ROC Data Processing Subsystem (DPS)

X.Bonnin
and the RPW Lead CoI analyzer/sensor teams

Outlines

1. RPW data products overview
2. ROC data processing subsystem (DPS) design
3. RPW Calibration Software (RCS) and data products

3.1. TDS calibration software (TDS_CALBA)

3.2. LFR calibration software unit (LFR_CALBUT)

3.3. THR calibration software (THR_CALBAR)

3.4. Bias calibration software (BICAS)

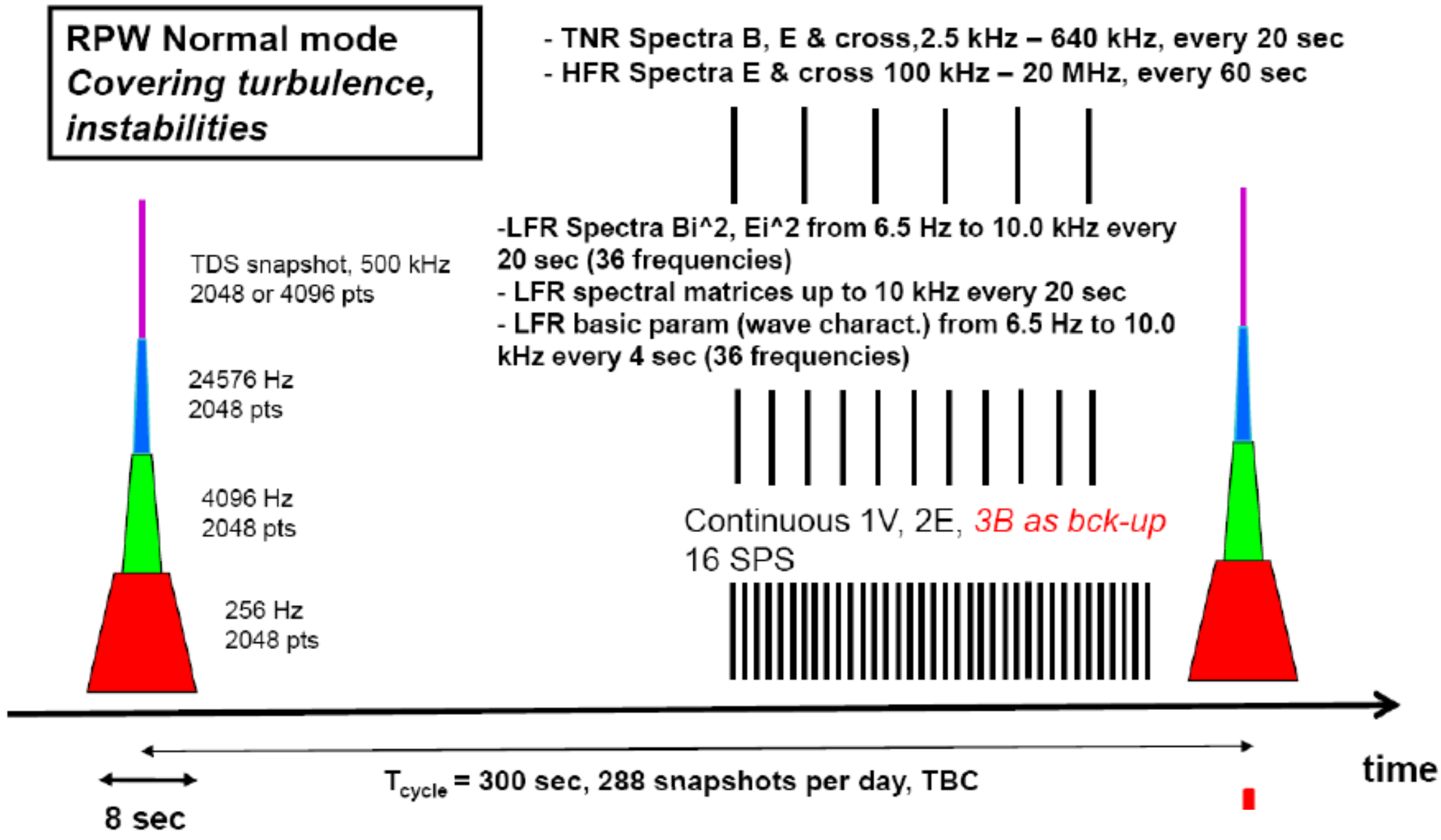
3.5. SCM calibration software (SCMCAL)

Outlines

1. RPW data products overview
2. ROC data processing subsystem (DPS) design
3. RPW Calibration Software (RCS) and data products
   - 3.1. TDS calibration software (TDS_CALBA)
   - 3.2. LFR calibration software unit (LFR_CALBUT)
   - 3.3. THR calibration software (THR_CALBAR)
   - 3.4. Bias calibration software (BICAS)
   - 3.5. SCM calibration software (SCMCAL)

RPW science data

RPW Normal mode
Covering turbulence, instabilities



- + TDS "derived" data (stats/histograms/psd/mamp)
- + SBM1 (shock) and SBM2 (LW) selective data

RPW data processing levels (1/2)

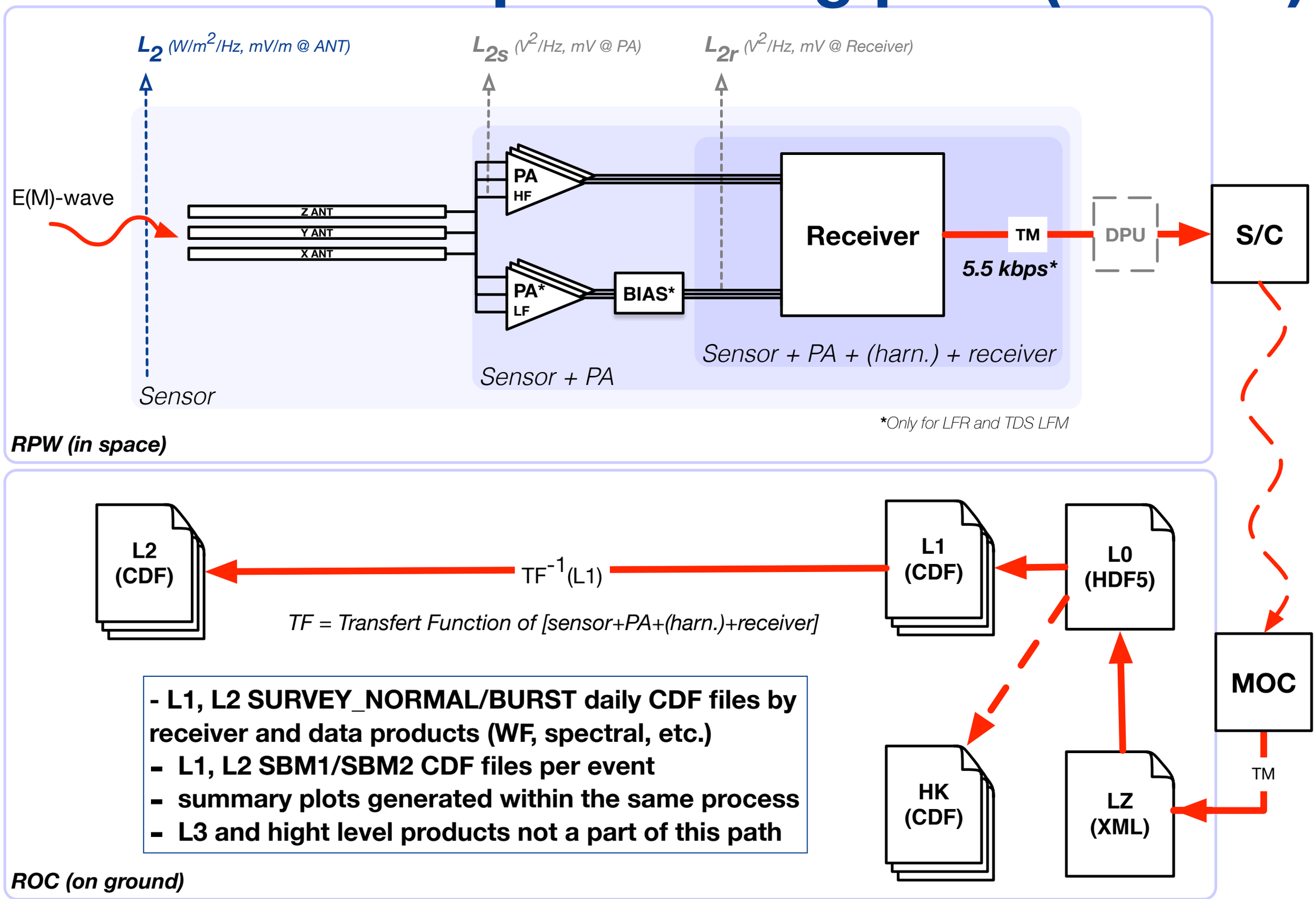
Level	Description	Source	Users
LZ	**Raw RPW telemetry (TM) as delivered by the Solar Orbiter MOC Data Dissemination System (DDS)** ApID-separated, sorted, cleaned. Daily XML format files (TM binary data encapsulated)	MOC	ROC
L0	**« Raw » data, unpacked and decompressed data** Daily HDF5 format files.	ROC	ROC
HK	**Uncompressed and decommuted HK, engineering values, UTC-tagged.** Daily SOC-compliant CDFs, 1 file per sub-system.	ROC	ROC, RPW lead-CoI teams
L1	**Uncompressed and decommutated L0 + UTC-tagged, waveform and spectral data in telemetry units (uncalibrated) in spacecraft coordinate system.** **Data affinity group. E + B components in the same files.** Daily SOC-compliant CDFs, Quik Look and daily/orbital summary plots	ROC	ROC, RPW lead-CoI teams, Solar Orbiter instrument PI teams, ESAC data archive, CDPP
L1R	**Same than L1, but only for waveform products and including information to link data with the associated calibration table. Separated E and B Waveform products.**	ROC	ROC, RPW lead-CoI teams
L2	**L1 + waveform and spectral data in fully calibrated physical units [V, mV/m, nT, W/m^2/Hz, nT^2/Hz] in spacecraft and heliophysical coordinates systems.** **Separated E and B Waveform products.** Daily SOC-compliant CDFs, Quik Look and daily/orbital summary plots	ROC, using the RPW Calibration Software (RCS)	ROC, RPW lead-CoI teams, Solar Orbiter instrument PI teams, ESAC data archive, CDPP
L3	**L2 + VxB removal for DC E-field measurement, offsets and corrections with data quality flags. Poynting flux.** **Plasma density. Spacecraft potential. Merged B. Merged density and temperature. Goniopolarimetry.** Daily SOC-compliant CDFs, Science data plots.	RPW sub-systems teams	ROC, RPW lead-CoI teams, Solar Orbiter instrument PI teams, ESAC data archive, CDPP
L4	**Event time tags and parameters**	TBD	TBD

