Getting ready to analyze LST data

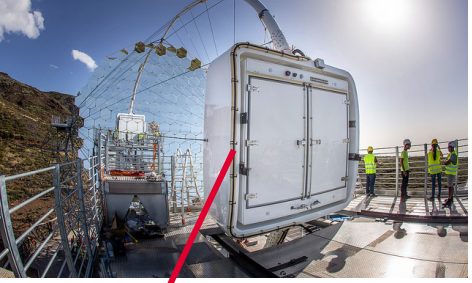
LAPP CTA, 19/12/2018

Thomas Vuillaume, on behalf of the LAPP team:
Jean Jacquemier, Pierre Aubert, Florian Gaté, Mikael Jacquemont, Gilles Maurin, Thomas Gasparetto

Xmas menu

- Summary of the analysis activities
- Some pretty images

Standard analysis chain



R0/DL0
- Simtel
- ZFITS

ctapipe

DL1
parameters
- hdf5

DL2
parameters
- hdf5

DL3
- fits

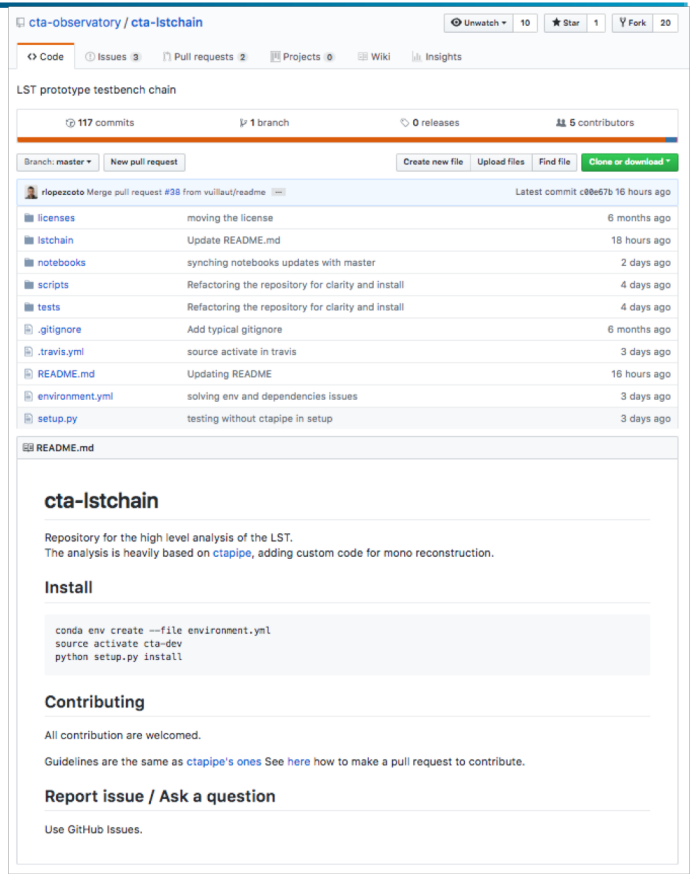
Model
(RF/BDT)
- joblib

Training
Production

cta-lstchain repository

- A side repository to prepare a low-level reconstruction for LST data
- Heavily based on ctapipe
- Custom code for mono reconstruction
- Easier and quicker to prototype here rather than in ctapipe
- Objective: move code to ctapipe when ready

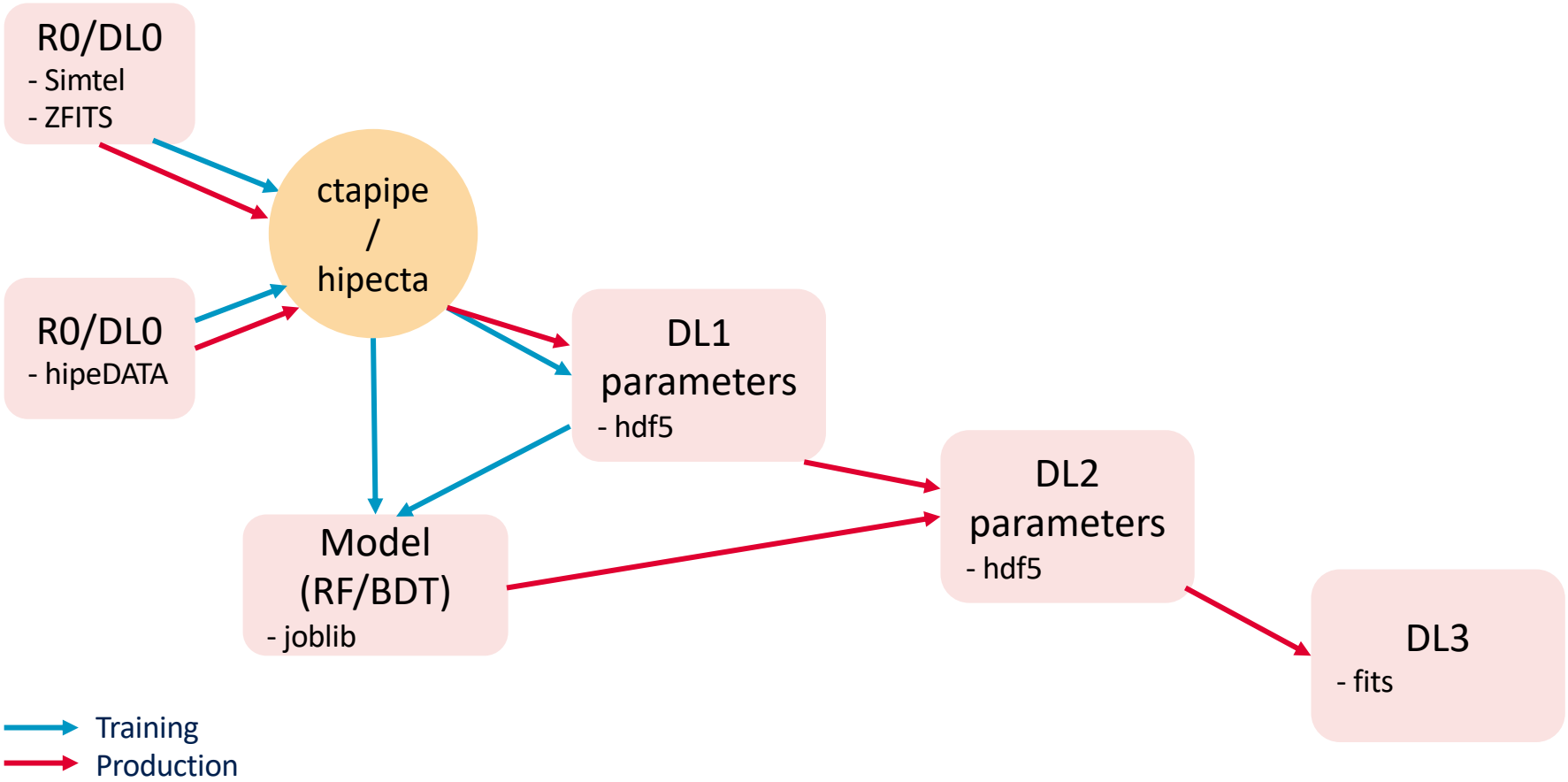
https://github.com/cta-observatory/cta-lstchain



