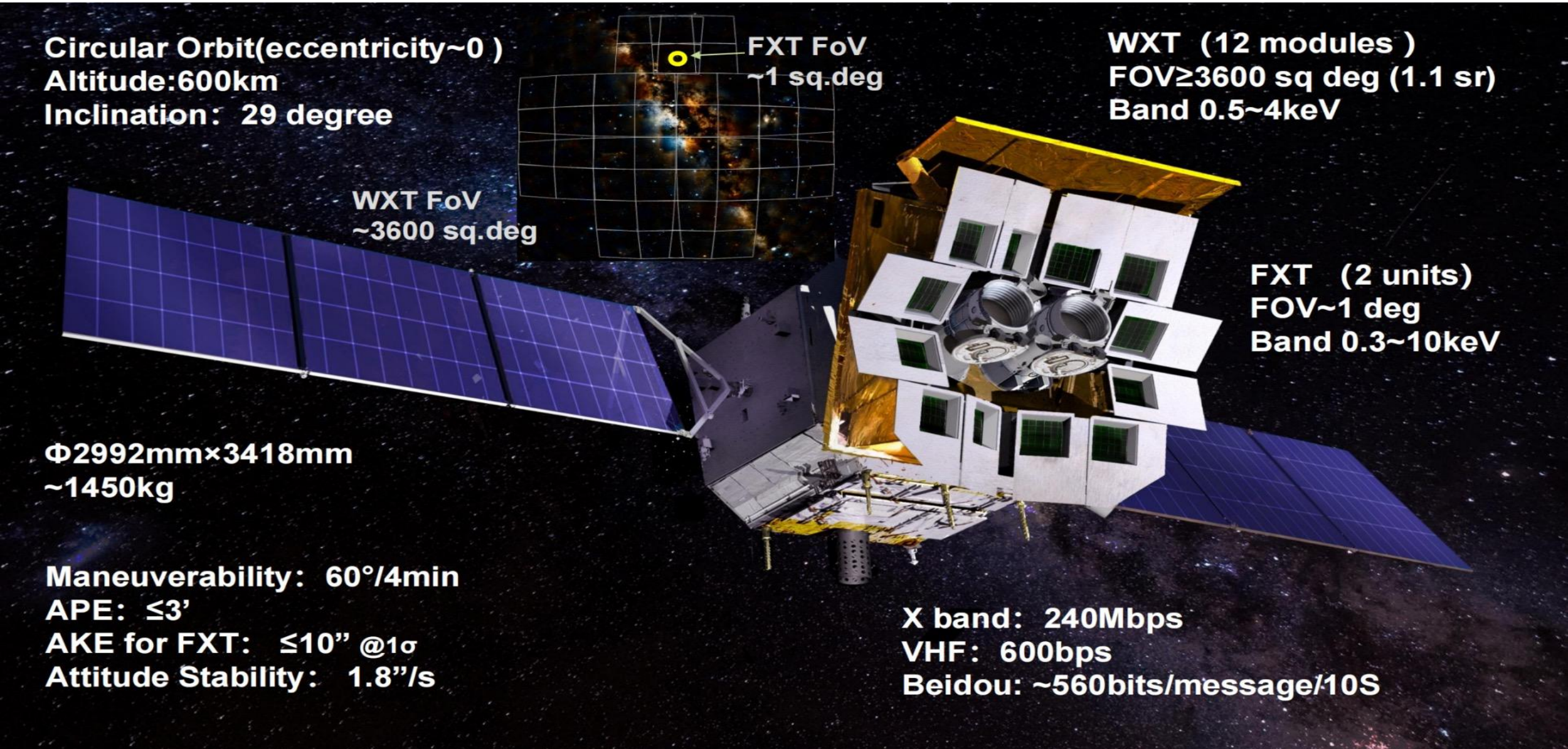
Experiences and **Lessons** learned from EP scientific workflow

Zhen Zhang on behalf of the China-VO team

Einstein Probe



Multi-Organizations

NAOC

The national astronomical observatories (NAOC) is responsible for the WXT data analysis softwares and data archives.