RPW data processing levels (2/2)

Level	Description	Source	Users
CAL	**Calibration table data in CDFs**	RCS teams	ROC, RPW lead-CoI teams, Solar Orbiter instrument PI teams, ESAC data archive
ANC	**Ancillary data in SPICE kernels and CDFs**	SOC	ROC, RPW lead-CoI teams, CDDP data archive centre (TBC)
LL01	**LL engineering data, output of the LL pipeline** **Specific CDF format files**	SOC (from RPW Low Latency Data Pipeline) ROC (backup instance)	SOC, ROC, Solar Orbiter instrument PI teams
LL02	**Operational LL data, enhanced with S/C HK** **Specific CDF format files**	SOC	SOC, ROC, Solar Orbiter instrument PI teams
LL03	**Visualisation of operational LL data, in "quicklook" format** **Data in PNG or JPG2000 (TBC)**	SOC	SOC, ROC, Solar Orbiter instrument PI teams

- Low Latency data (LLD) are not science data, but used for operations (e.g., Bias sweeping and SBM event data selection for RPW) — 64 bits/s (692 kB/day)
- SOC will supply a Web page to view LLD, but might be restricted to instrument teams (i.e., ROC for RPW)
- LLD generated at SOC will be retrieved asap by the ROC and make available to the RPW teams.
- ROC will als generate RPW LLD at LESIA in backup (and for comparison)

RPW data processing path (E-case)



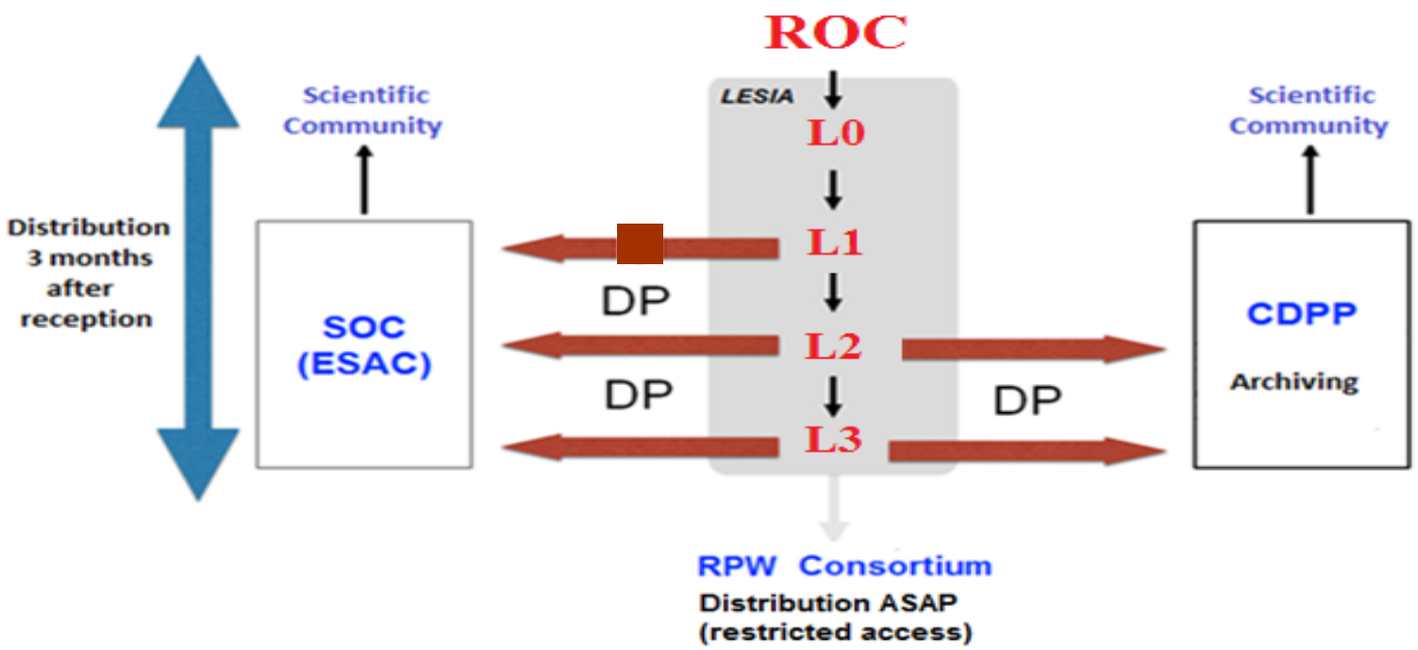
Data production and distribution concept

Data production responsibilities

Data Products (DP):

- *survey/sbm (CDF)*
- *summary plots*
- *geometry(antenna)(SPICE)*

+ Low Latency Pipeline
For RPW radio flux
@ 1 MHz



During Calibration

L0 → LESIA → **L1** → BIAS, TDS, LFR & THR Teams → **L2s**

After launch

L0 → LESIA → **L1** → Softwares provided by the BIAS, TDS, LFR & THR Teams Data produced @ LESIA → **L2**