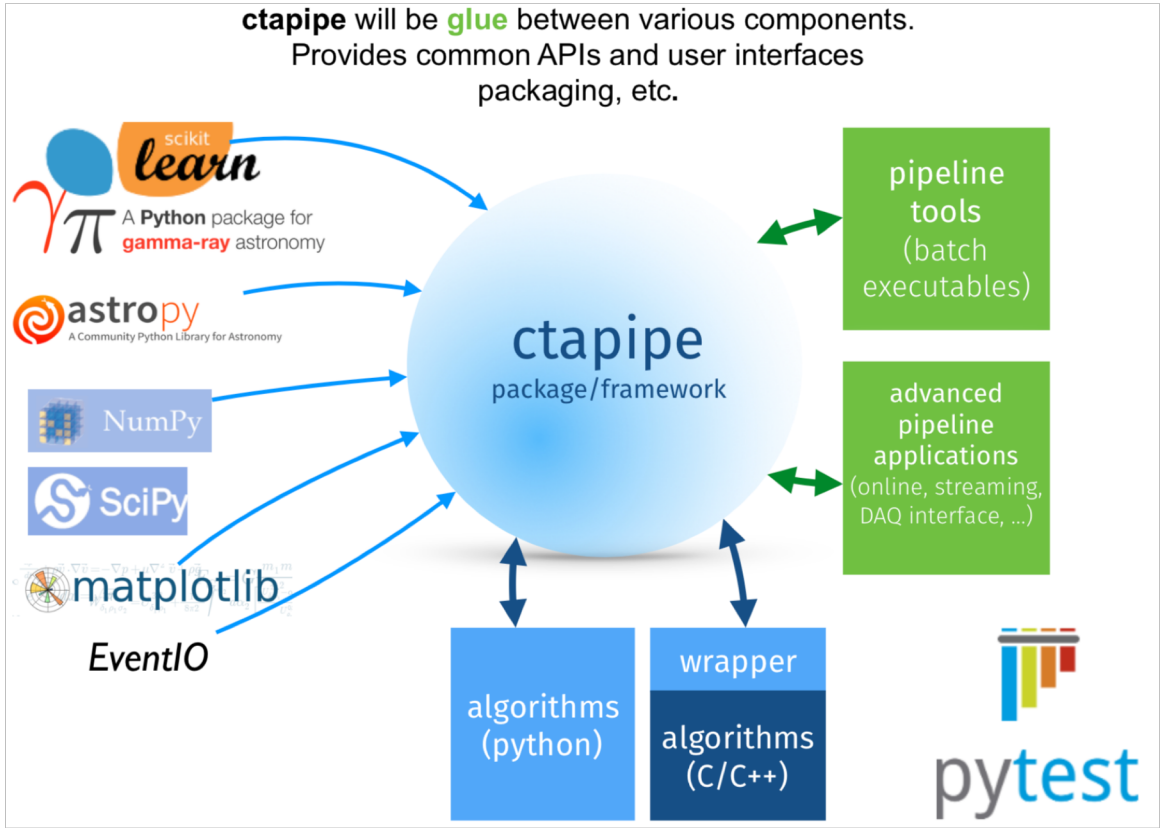
cta-lstchain repository

- Contributors:
  - Thomas Vuillaume (@vuillaut)
  - Mab Bernardos (@misabelber)
  - Ruben Lopez-Coto (@rlopezcoto)
  - Lab Saha (@labsaha)
  - Yuki Iwamura (@yiwamura)