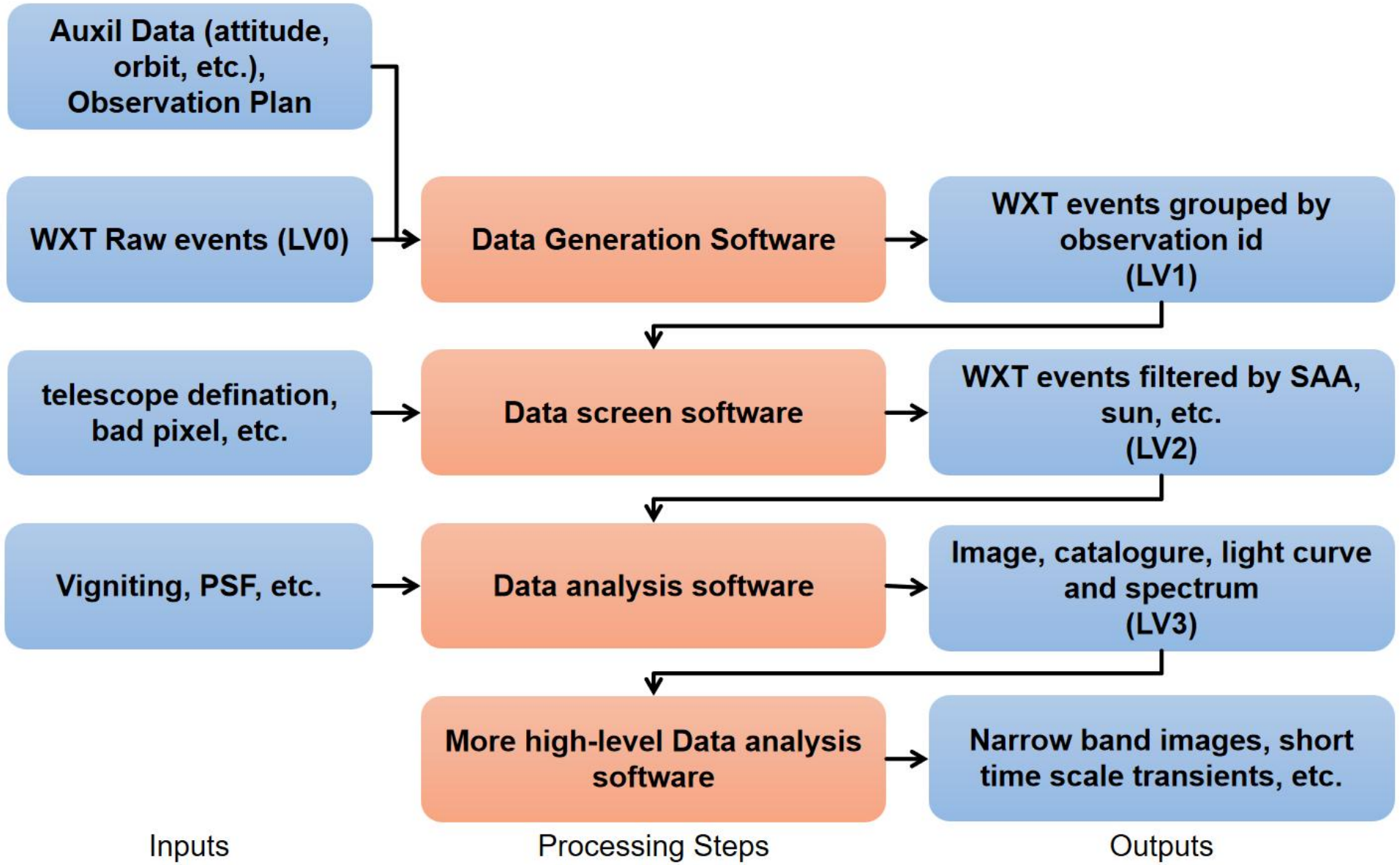
NSSC

The national space science center (NSSC) is responsible for data conversion and transfer.

