SDC-CH:

Processing Infrastructure and Pipelines

Florian Dubath
Université de Genève

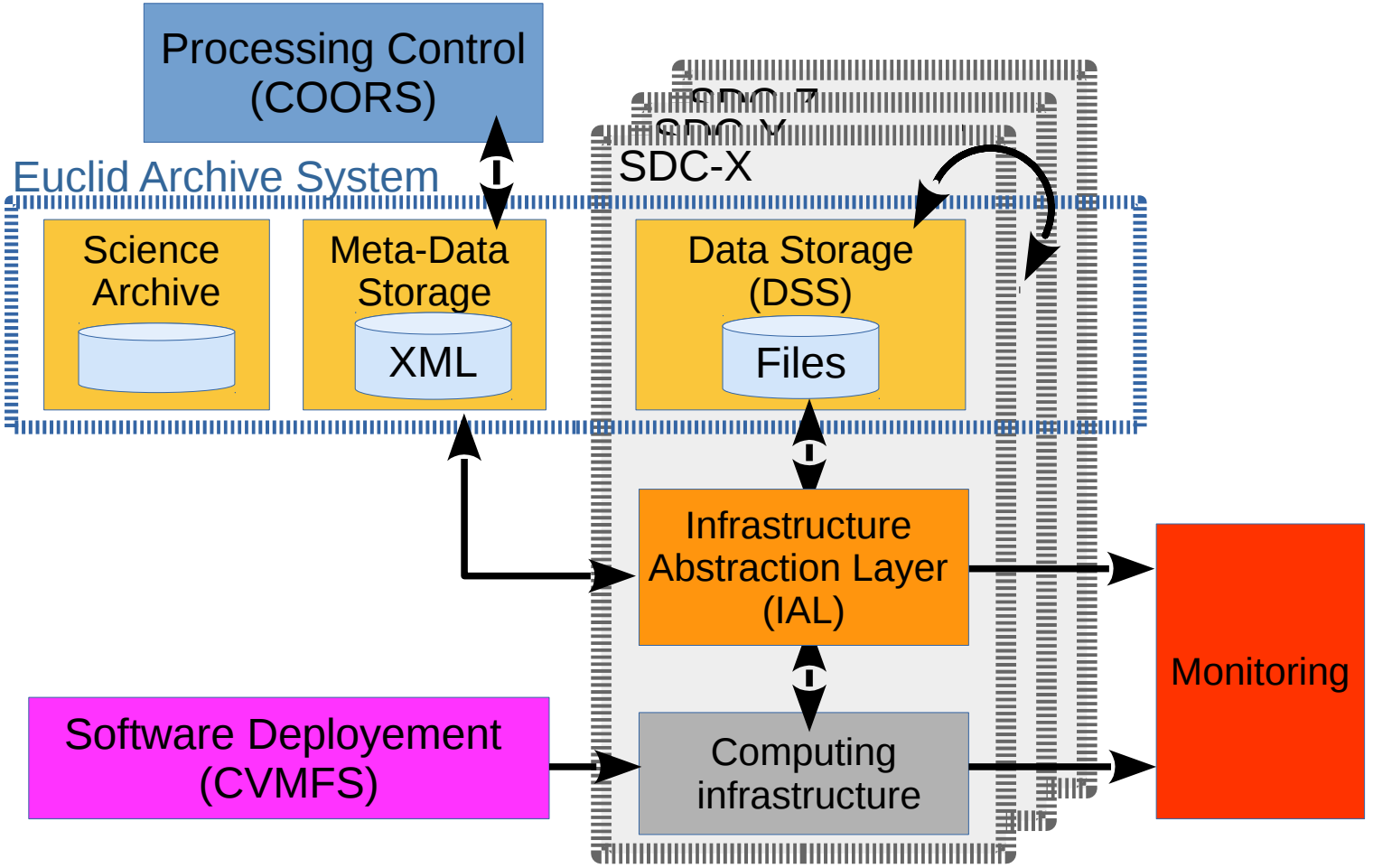
Swiss Euclid Days, 02-2020

OUTLINE

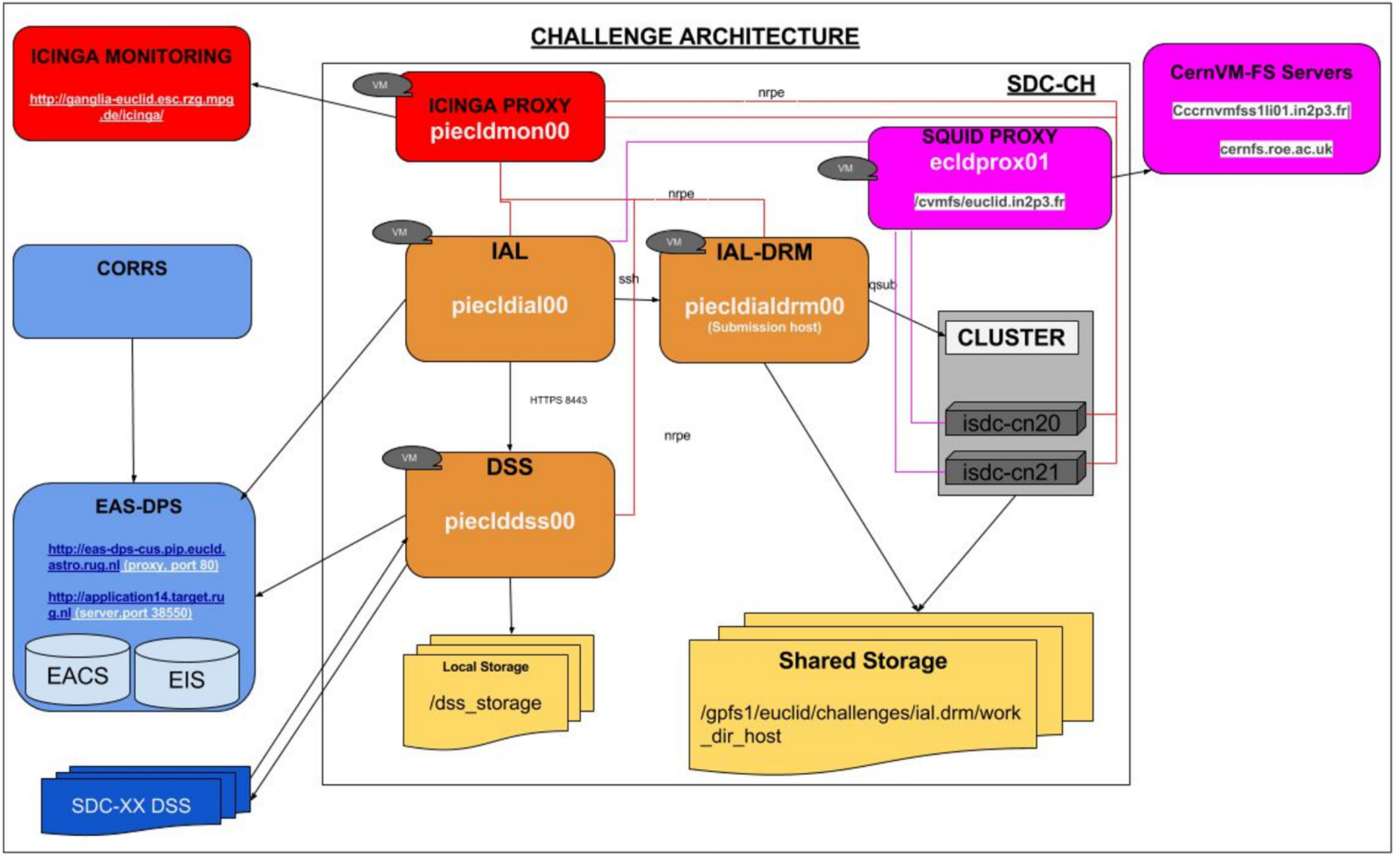
- **Data Processing Infrastructure**
- **SDC-CH Implementation**
- **Processing Functions & Data Model**
- **Pipelines**
- **SDC-CH Packages & PHZ Pipeline**

Data Processing Infrastructure

- Too many data to move them around
- Data stored on SDCs, Meta-Data in a common DB
- Each software can be run on any SDC