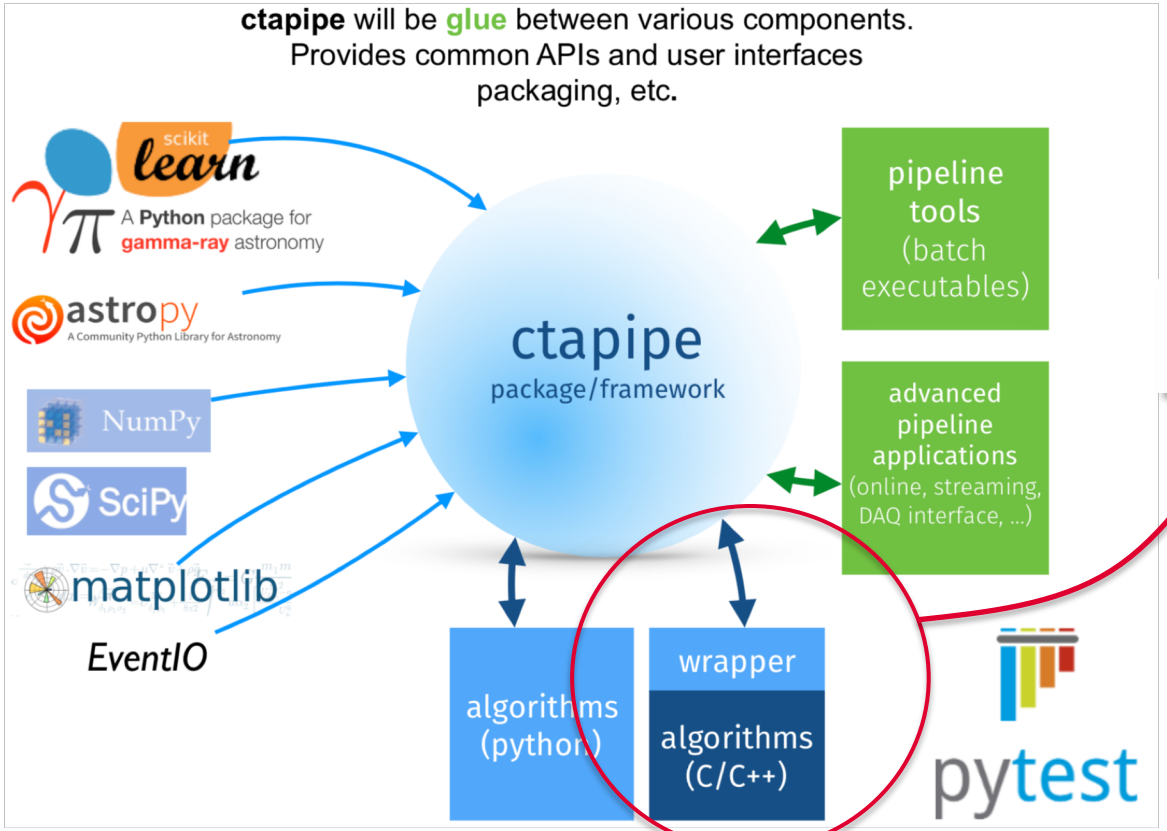
HiPeCTA: High Performance Computing for ctapipe



ctapipe


ctapipe will be glue between various components.
Provides common APIs and user interfaces
packaging, etc.
scikit learn
A Python package for gamma-ray astronomy
astropy
A Community Python Library for Astronomy
NumPy
SciPy
matplotlib
EventIO
ctapipe
package/framework
pipeline tools (batch executables)
advanced pipeline applications (online, streaming, DAQ interface, ...)
algorithms (python)
wrapper
algorithms (C/C++)
pytest