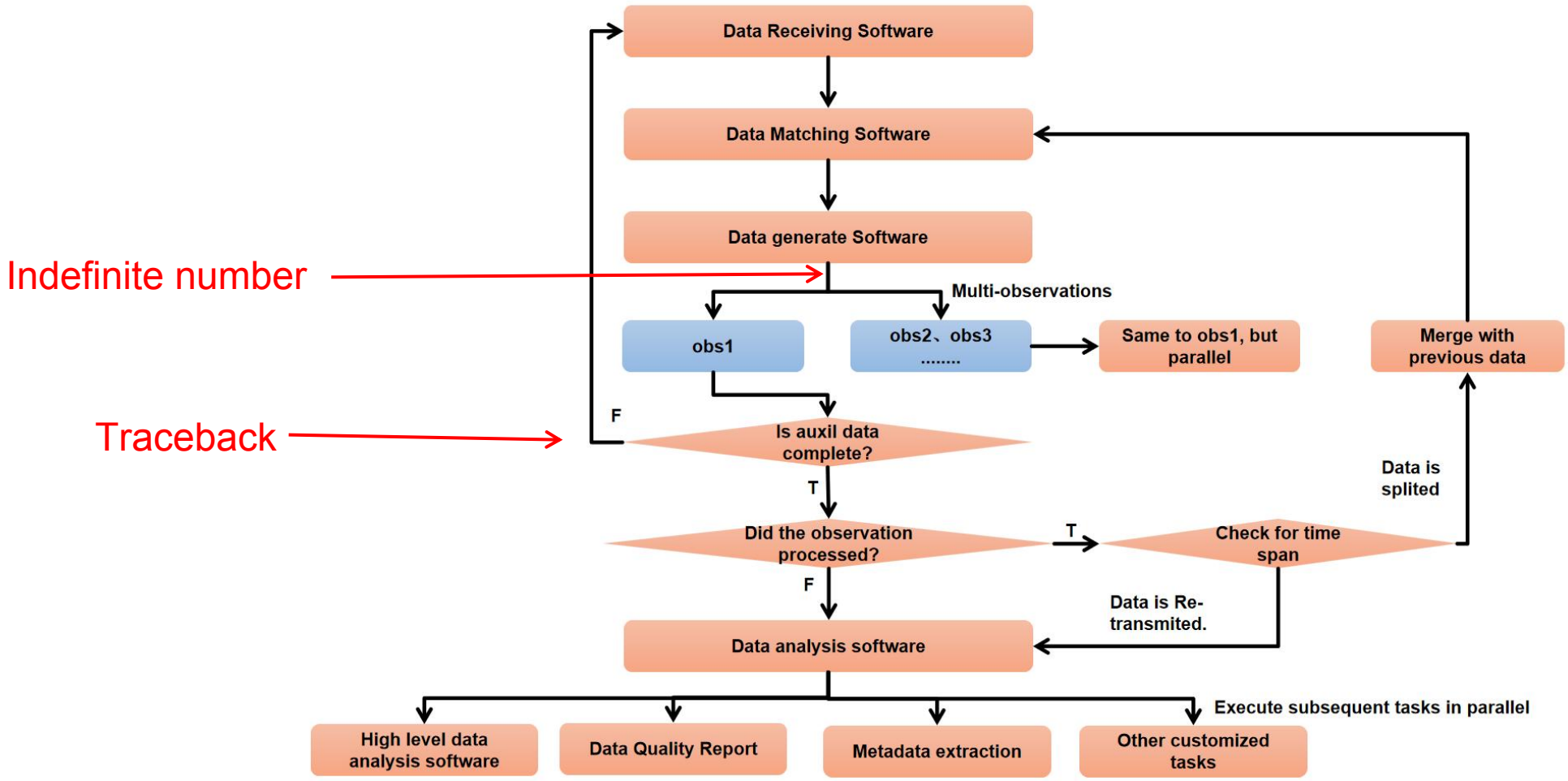
IHEP

The insitute of high energy physics (IHEP) provides FXT data analysis softwares.

Data Level


Auxil Data (attitude, orbit, etc.), Observation Plan
WXT Raw events (LV0)
Data Generation Software
WXT events grouped by observation id (LV1)
telescope defination, bad pixel, etc.
Data screen software
WXT events filtered by SAA, sun, etc. (LV2)
Vigniting, PSF, etc.
Data analysis software
Image, catalogure, light curve and spectrum (LV3)
More high-level Data analysis software
Narrow band images, short time scale transients, etc.
Inputs
Processing Steps
Outputs

Real procedure for high-energy astrophysical data



Four challenges we faced

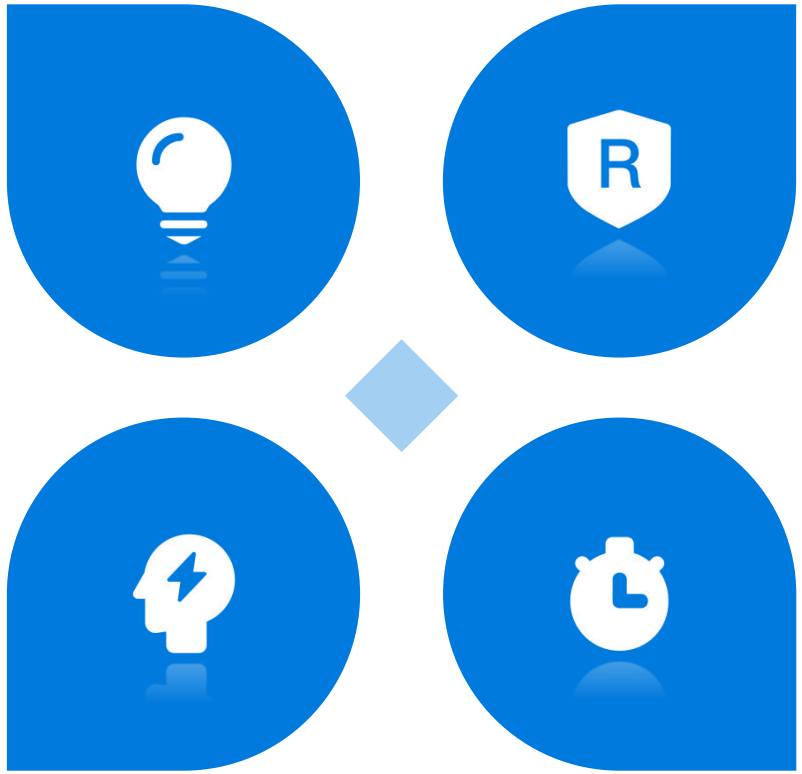
National Astronomical Data Center
国家天文科学数据中心