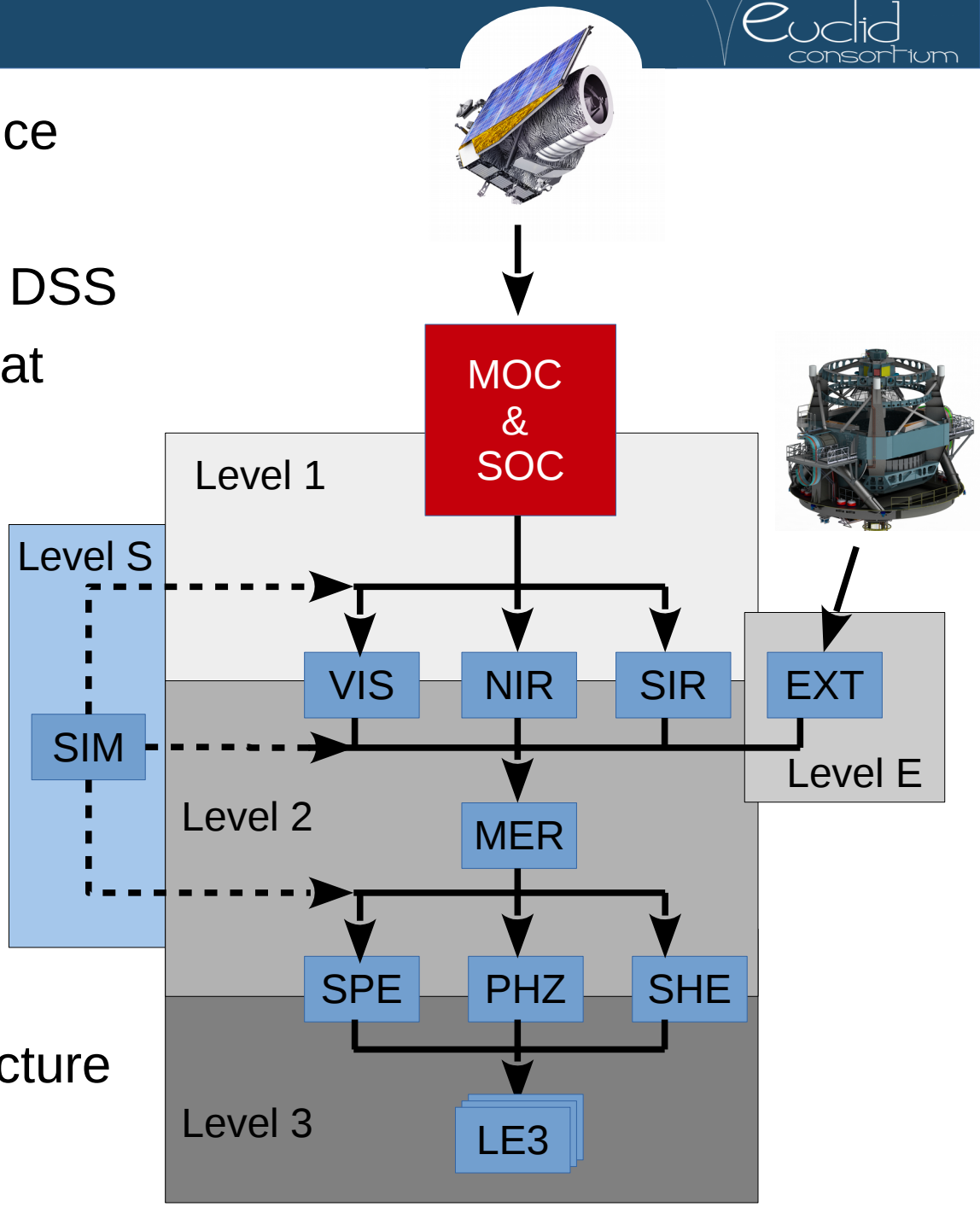


SDC-CH Implementation



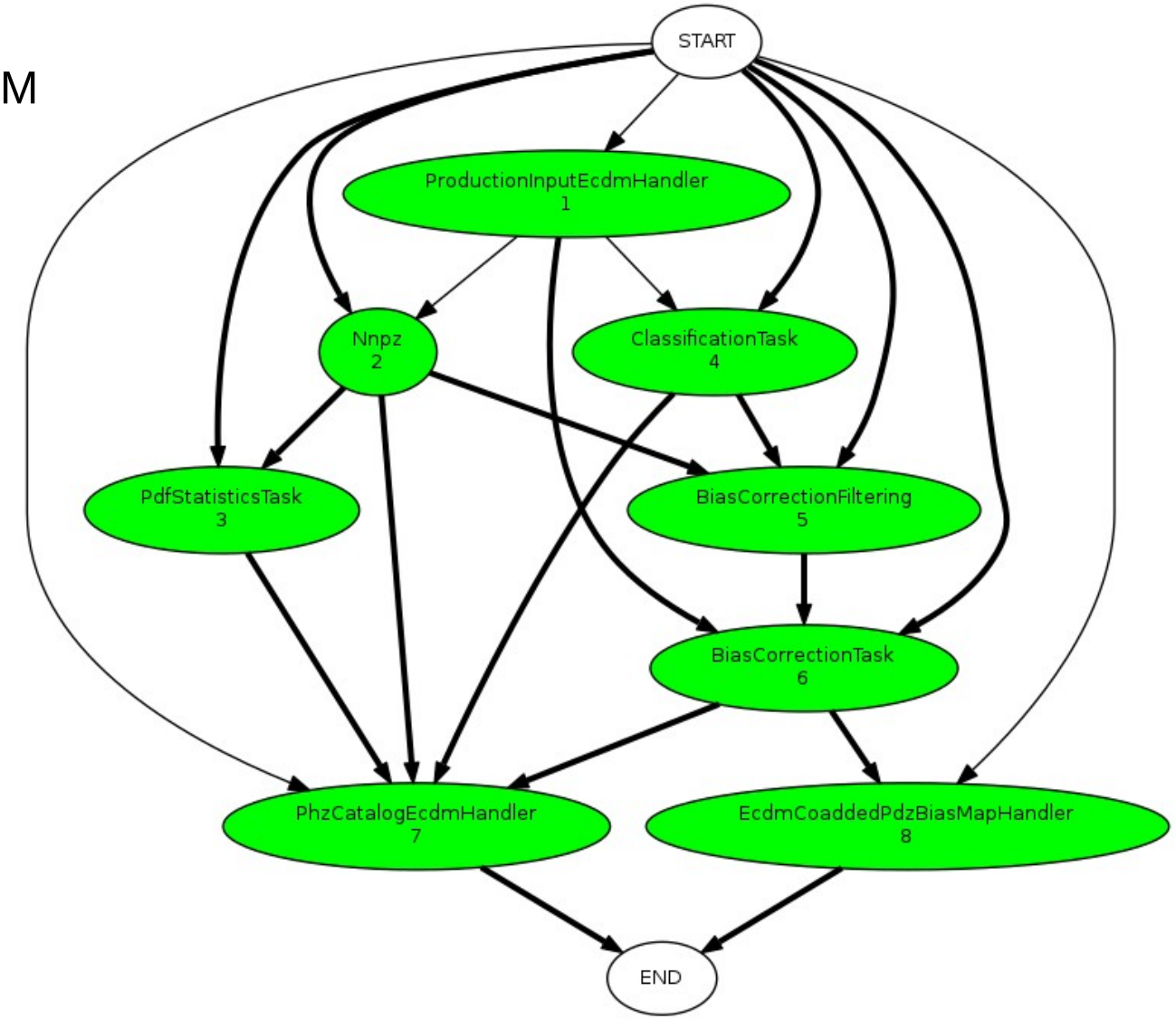
PF & ECDM

- Data are processed in sequence
- Intermediate results: stored in DSS following pre-established format (Euclid Common Data Model)
- No loop
- Later PF's have only access to upstream results
- Each PF build pipelines to be run in the processing infrastructure