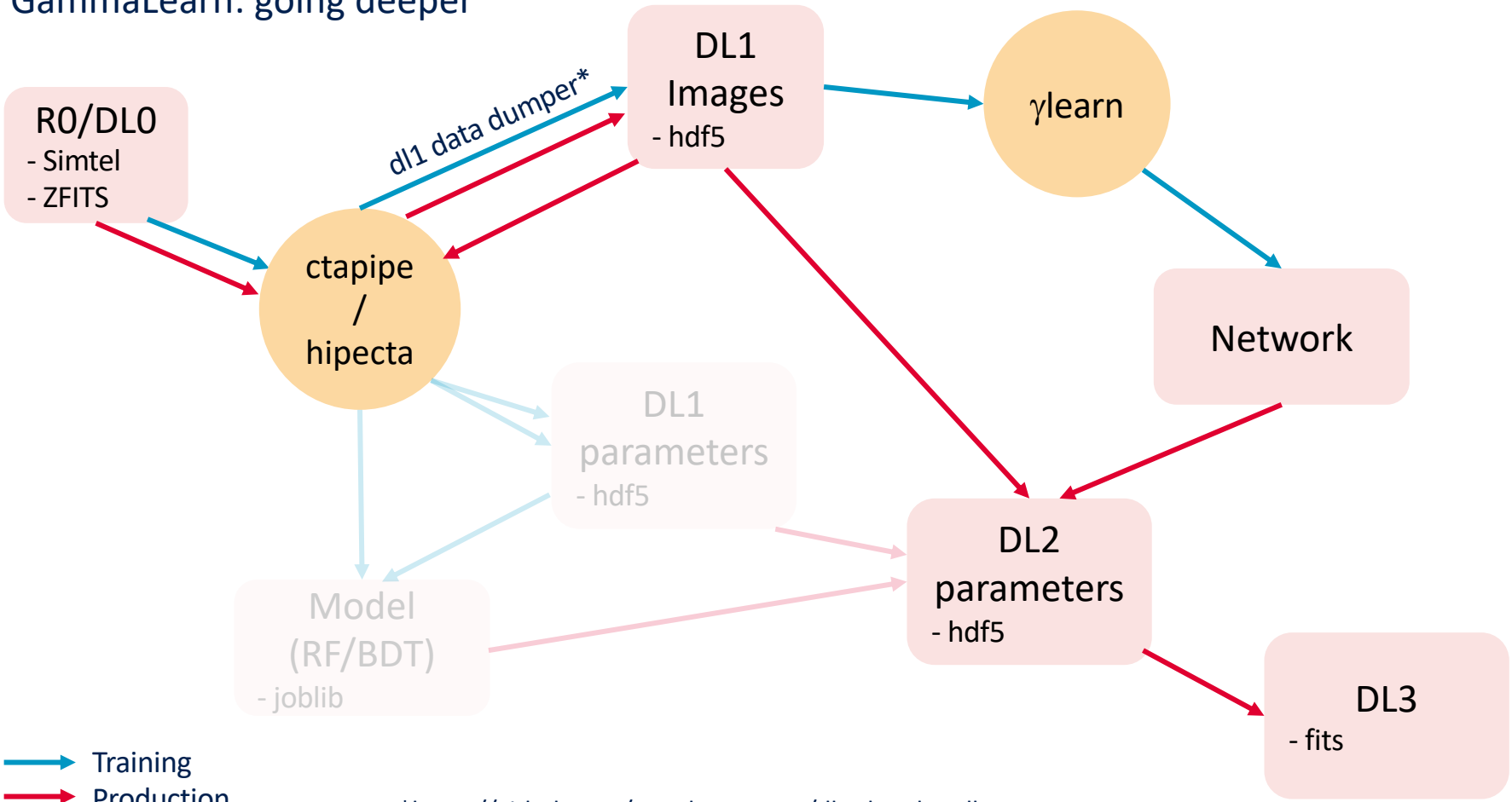
ctapipe + hipeCTA



hipeCTA

High performance algorithms for CTA

GammaLearn: going deeper



*https://github.com/cta-observatory/dl1-data-handler

GammaLearn: going deeper

R0/DL0
- Simtel
- ZFITS

R0/DL0
- hipeDATA

ctapipe / hipecta

dl1 data dumper

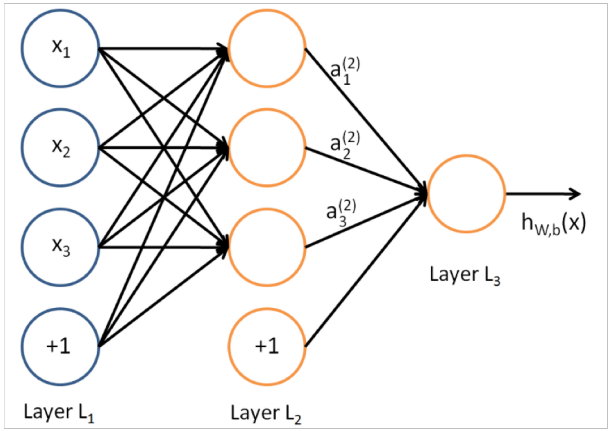
DL1 Images
- hdf5

γlearn

Network

DL2 parameters
- hdf5

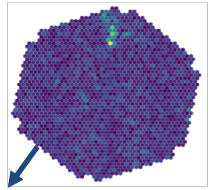
DL3
- fits

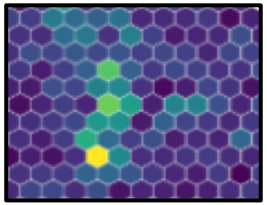


Training
Production

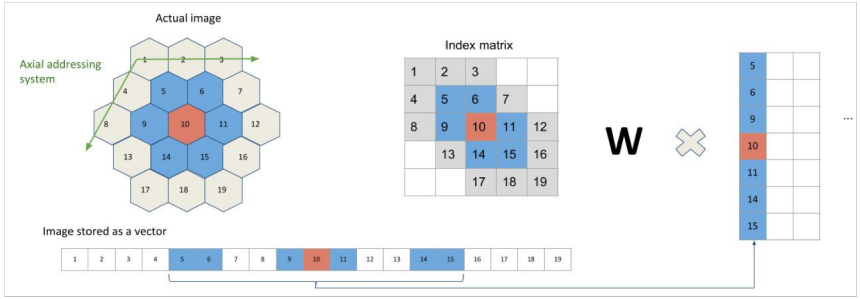
IndexedConv package for deep learning application

Processing unconventional images





- Indexed convolution and pooling
  - Validated on CIFAR and AID datasets
  - Accepted for publication @VISAPP 2019 (computer science)
  - Pros: process unaltered images, easily applicable to other detectors shape
  - Cons: slower (Python for now)
    - C++ / cuda version being developed