Outlines

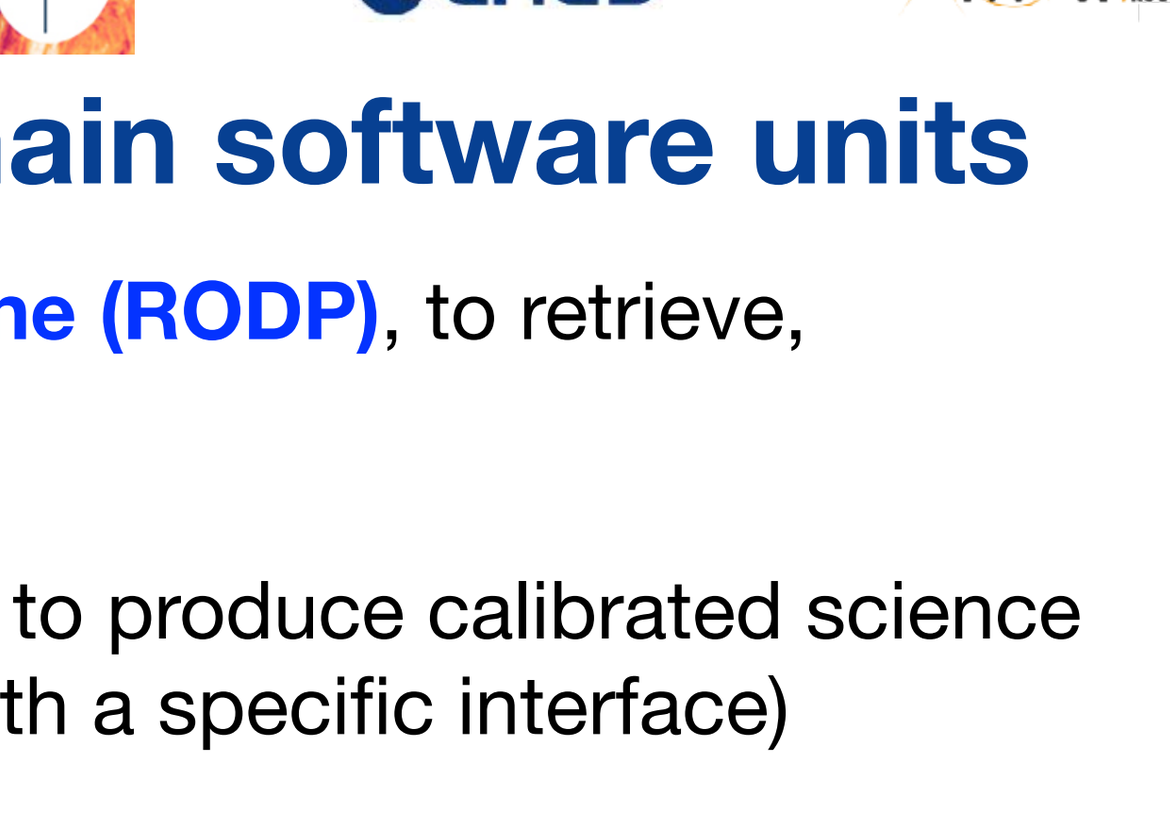
1. RPW data products overview
2. ROC data processing subsystem (DPS) design
3. RPW Calibration Software (RCS) and data products
   - 3.1. TDS calibration software (TDS_CALBA)
   - 3.2. LFR calibration software unit (LFR_CALBUT)
   - 3.3. THR calibration software (THR_CALBAR)
   - 3.4. Bias calibration software (BICAS)
   - 3.5. SCM calibration software (BICAS)

DPS main software units

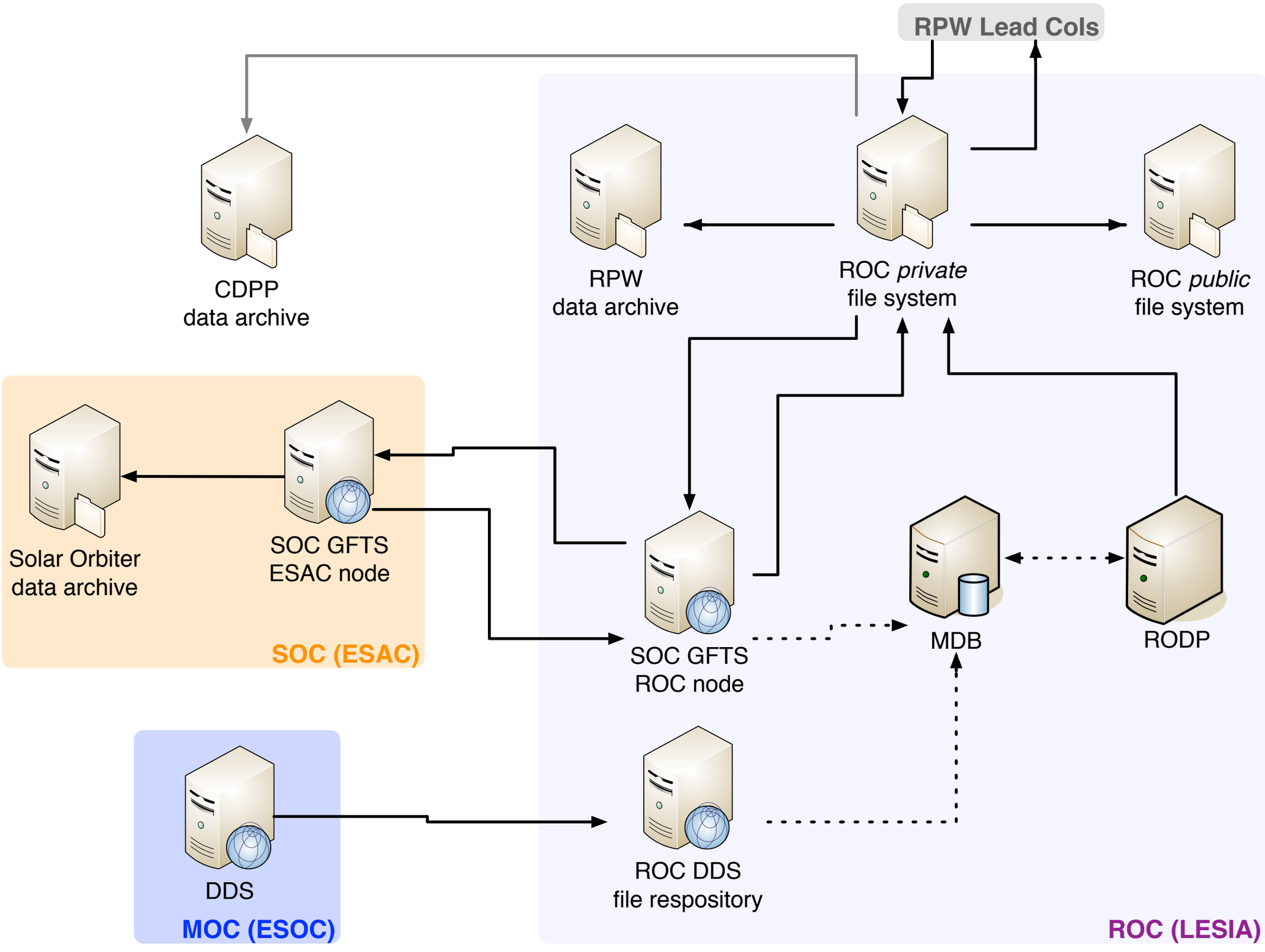
- **ROC Operations and Data Pipeline (RODP)**, to retrieve, process and distribute RPW data.