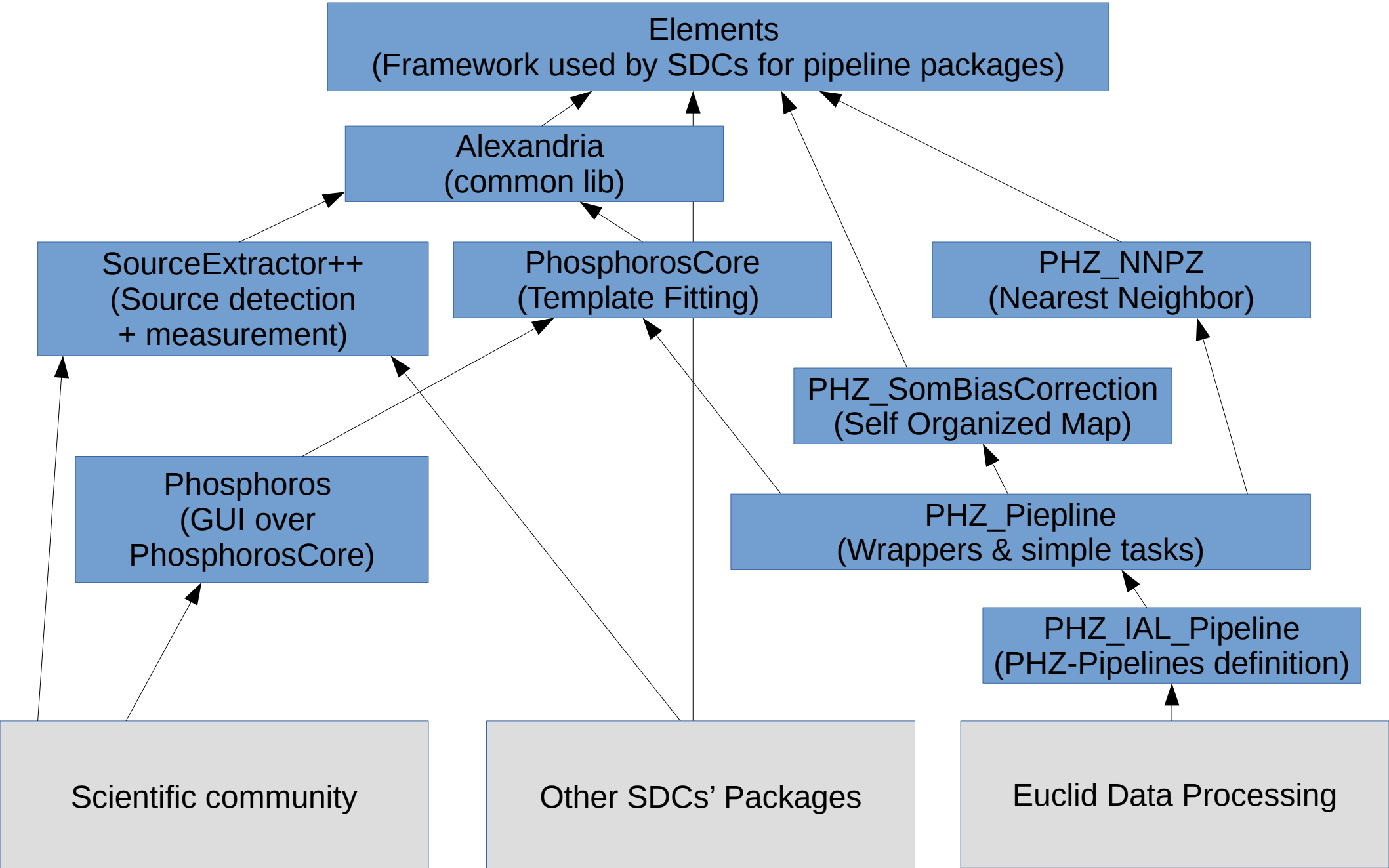
Pipelines

- PF Pipelines: build by chaining software
- Input/output data format: ECDM
- Scientific software must be wrapped so they follow interfaces and understand ECDM
- Pipeline use fixed version of the software packages => tracability of products