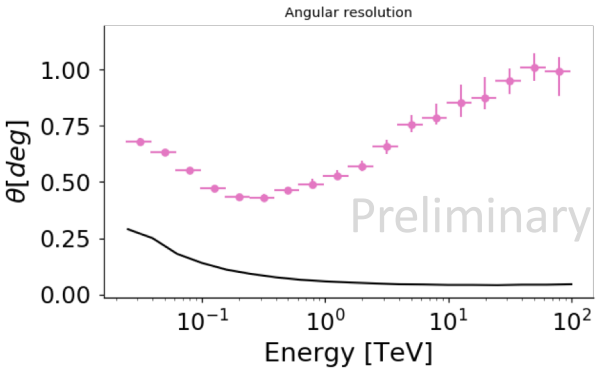


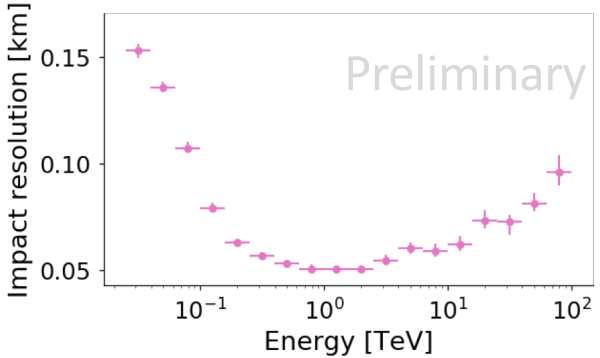
Example from CIFAR dataset

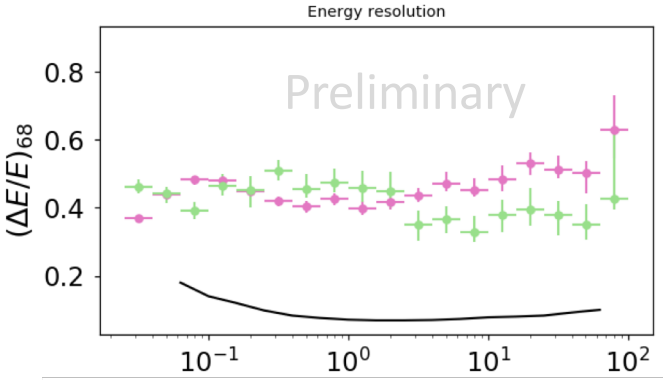
- Of interest for CTA but not limited to
  - Can be applied to any grid of pixels, given that pixel neighbours are known

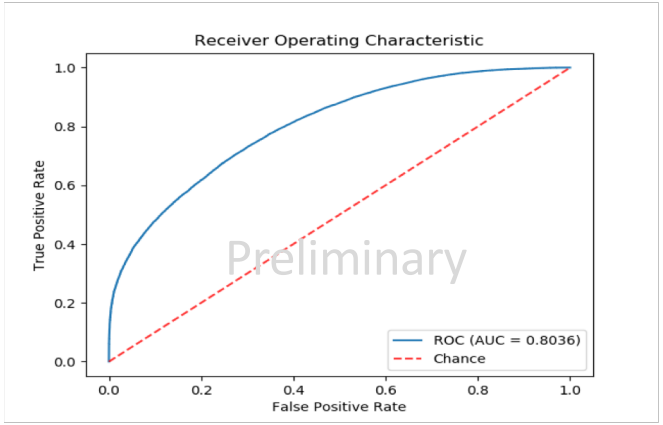
→ *https://github.com/IndexedConv/IndexedConv*

Example: Reconstruction for LST1 (mono)





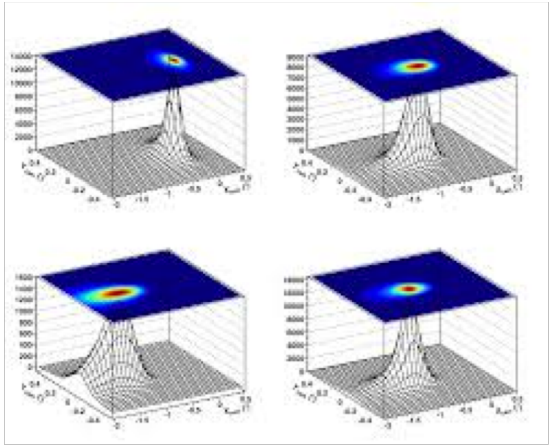




- Diffuse gamma, images selected based on tailcut cleaning
  - convnet (250 k images)
  - random forest (1 M images)

- Also working on stereo with different strategies
  - See presentations done in the group