- **RPW Calibration Software (RCS)** to produce calibrated science data (run at LESIA by the RODP with a specific interface)

- **Low Latency Virtual Machine (LLVM)** to process RPW LL01 data (primary instance run at MOC)



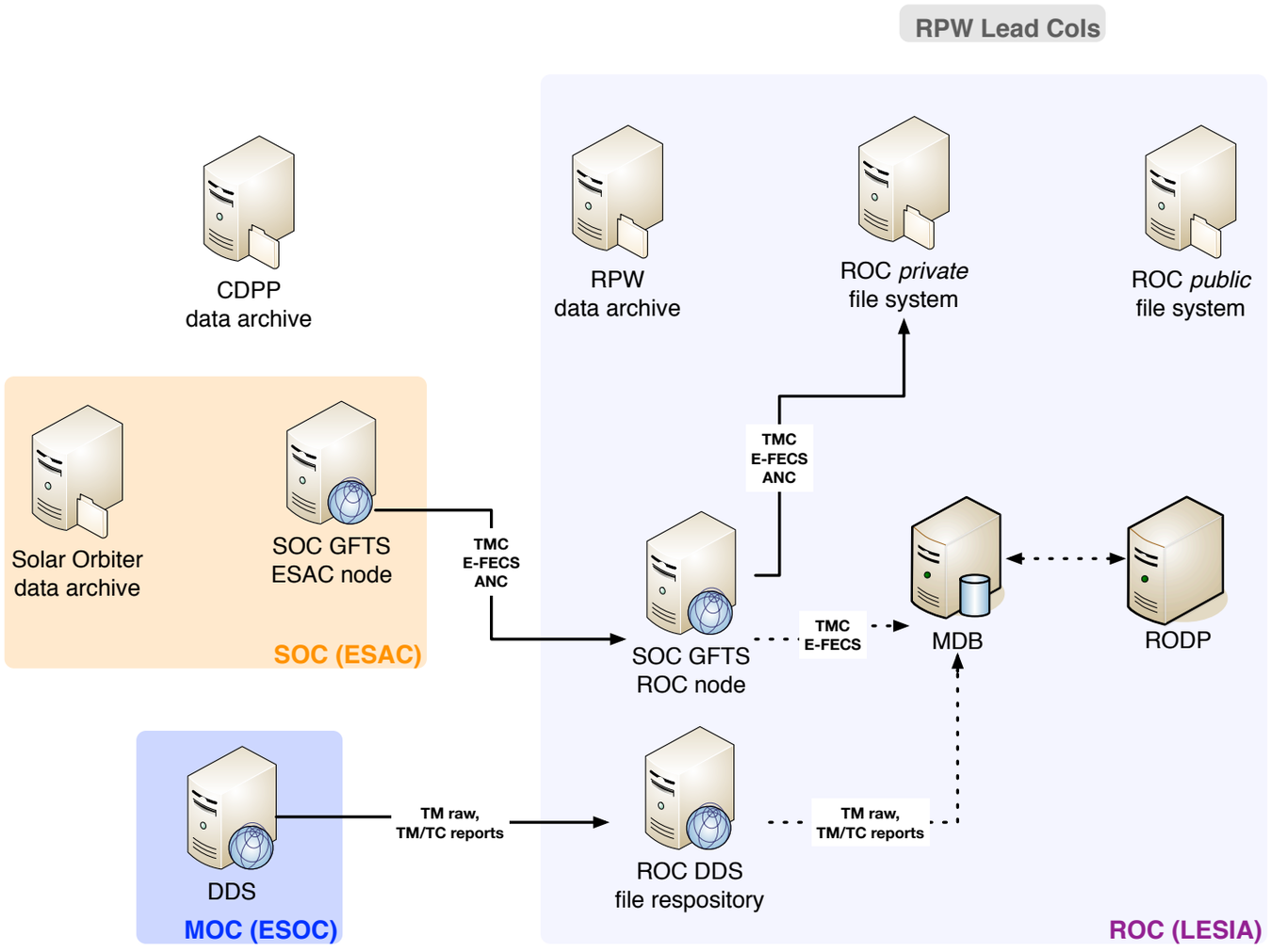
RODP infrastructure



RODP nominal use case - Step #1

- Retrieve and pre-process data

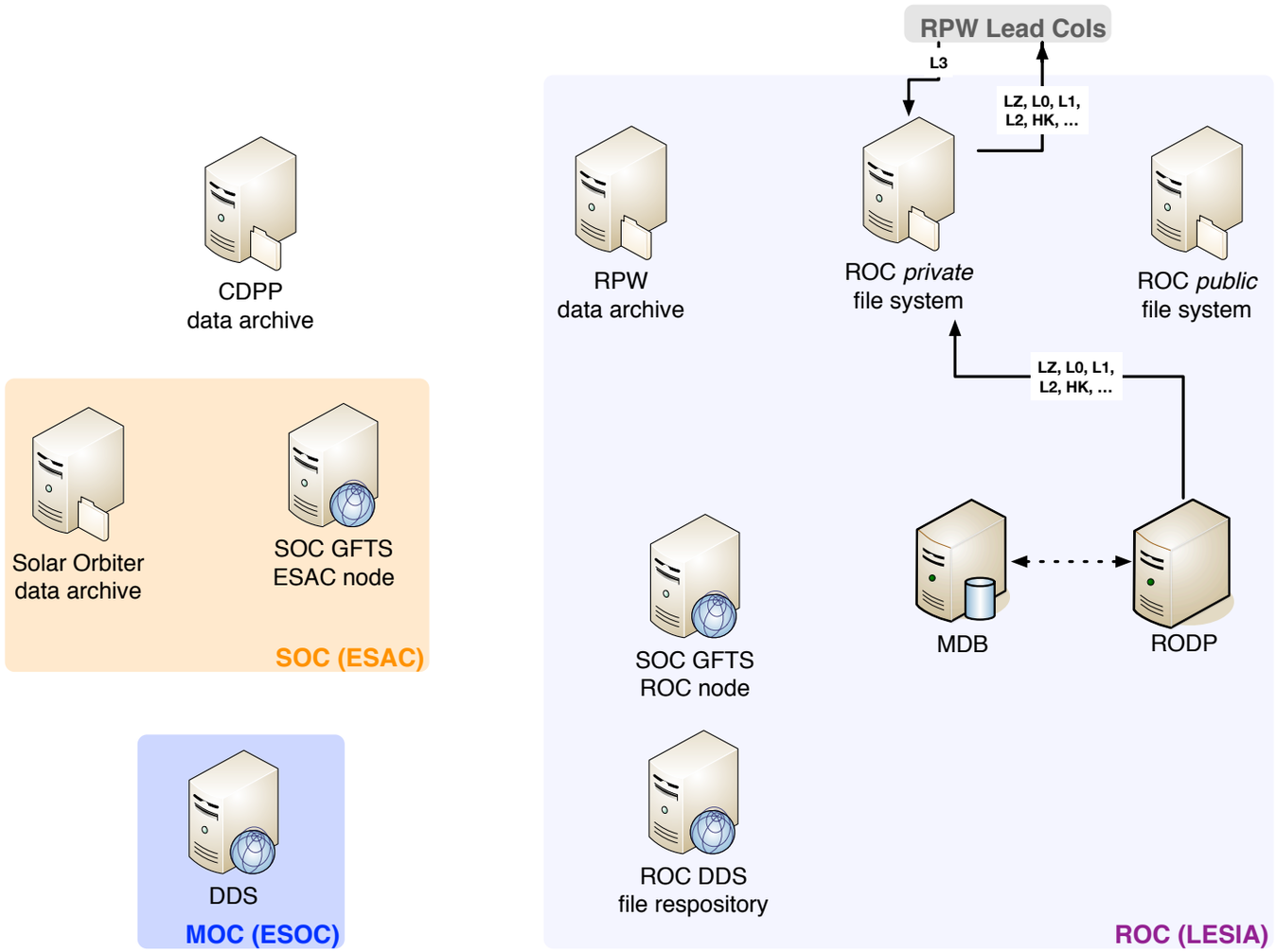
1. Operations inputs (E-FECS, TMC) from SOC
2. Ancillary data (SPICE kernels and CDF-digest) from SOC
3. TC/TM report data from MOC
4. TM raw data from MOC



- Data are checked and saved [TMC, E-FECS, ANC] in the appropriate folder in the *private* file system or stored directly in the mission database (MDB) [TM raw, TM/TC reports, E-FECS, TMC]