SC456-SC7 PHZ Pipeline

SDC-CH Packages



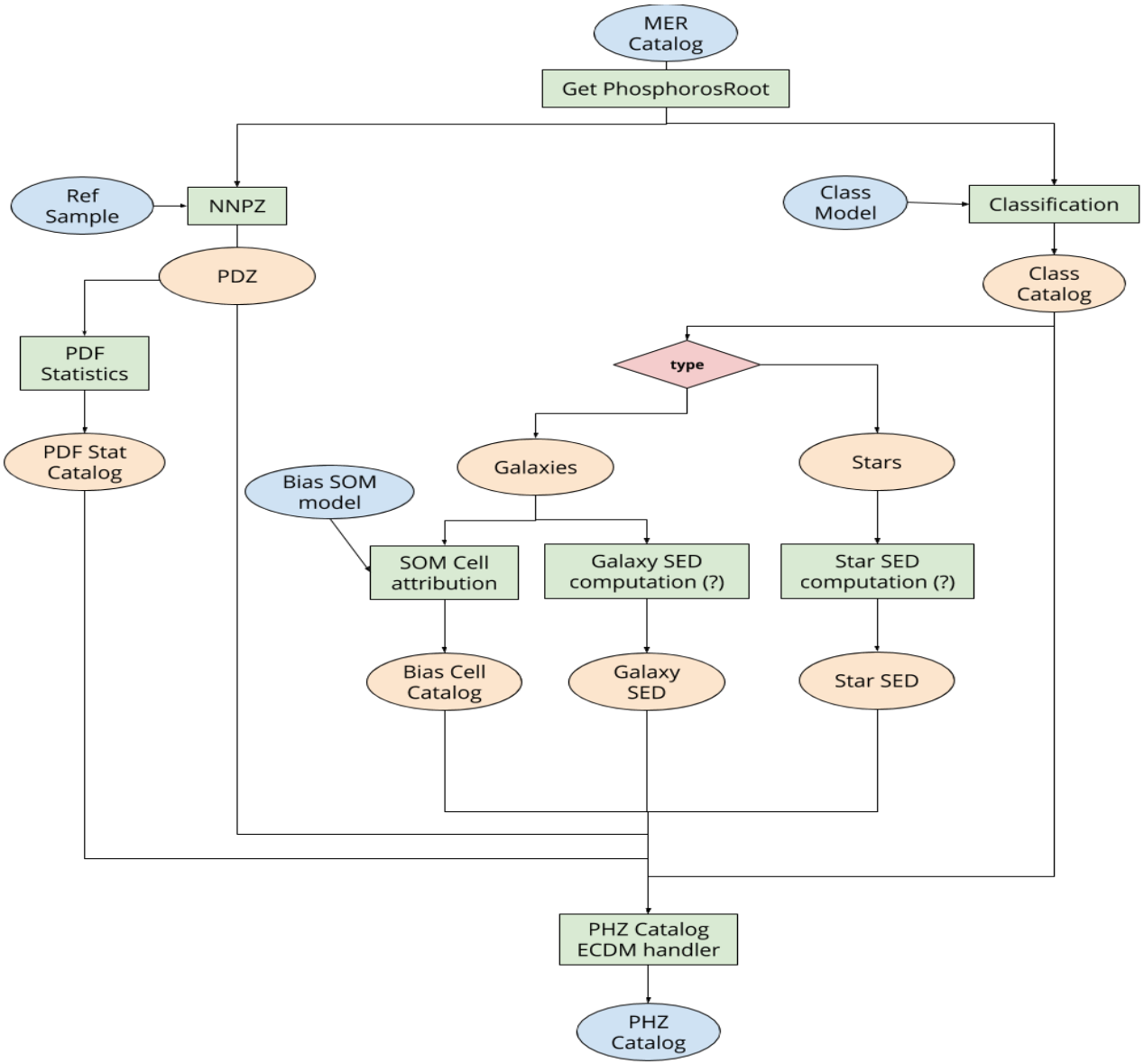
PHZ Pipeline

Example: Pipeline computing the Photometric Redshift (SC8 forecast)