Impact: template analysis



R0/DL0
- Simtel
- ZFITS

R0/DL0
- hipeDATA

ctapipe / hipecta

dl1 data dumper

DL1 Images
- hdf5

DL1 parameters
- hdf5

Model (RF/BDT)
- joblib

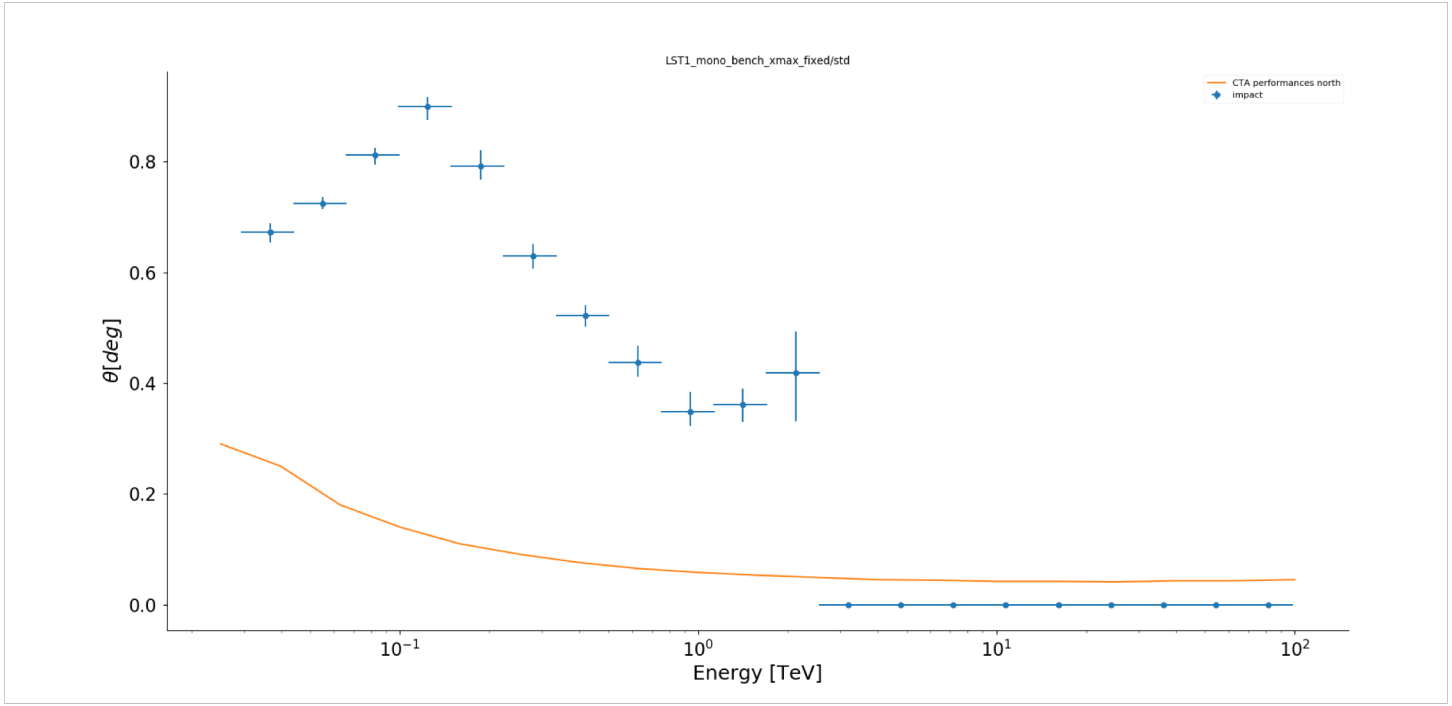
Impact

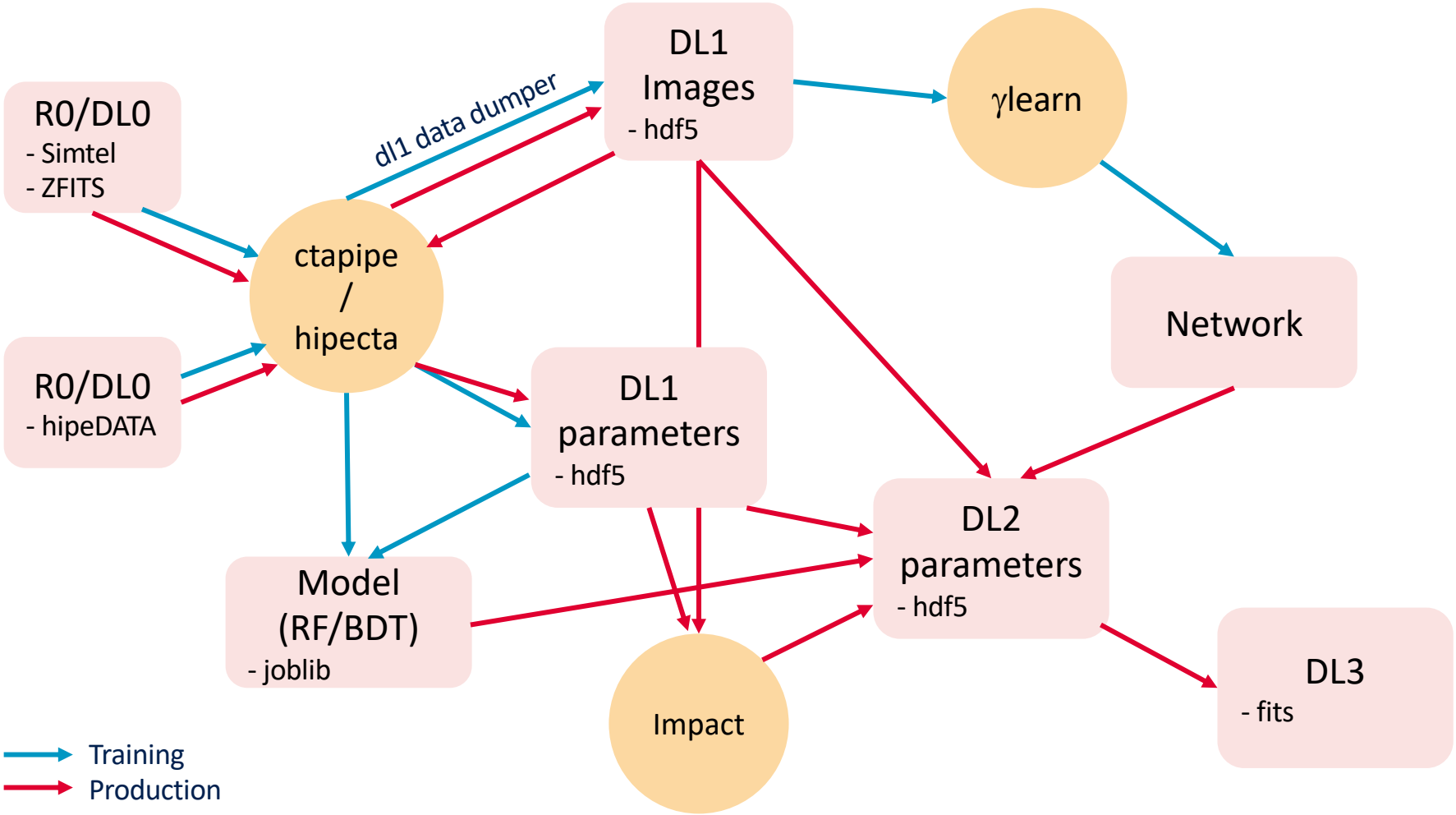