- TM raw data are converted into LZ XML data file

RODP nominal use case - Step #2



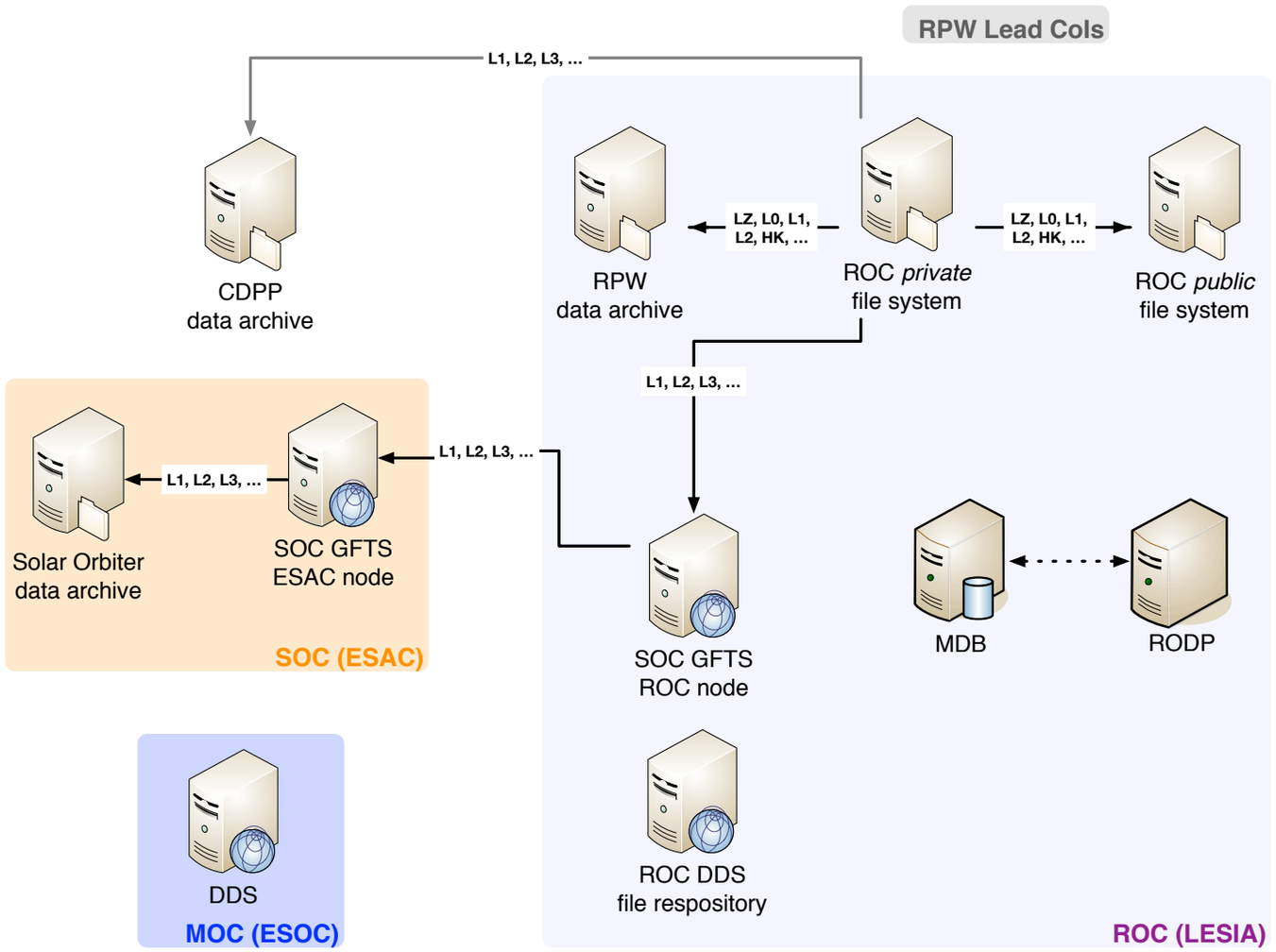
- Produce L0, L1, L1R, L2, L3 data files

1. Produce L0, L1 and HK data files from a parent LZ file
2. Produce L1R and L2 non-WF data files from parents L1
3. Produce L2 WF data files from parents L1R
4. Retrieve L3 data files produced by RPW lead CoI teams (TBD)

- Systematic input/output data file verification before/after the production
- Invalid/corrupted files are discarded
- Valid data files are stored into the ROC *private* file system as soon as the production has ended
- Summary plots are generated next to the file creation
- System allows data re-production (with restriction, controls and data versioning)

RODP nominal use case - Step #4

- Disseminate/archive data



1. For data files that reach the end of the proprietary period (3 months)

2. Push the data files into the ROC public file system

3. Archive the data files at the RPW data archive at LESIA

4. Archive the data files at the Solar Orbiter data archive

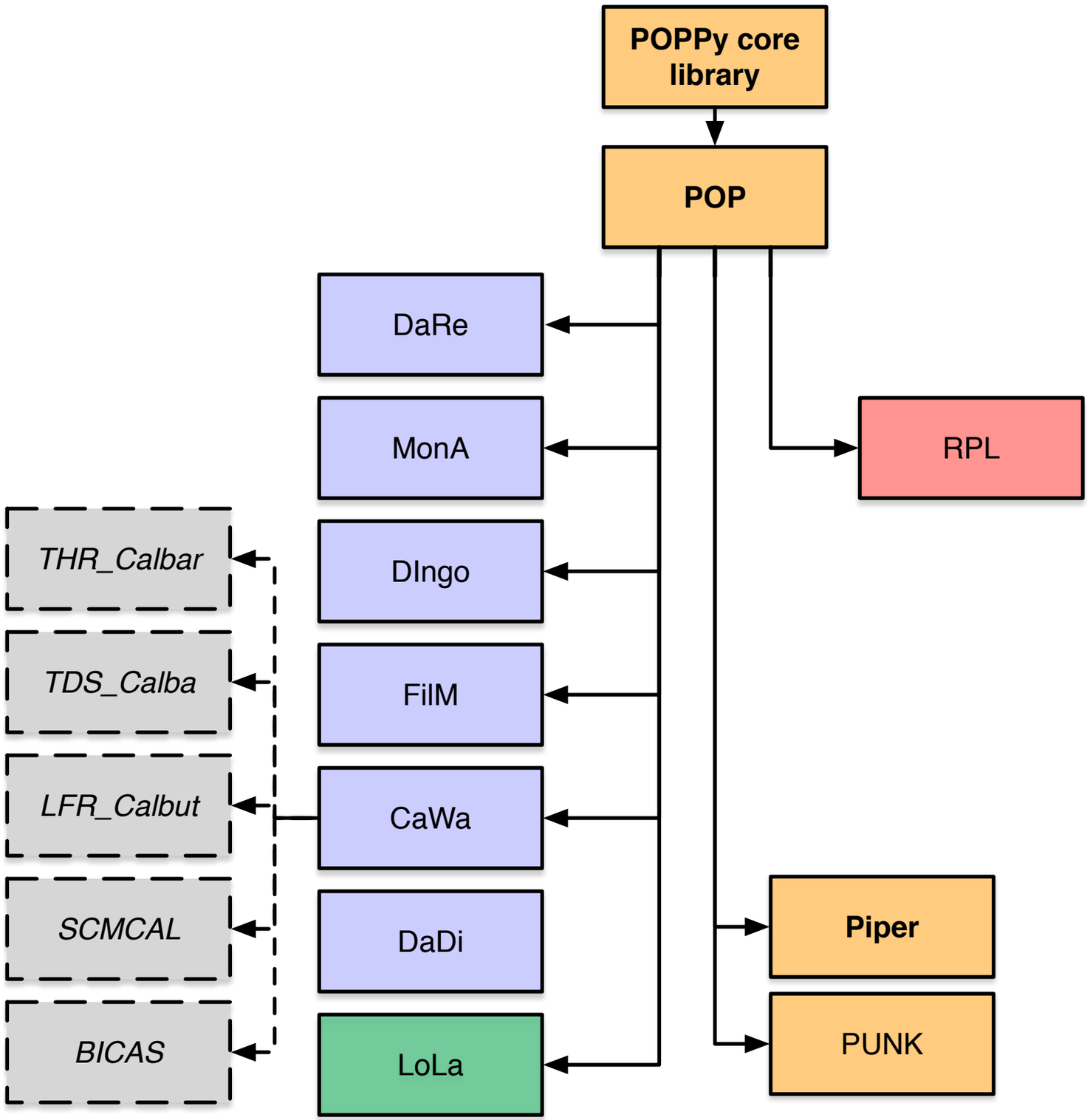
5. Archive the data files into the CDPP data archive (Toulouse)

- Archived data as tagged as *definitive*

RODP nominal use case - Expected execution frequency

Type of RODP run	Executed steps	Execution time	Motivations
Pass run	1, 2, 3 and 4	After each pass	Ensure to have the newest RPW data downlinked from the latest pass. Note that the TM packets from a pass might by available with some delay via the DDS. The RODP must hence also request TM from the previous pass.
Daily run	1, 2, 3 and 4	Every day at 2am (TBC; the exact hour may take account of the schedule of passes)	Generate complete daily data files for the previous day (assuming here that all o the TM have been received). Re-process daily data files, which are incomplete – because all of the TM had not been retrieved - or need upgrades (e.g., due to calibration table refining for instance) Ensure the data distribution and archiving at least every 24h.
Single run	2,3, 4	On demand	In the very specific case where a re-processing of the totality or a part of data archive is required.