ECDM Data

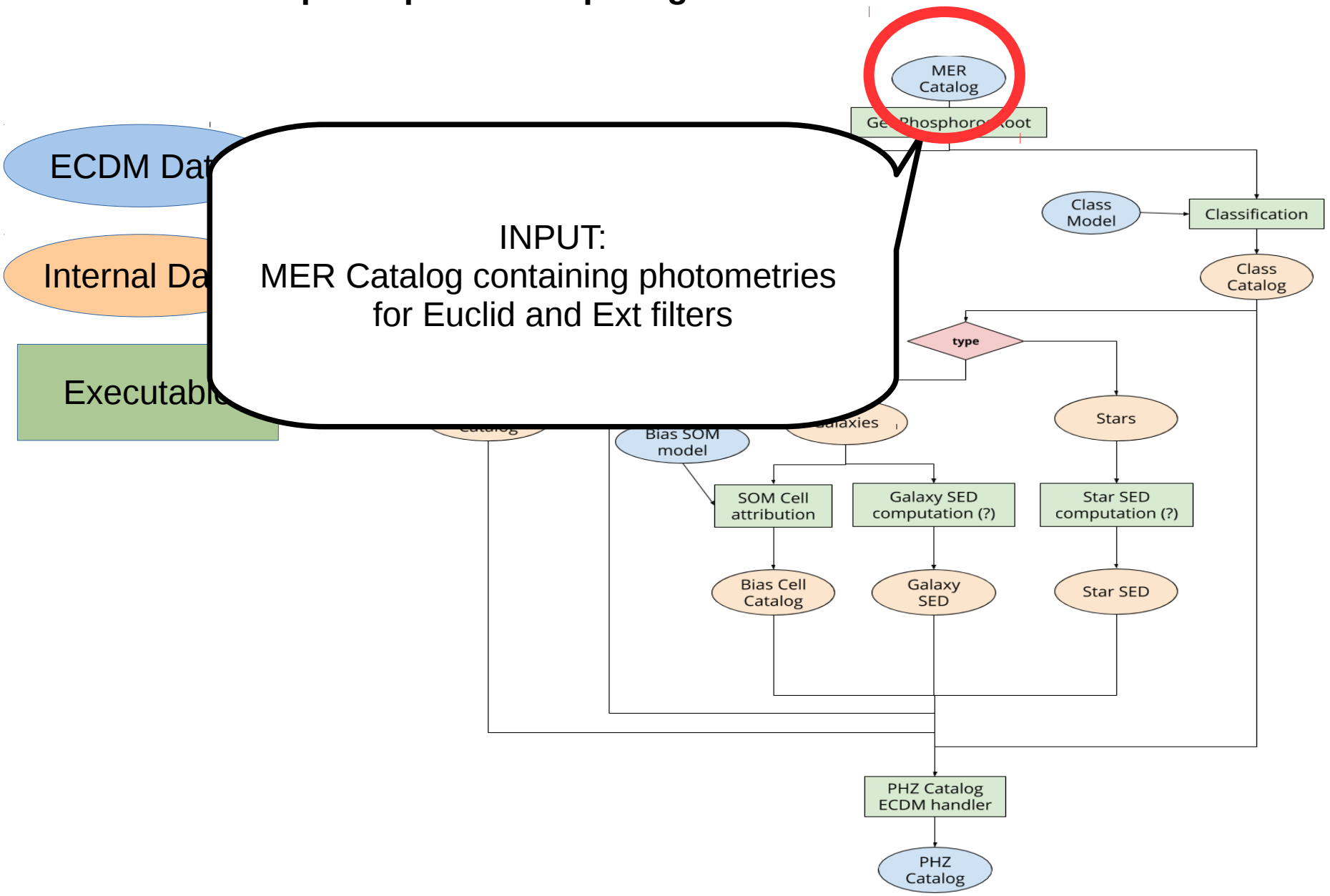
Internal Data

Executable



PHZ Pipeline

Example: Pipeline computing the Photometric Redshift



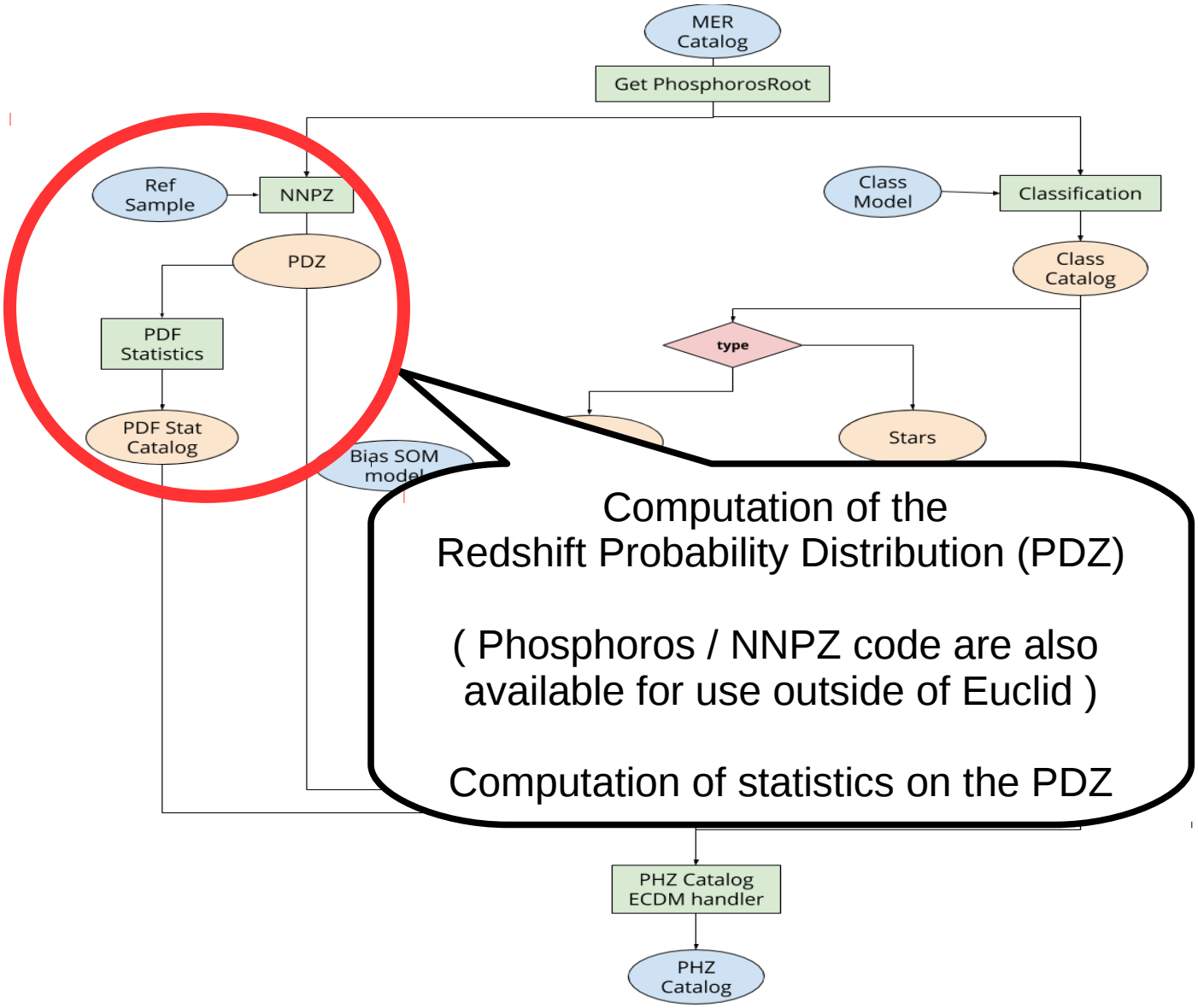
PHZ Pipeline

Example: Pipeline computing the Photometric Redshift

ECDM Data

Internal Data

Executable