DL2 parameters
- hdf5

DL3
- fits

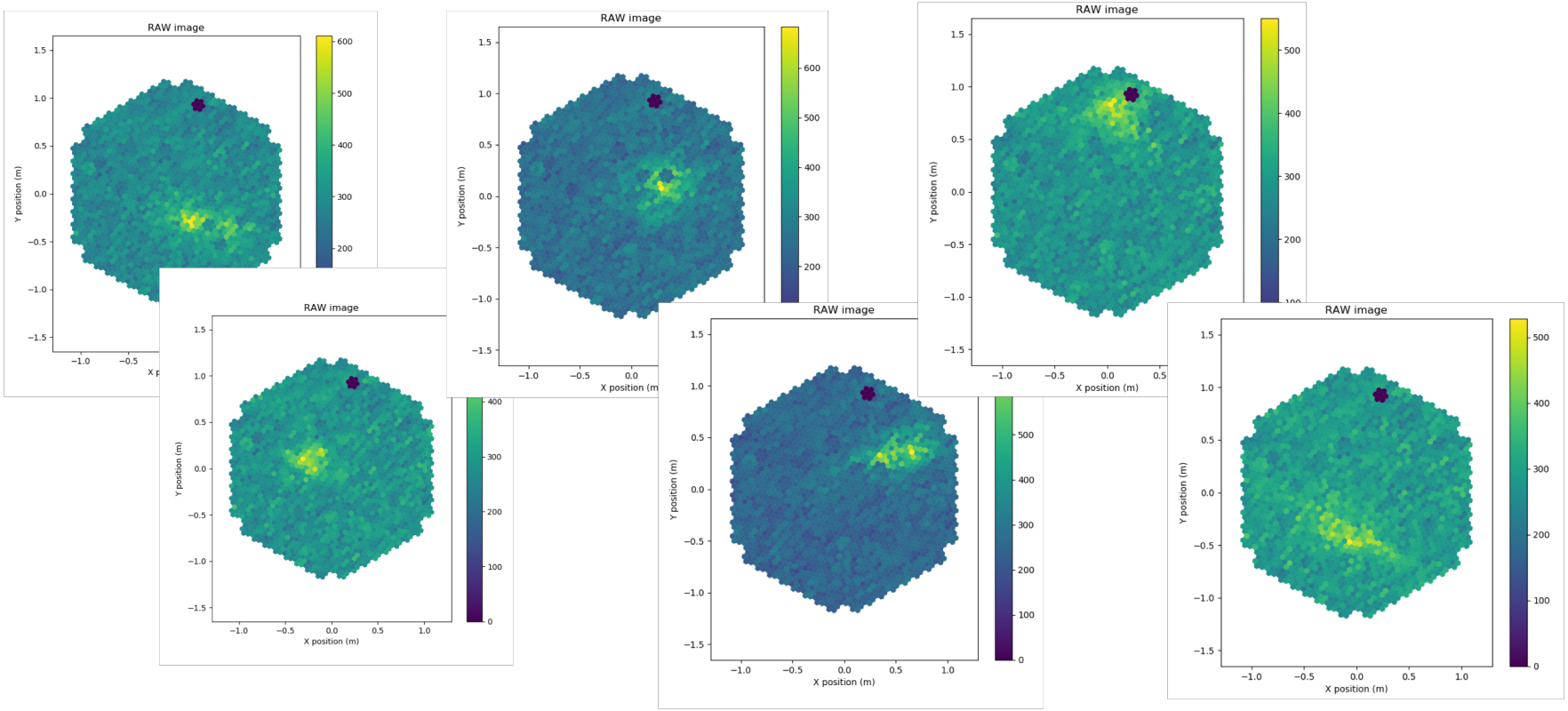
Training
Production

Impact: template based analysis

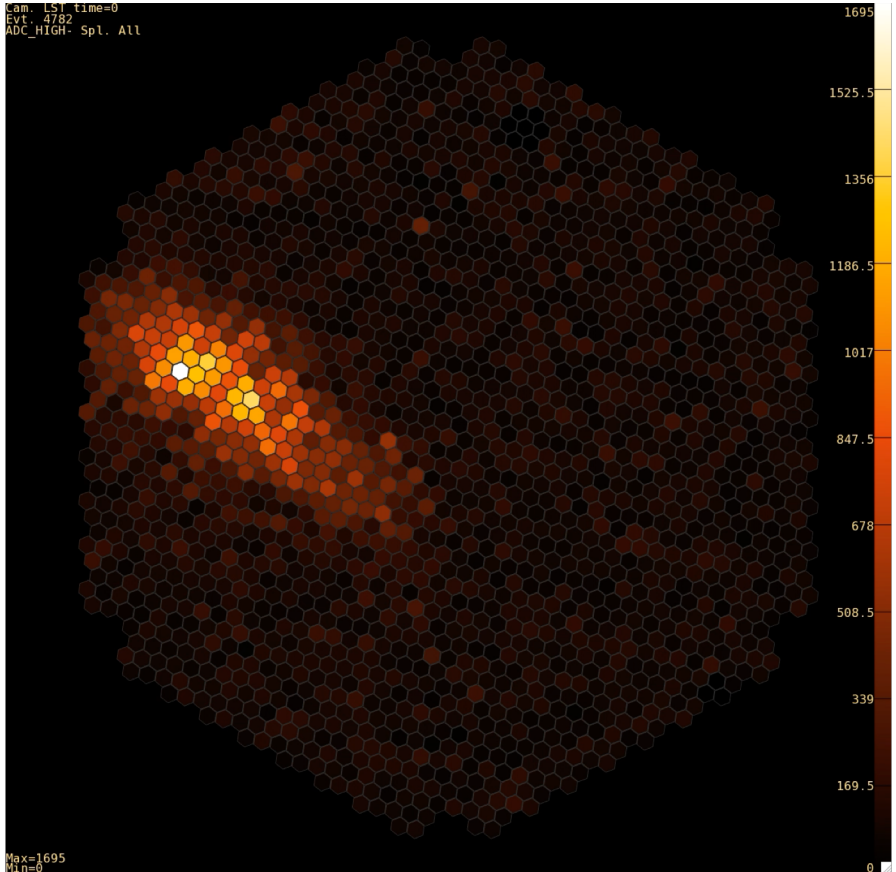


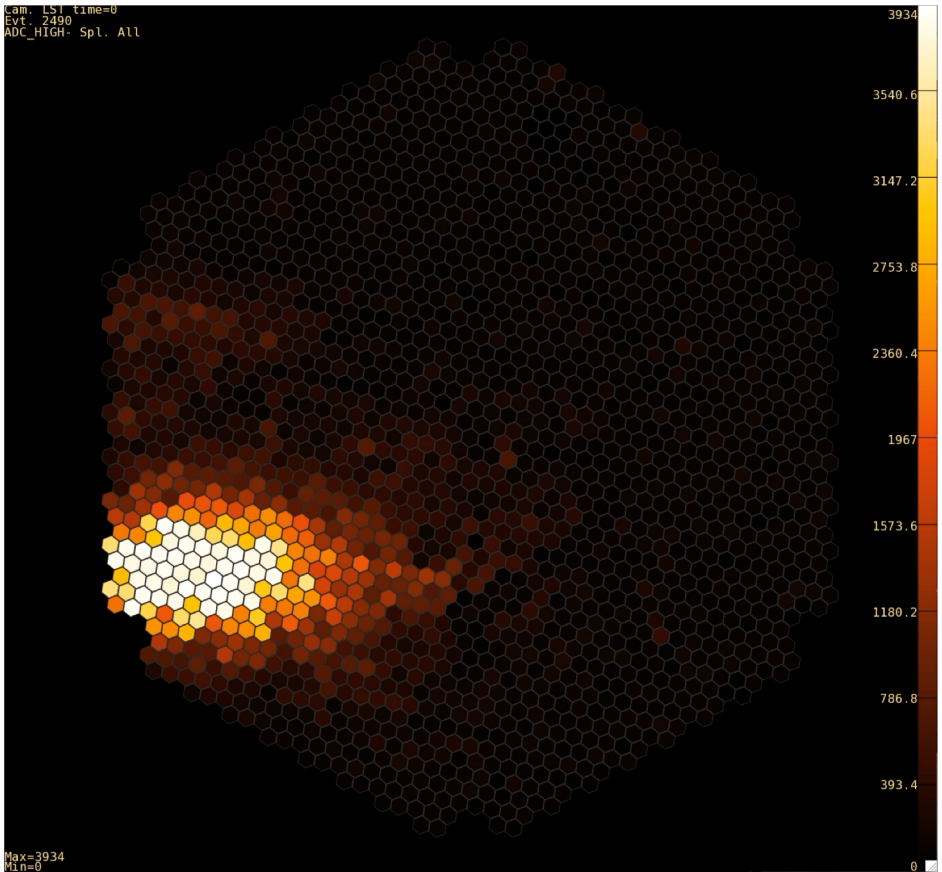

R0/DL0
- Simtel
- ZFITS