Integratibility

Dependency bugs caused by integration of programs from multi-organizations.

Schedulability

Directed acyclic graph(DAG)is difficult to describe some control cases.



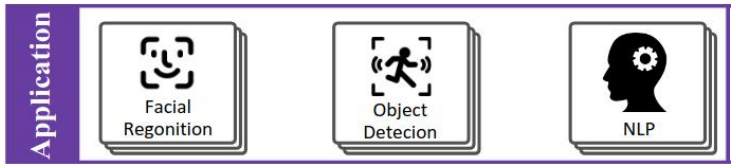
Traceability

Fault tracing when data products are not expected.

Extendibility

Frequent updates of algorithms and customized software require agile deployment.

The Execution Module

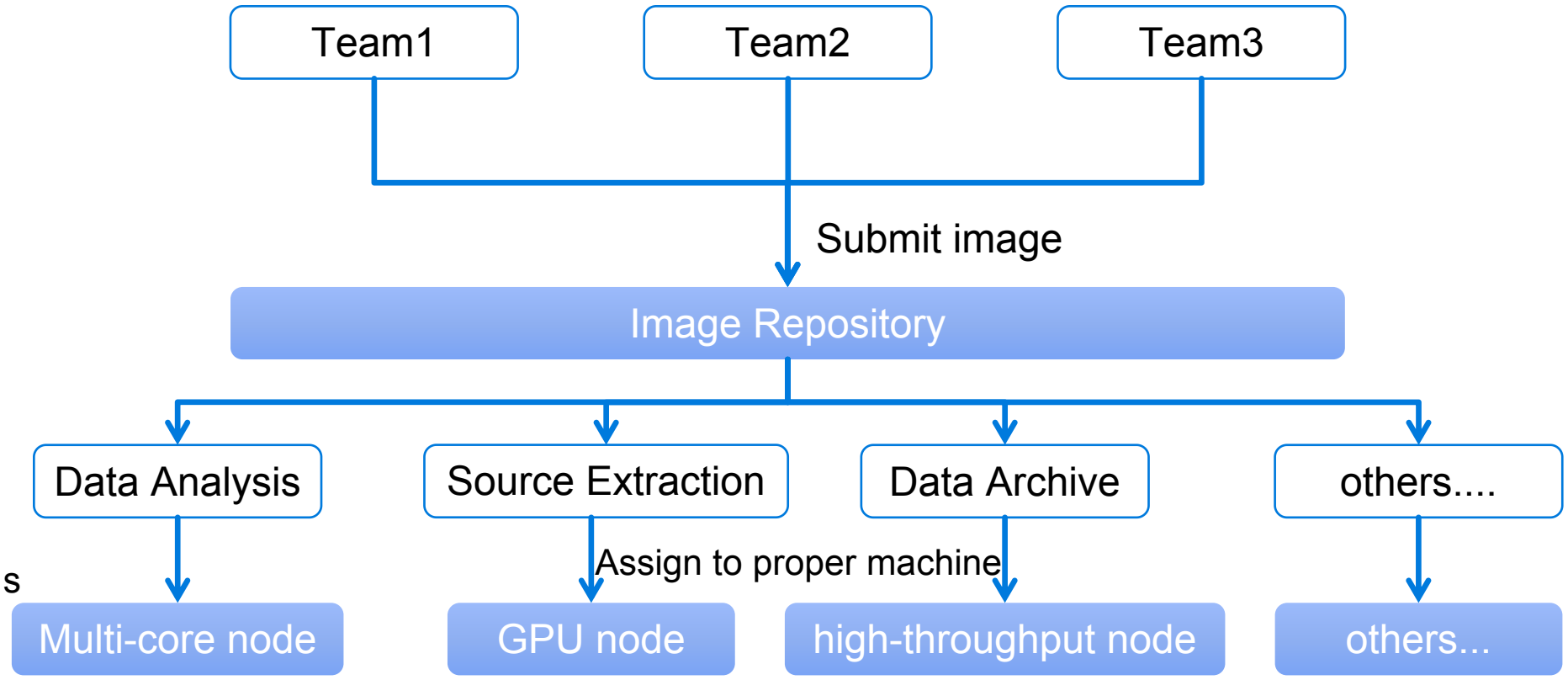
Application		Alter calling
Library		Virtual Enviroment
Runtime		Containerization
OS/Container		Virtualization
Hardware		Switch Hardware