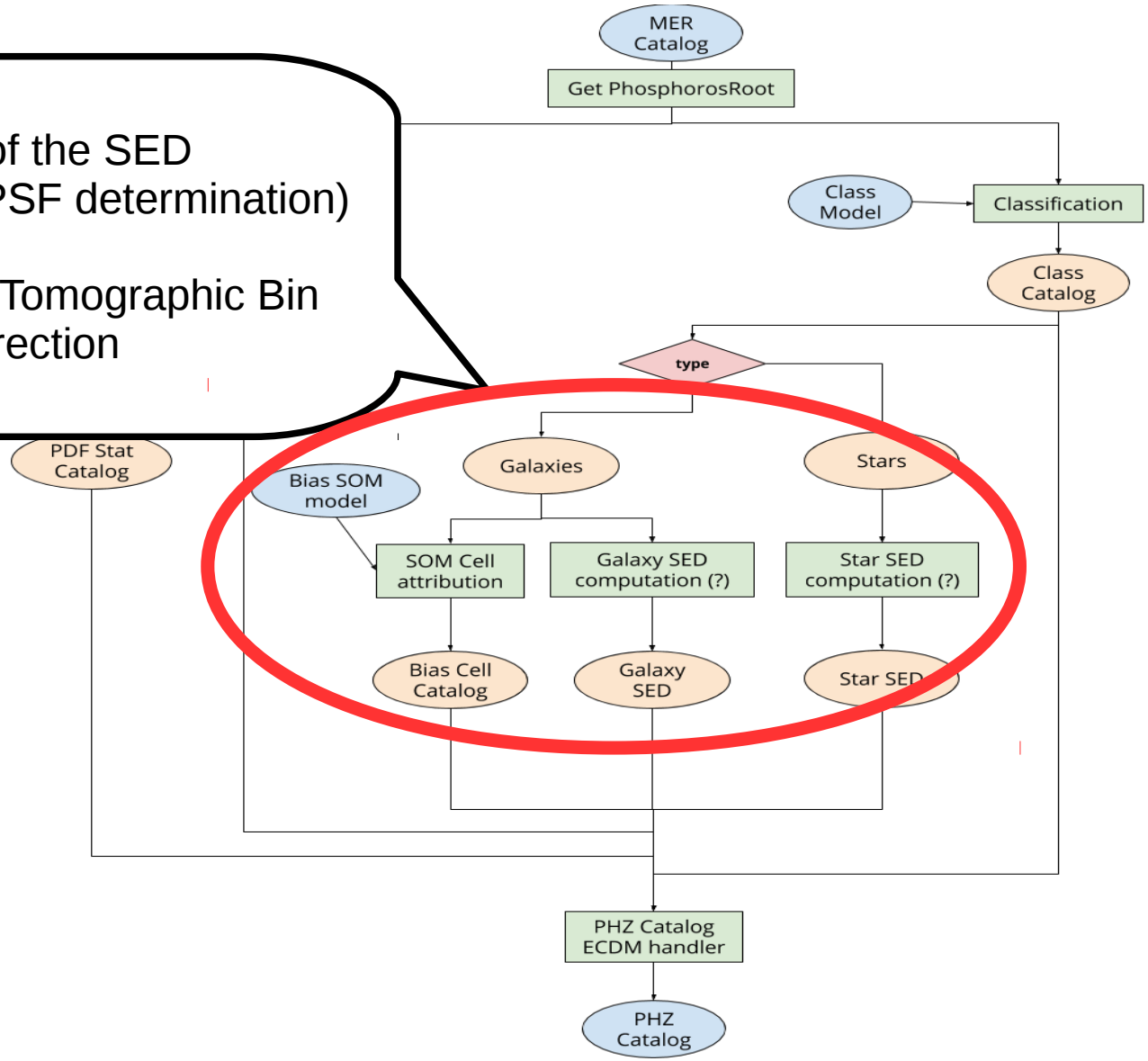


PHZ Pipeline

Example: Pipeline computing the Photometric Redshift

Computation of the SED
(Needed by SHE for PSF determination)

Determination of the Tomographic Bin
& bias correction



PHZ Pipeline

Example: Pipeline computing the Photometric Redshift