RODP static architecture



- Main RPW data processing pipeline
- Plugin-oriented architecture, based on the POPPy framework for Python 3
- Each plugin has an assigned function (retrieved data, produce file, distribute data, etc.).
- Standardized data exchange interface between plugins
- Monitor of the executed processes/ jobs and of plugin inputs/outputs using the *Tasks* and *Targets* POPPy concept
- Control the pipeline calling with the *Commands* POPPy concept
- Design based on the Python package mechanism

RODP functional architecture: POPPy workflows

- *Tasks* of plugins can be combined to build *workflows* (i.e., sequential run of jobs)

- Example:

  - Task A —> request RPW TM raw data from the DDS for a given time rang

  - Task B —> check TM data into the local DDS repository

  - Task C —> analyse TM data and insert into the MDB

  - Task D —> generate RPW LZ file

- *workflows* can be called from the terminal in a standard ways using C*ommands* classes

Example

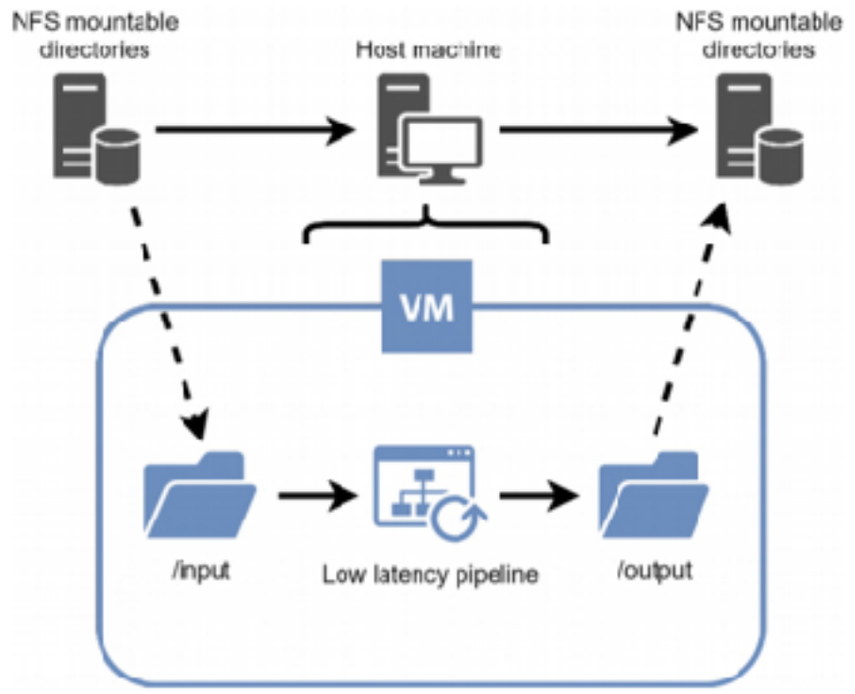
Schema


```
A = TaskA()
B = TaskB()
C = TaskC()
D = TaskD()
pipeline | A | B | C | D
```

```
pop dare retrieve --dds all --backup-dds --start 2022-01-12T20:50:39 --end 2022-01-14T13:43:22
```

RPW Low Latency Virtual Machine (LLVM)

- Virtual machine to be delivered to the SOC
- Contain an instance of the RODP, running in a very specific mode (stateless, dry run) to produce RPW LL01 CDF data files from input TM
- Primary instance autonomously deployed and run at SOC
- Backup instance at LESIA in close prod. environment
- Specification defined in SOL-SGS-TN-0006

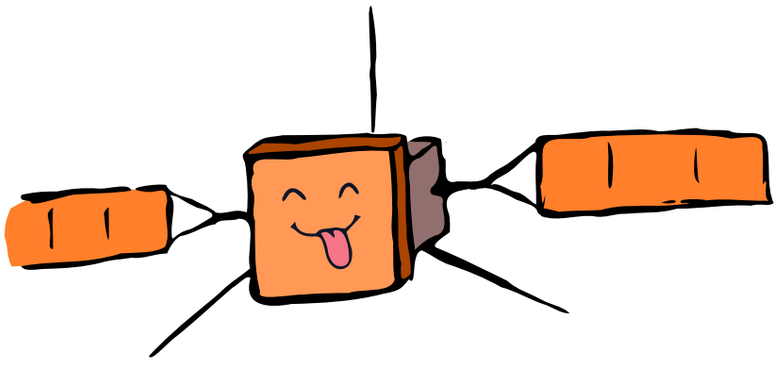


RPW Calibration Software (RCS)

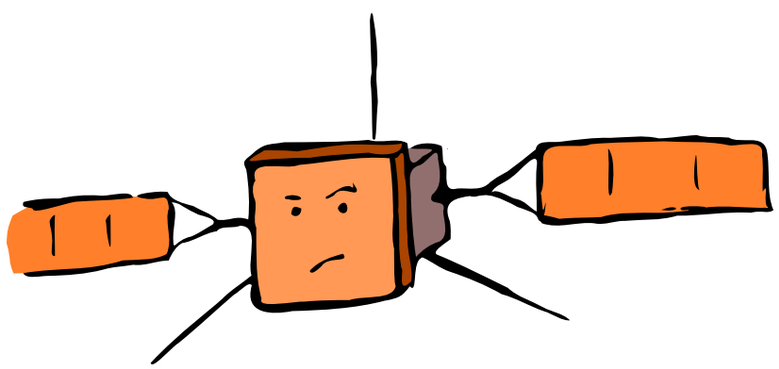
RCS name	Function	Team in charge	Programming Languages
THR data CALiBrAtion SoftwaRe (THR_CALBAR)	Produce TNR-HFR L2/L2S calibrated science data files in the CDF format	TNR-HFR team (LESIA, Meudon)	IDL
TDS data CALiBrAtion Software (TDS_CALBA)	Produce TDS L1R/L2/L2S calibrated science data files in the CDF format. L2/L2S only concern the non-waveform (WF) data products.	TDS team (IAP, Pragues)	IDL
LFR data CALibration UniT (LFR_CALBUT)	Produce LFR L1R/L2/L2S calibrated science data files in the CDF format. L2/L2S only concern the non-WF data products.	LFR team (LPP, Palaiseau)	Python
SCMCAL	Produce TDS/LFR L2/L2S magnetic WF calibrated data files in the CDF format	SCM team (LPC2E, Orléans)	IDL
BIas CAlibration Software (BICAS)	Produce TDS/LFR L2/L2S electrical WF calibrated data files in the CDF format. The BICAS is a part of the IRFU_MATLAB package.	Bias team (IRF, Uppsala)	Matlab

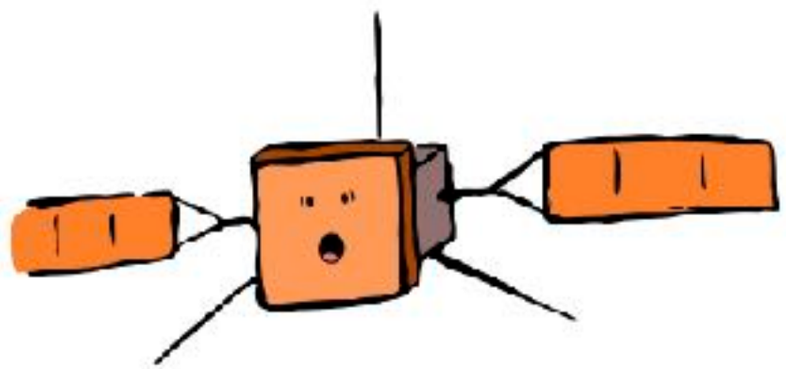
Outlines

1. RPW data products overview
2. ROC data processing subsystem (DPS) design
3. RPW Calibration Software (RCS) and data products
   - 3.1. TDS calibration software (TDS_CALBA)
   - 3.2. LFR calibration software unit (LFR_CALBUT)
   - 3.3. THR calibration software (THR_CALBAR)
   - 3.4. Bias calibration software (BICAS)
   - 3.5. SCM calibration software (BICAS)