The classification of dependency bug (DB)

Solutions for DB

The Execution Module

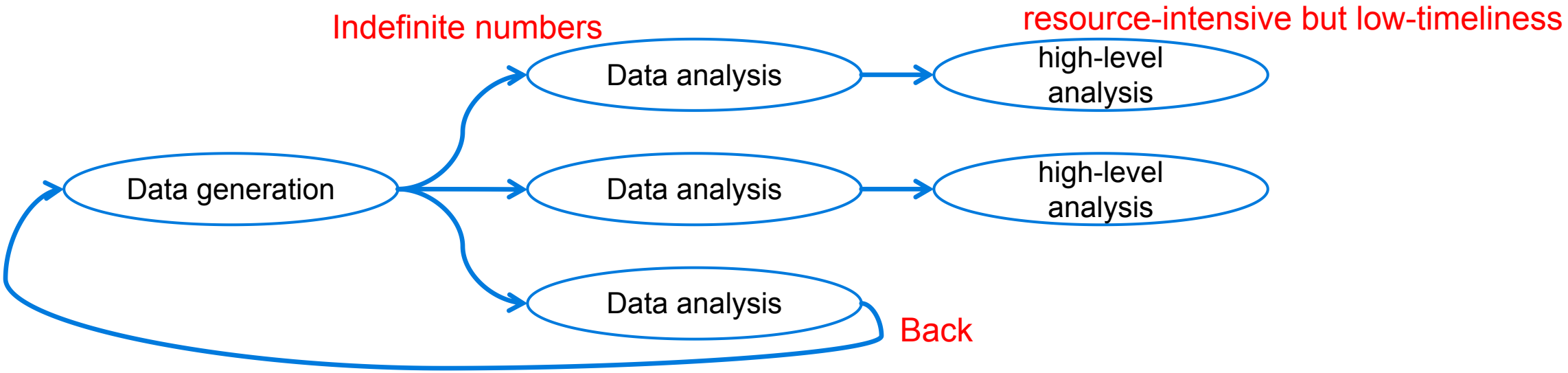
- **Solution:**
  - Containerization as Primary
  - Virtualization as Supplement
  - Allocate containers to appropriate nodes by K8S+Argo
- **Advantages:**
  - Autonomous Development Team
  - Avoid dependency bugs
  - Improve integration efficiency



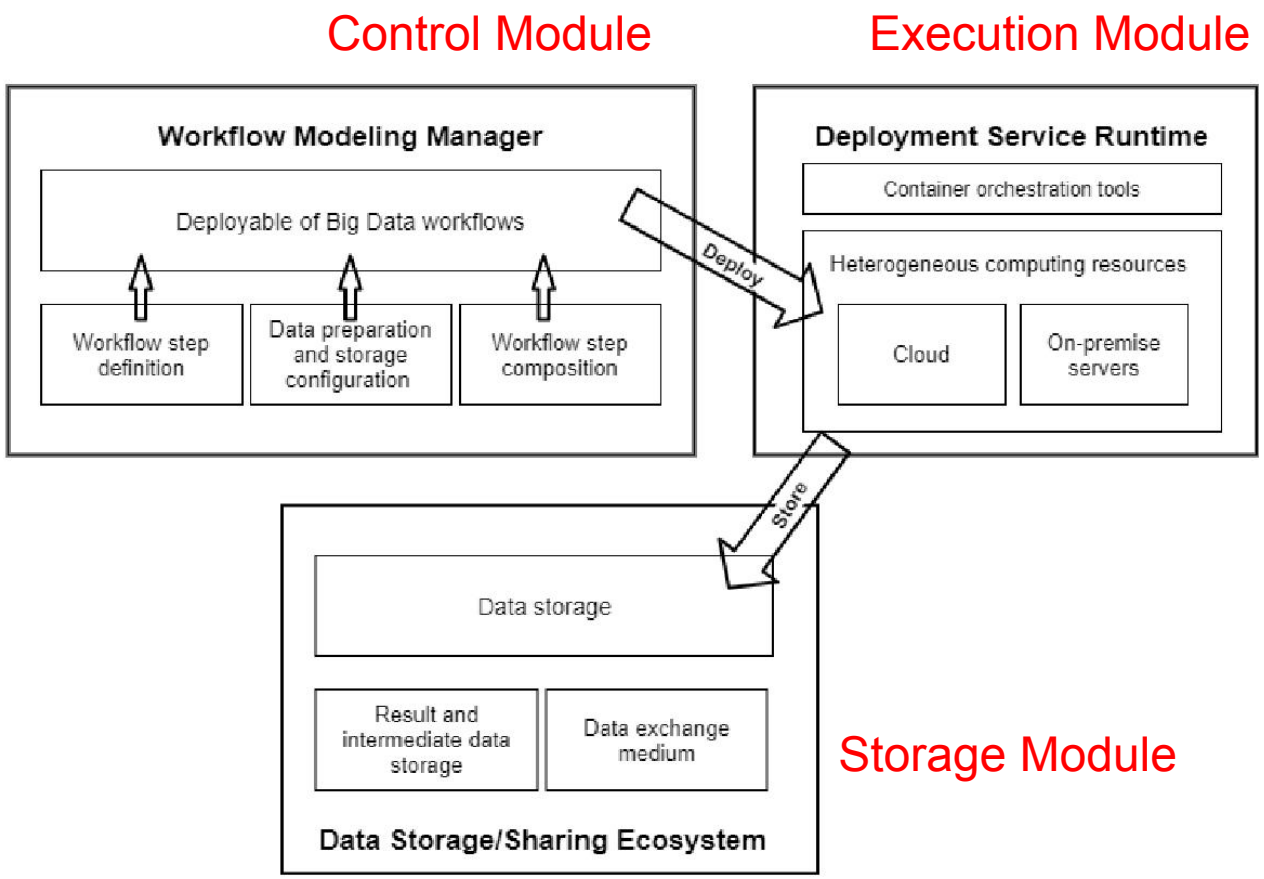
The Control Module

DAG is difficult to describe some control cases:

1. **BackTrace**: Auxiliary data are insufficient or data need to be merged, back to data receiving software.
2. **Parallel successors with indefinite** : Indefinite number of observations need to be executed in parallel.
3. **Priority**: Allocating resources to timeliness software from resource-intensive but low-timeliness applications.



The Control Module



A control/execution separation architecture（Yared, 2020）

- Execution module sends messages to the message bus.
- Control module listens messages and dynamically update workflow topology .

Problem: Invasive to algorithms.