Extra slides





RODP nominal use case - Step #2

- Produce L0, L1/HK data files

    1. LZ parent file pushed into the RODP / INPUT directory

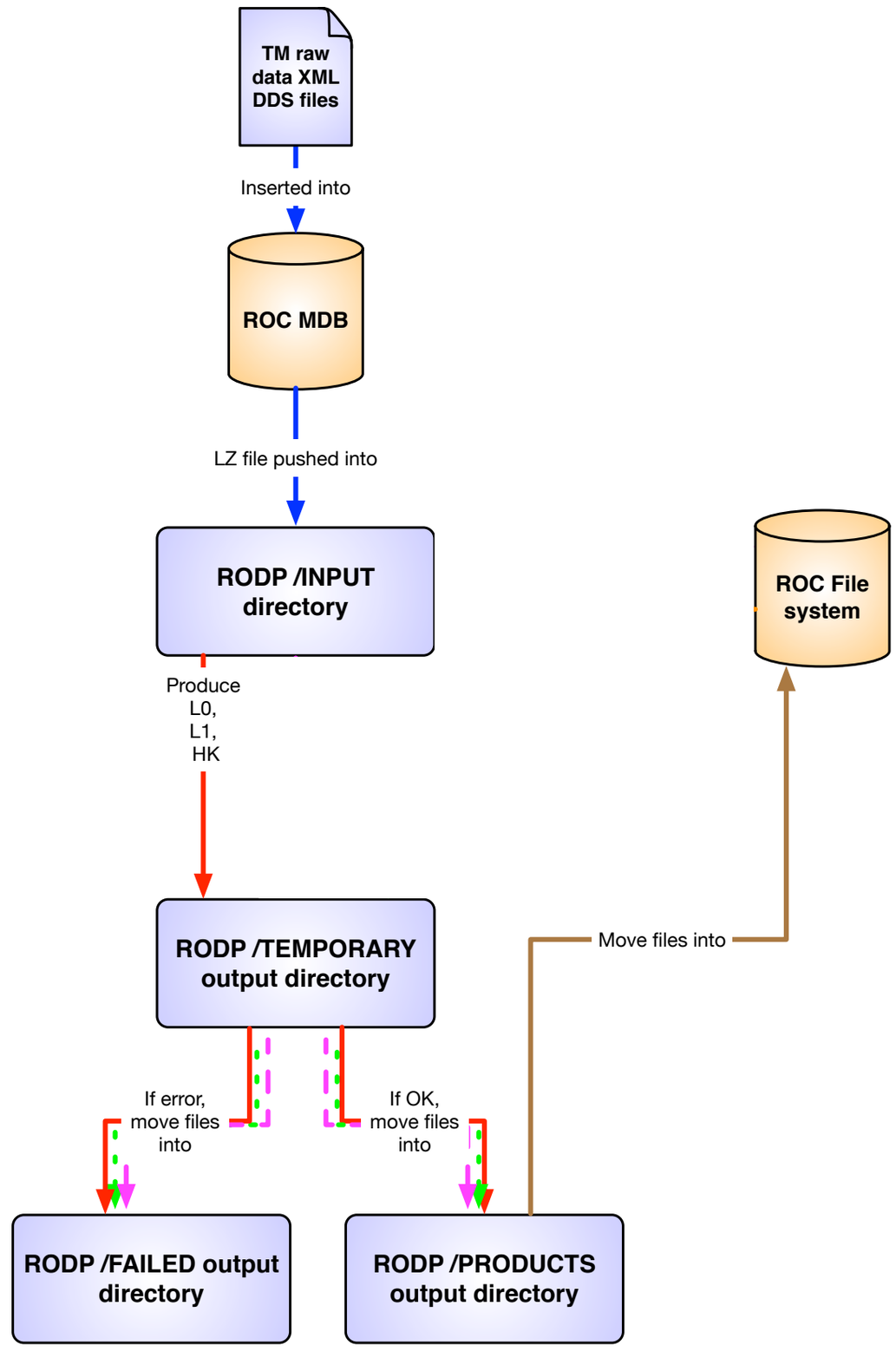
    2. L0, L1 and HK child files are written into the RODP /TEMPORARY output directory

    3. If files OK, moved into the /PRODUCTS output directory. Otherwise dropped in / FAILED

- Output files in /PRODUCTS are transferred in the ROC *private* file system

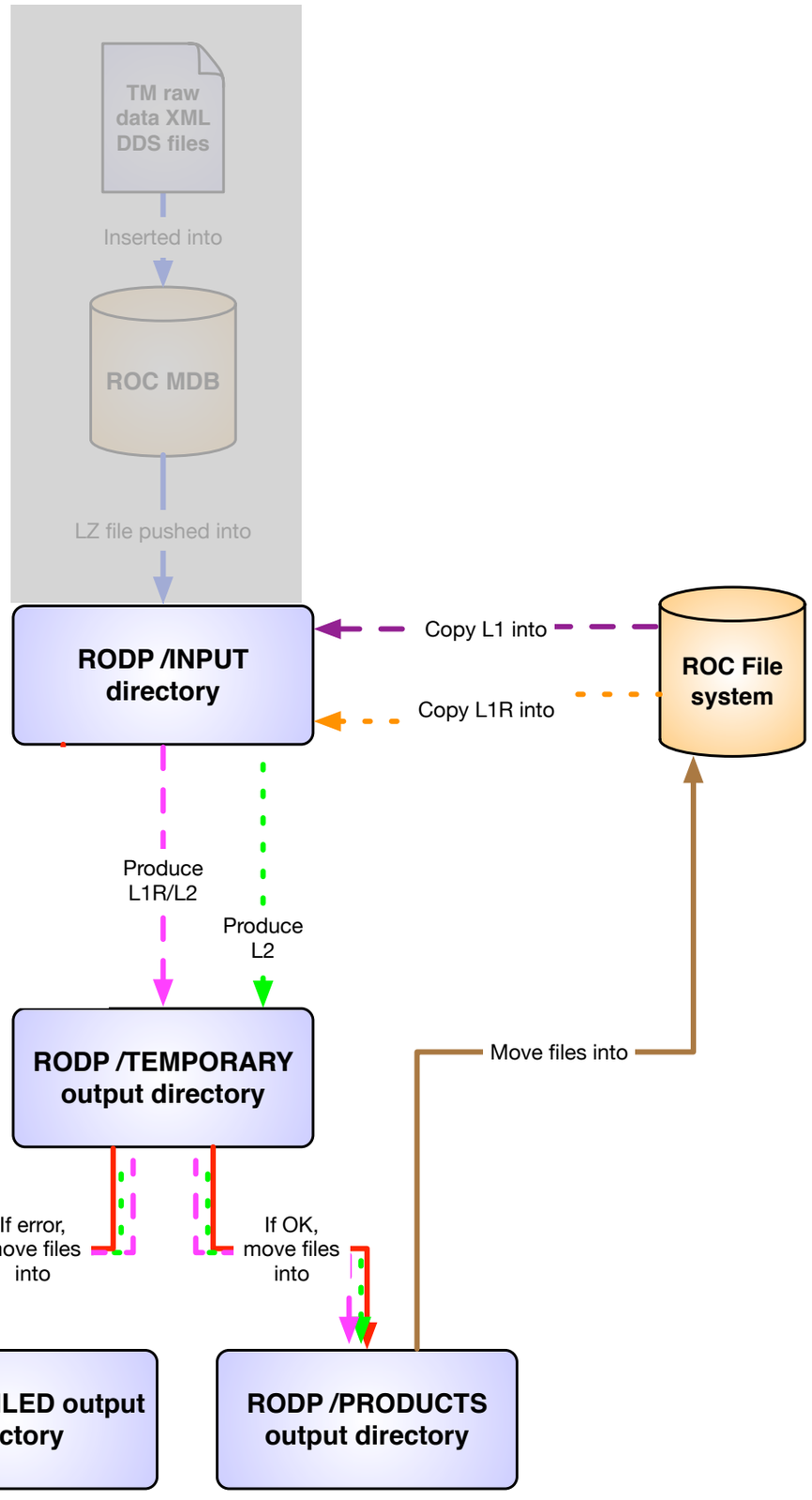
- Summary plots are generated in the same time