The Control Module

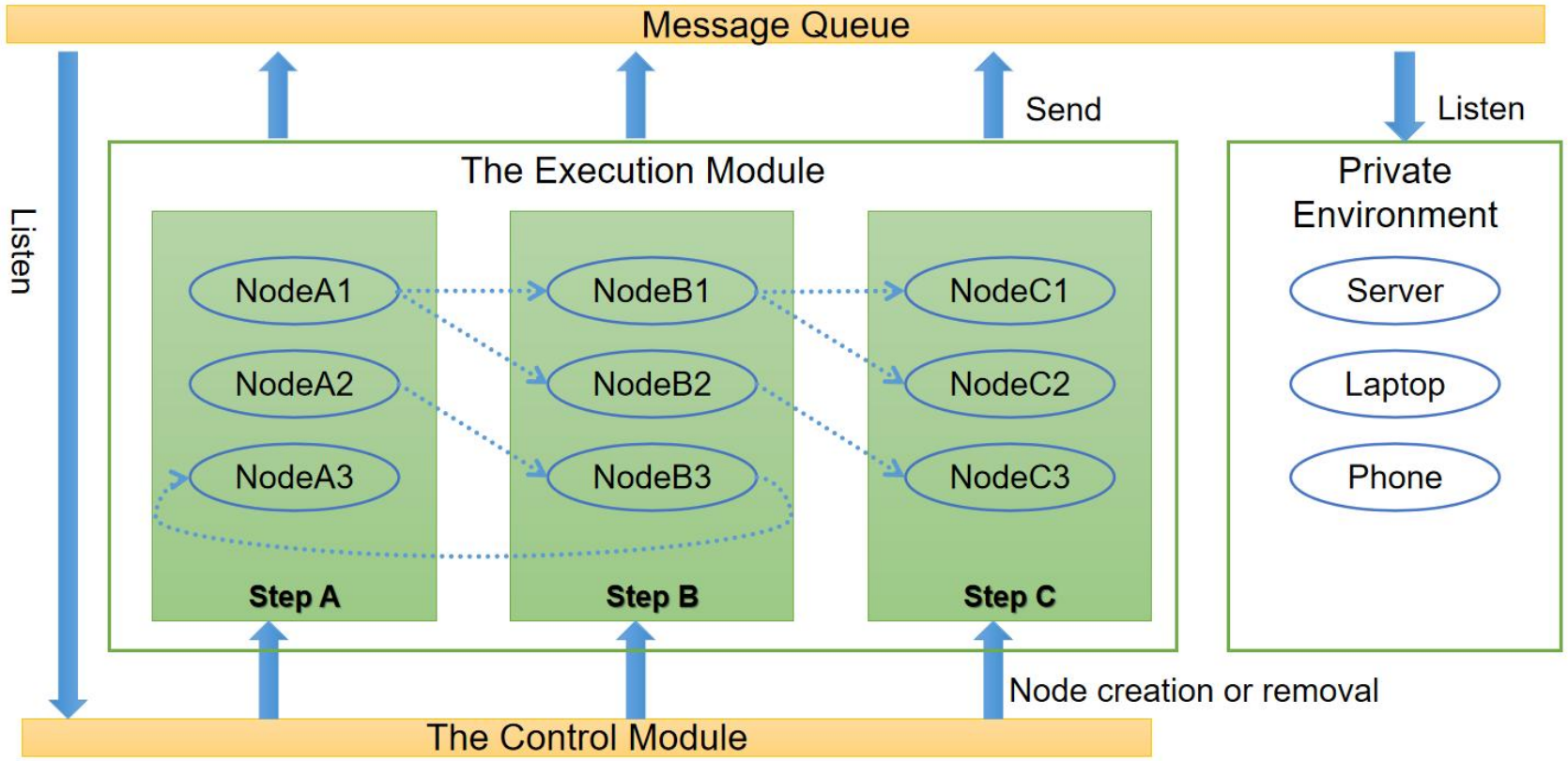
Communication Aspect

- ***Pre-processing Aspect***: Expected resource occupancy, input parameters, etc.
- ***Post-processing Aspect***: Execution status, output products, etc.

Controller

- ***Message parser***: Parses messages from the communication aspect.
- ***Topology Updater***: Dynamically update workflow topology (circular execution, indefinite parallel, etc).
- ***Scheduler***: Select tasks from the queue based on priority, resources.

The Extendibility



Two Methods to extend

- **Deploy as a step**:Package the new algorithm as an image and schedule it through the pipeline.
- **Listen on private environment**: Autonomously listen to the messages of the communication aspect and process them, making it more flexible.

The Traceability

Data and workflow activities are linked by metadata.

Data

Data can be addressed using data type and meta data.

- **Data Type**: such as light curve, image, catalogue.
- **Metadata**: locate specific data within a data type, such as observation id, CMOS id, version.

Activities

Workflow activities are also rich in metadata.

- **parent activity**: The activity that triggered this activity.
- **Input parameters**: Such as observation id, CMOS id, etc.
- **Runtime information**: Logs, startup time, etc.

Advantages

	Development	Deployment	Efficiency	Fault Tolerancy	Reproducibility
Traditional Methods	Each module is written into a package, forcing a unified environment	Global recompilation is time-consuming and prone to dependency conflict issues.	Difficult to achieve elastic expansion of resources and dynamic scheduling of tasks.	A single module failure may cause the entire system to crash.	Recordation are incomplete and invasiveness.
New Framework	Modules adopts different technology stacks and environments. Team autonomy and decoupling.	Modules are packaged into images, achieving agile deployment.	Intelligent scheduling based on resources and priorities, playing a "peak cutting" role.	Containerization ensures that the failure of a single activity does not affect the non-successor activities.	Metadata are structured, making it easy to trace back.

Experiences we learned

National Astronomical Data Center
国家天文科学数据中心

- Containerization of algorithms
- Separation of control and execution
- Message queue to support extension
- Structured metadata to find provenance

Achievements

01

For pipeline itself

- Updated 300+ times
- Integration time shortened from hours to minutes.
- Triggered 60000+ times