- RODP performs data verifications before/after each process



RODP nominal use case - Step #3

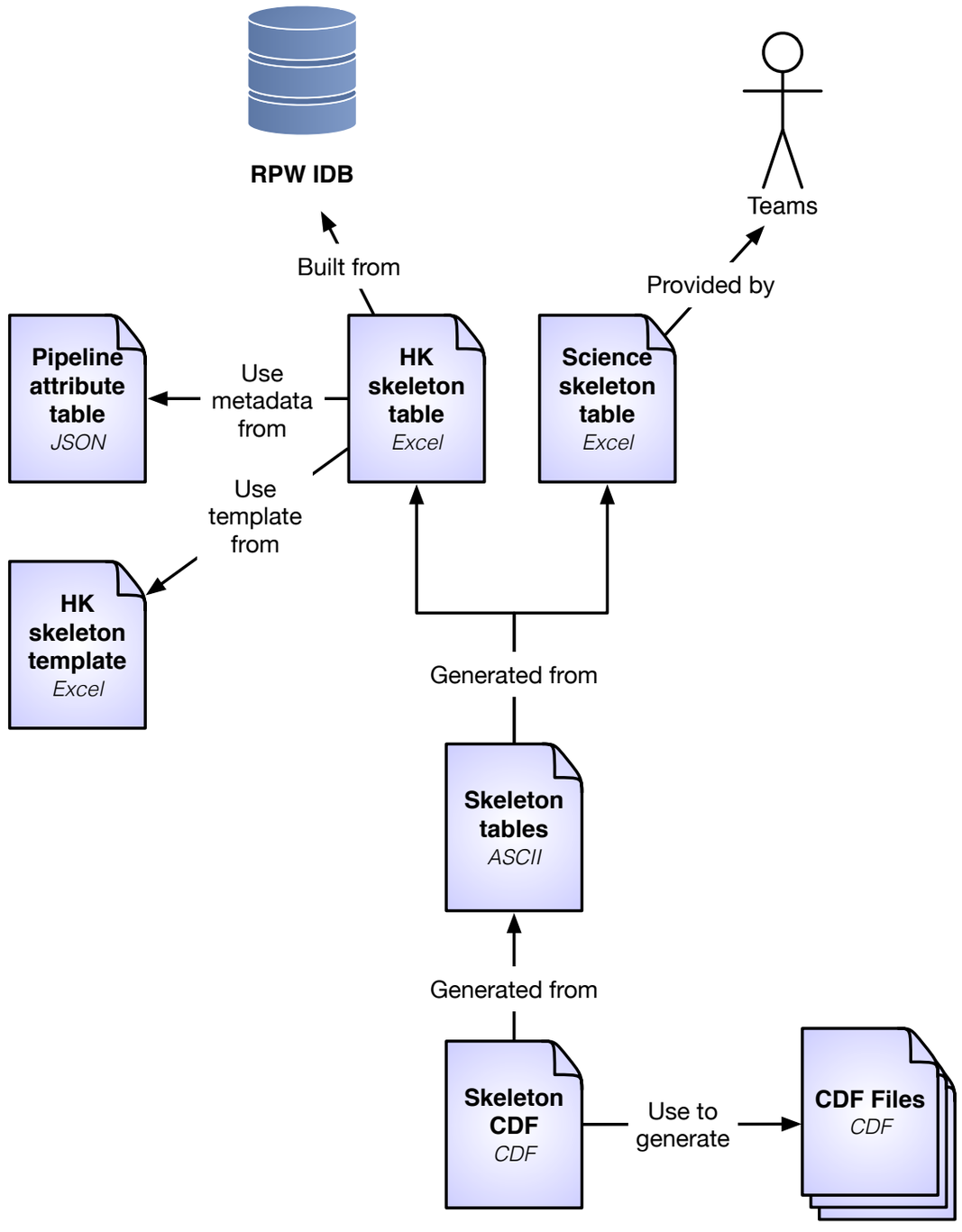
- Produce L1R/L2 data files
  - L1 parent files are copied into the /INPUT directory
  - L1R and L2 (non-WF) child files are written into the /TEMPORARY output directory
    - If files OK, moved into the /PRODUCTS output directory. Otherwise dropped in /FAILED
- Output files in /PRODUCTS are transferred in the ROC *private* file system
- Same mechanism is applied to generate L2 WF from L1R WF
- Summary plots are generated in the same time
- RODP performs data verifications before/after each process



CDF data production workflow

CDF skeleton philosophy

- CDF data files produced using CDF template files called « CDF skeleton » (or « master CDF »)
- ROC and RPW teams produces Excel format files which are converted into CDF skeleton, using a dedicated program*
- CDF skeletons of HK are automatically generated from the Instrument database (IDB) by the ROC-SGSE
- There is one skeleton per ROC data set
- List of ROC-SGSE CDF skeletons are stored in the ROC SVN repository
- Structure and content of the CDF skeleton automatically check by a dedicated program* (in progress)



Solar Orbiter ancillary data

Ancillary data

- Orbit
- Attitude
- Roll
- Reference frames (S/C, instruments)
- OBT->UTC time conversion

Delivered by the SOC as SPICE kernels (but also as CDF-digest)

S/C Reference Frame

3.4.2.1 SOLO_SRF – "Spacecraft Coordinates"

SOLO_SRF (Solar Orbiter Spacecraft Reference Frame) is the name of the reference frame that is commonly referred to as spacecraft coordinates and is defined in [RD9]. It is a

Cartesian frame centred on the point of intersection of the launcher longitudinal axis with the separation plane between the launcher and the composite, and has the following components:

- X points along the longitudinal axis of the spacecraft, positive towards the heat shield.
- Z points perpendicular to the launcher interface plane, positive towards the face of the spacecraft containing the MGA and one of the RPW antennae.
- Y completes the right handed set.

By default, assuming no off-points or rolls, SOLO_SRF_X will point towards the centre of the Sun; SOLO_SRF_Z will point normal to the plane of the orbit, positive towards solar north; SOLO_SRF_Y will complete the right-handed set, opposing the direction of motion of the spacecraft.

S/C orbit digest

4.1 Spacecraft Orbit Digest

SOC will produce a single 'orbit digest' file for the entire mission from the Solar Orbiter SPK, which in turn will be based on the OEM orbit file produced by the Solar Orbiter MOC at ESOC. Since, between GAMs, Solar Orbiter's orbit is essentially ballistic, it is anticipated that this file will not need to be updated with high cadence, and as such a new version will be nominally be issued after launch and after each GAM.

The orbit digest will follow the Solar Orbiter metadata standard and its filename will have the format:

```
solo_ANC_soc-orbit_YYYYMMDD-YYYYMMDD_V01.cdf
```

Where YYYYMMDD-YYYYMMDD represents the start date and end date of the coverage of the file.

The orbit digest will contain the following parameters at 1-hour resolution:

- HCI Position, XYZ, km.
- HCI Velocity, XYZ, km/s.
- HEE Position XYZ, km.
- Spacecraft Heliocentric distance.
- Spacecraft Heliographic latitude.
- Spacecraft HCI longitude.
- Spacecraft Carrington Heliographic longitude.
- Angle between Sun-spacecraft line and its projection on the ecliptic plane.
- Angle between projection of Sun-spacecraft line on the ecliptic plane and the Sun-Earth line.

The ecliptic of J2000 will be used in defining the orbit digest file. See APPENDIX B for the definition of the HCI, HEE and Carrington Heliographic coordinate systems.