02

For scientific outputs

- Produced 100+ TB products
- 12 scientific papers

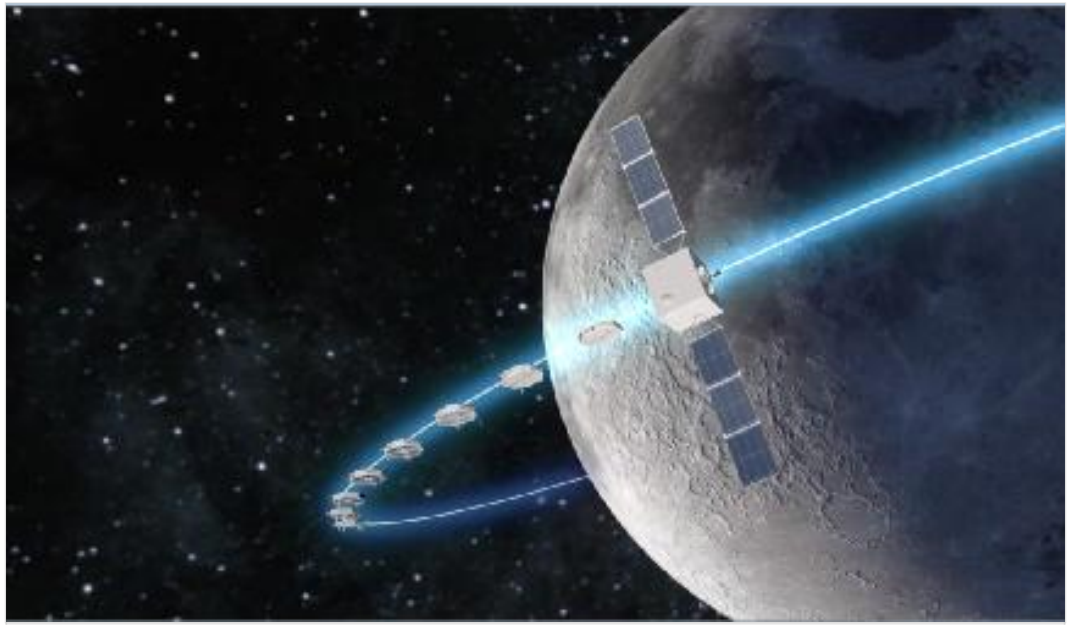
Future: VO Resource

Assign VO identifier for each activity

Standardize the activity metadata

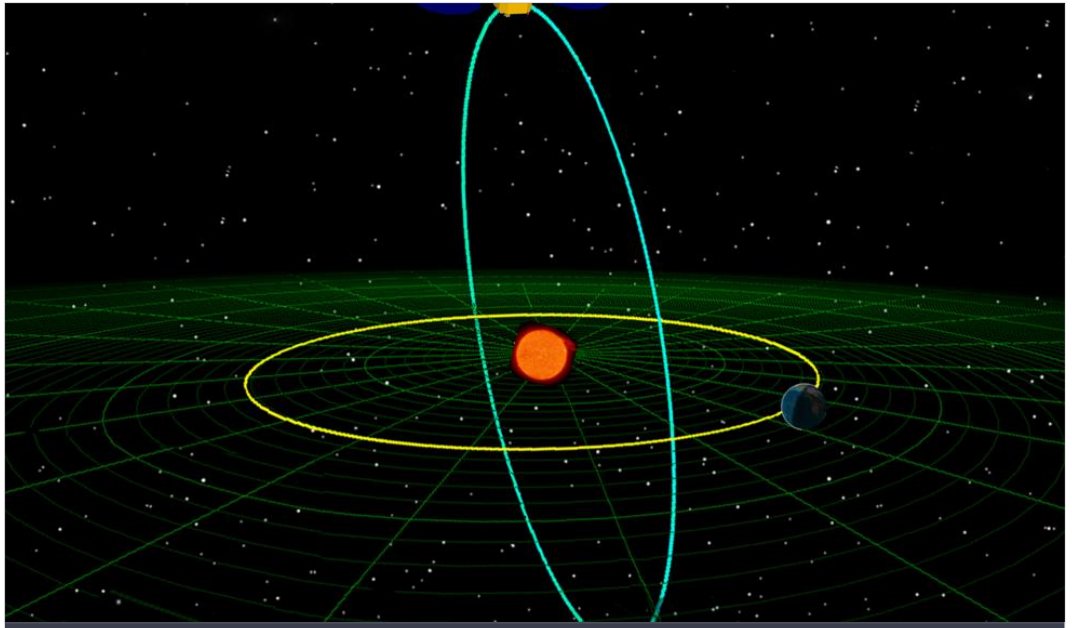
Associate the VO provenance information

Future Applications



Discovering Sky at the Longest wavelength（DSL）

A distributed interferometric array composed by 10 satellites.



Solar Polar-orbit Observatory (SPO)

A satellite with large orbital inclination that images the solar polar region.

Thank You!

Zhen Zhang

zhangzhen@nao.cas.cn