R0/DL0
- hipeDATA
dl1 data dumper
ctapipe
/
hipecta
DL1
Images
- hdf5
γlearn
Network
DL1
parameters
- hdf5
DL2
parameters
- hdf5
Model
(RF/BDT)
- joblib
Impact
DL3
- fits
Training
Production

First events !

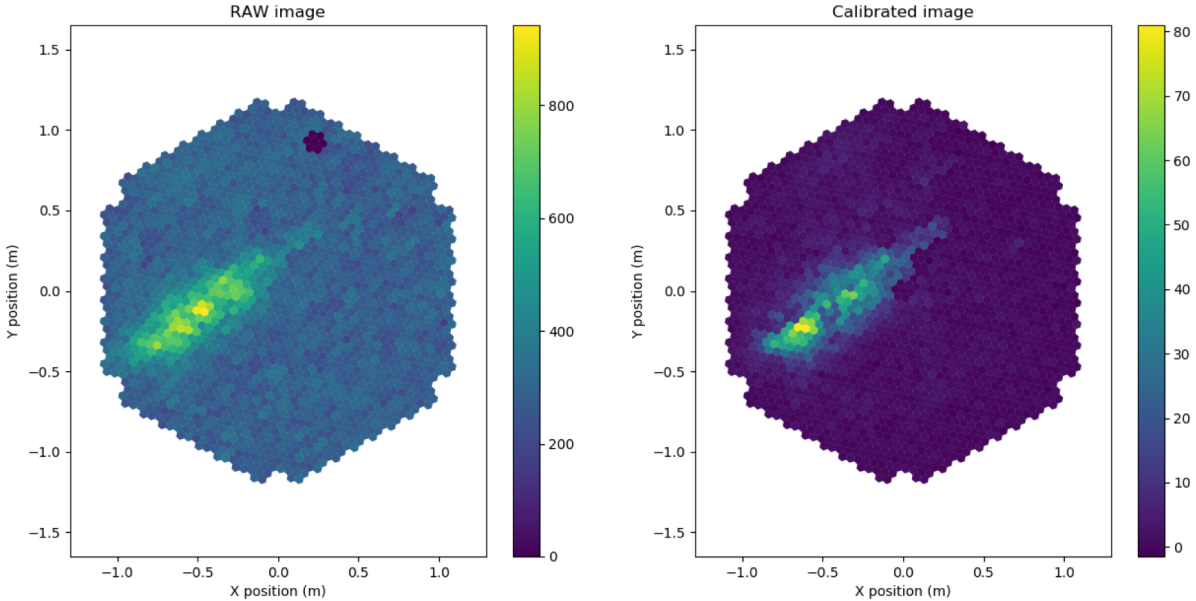


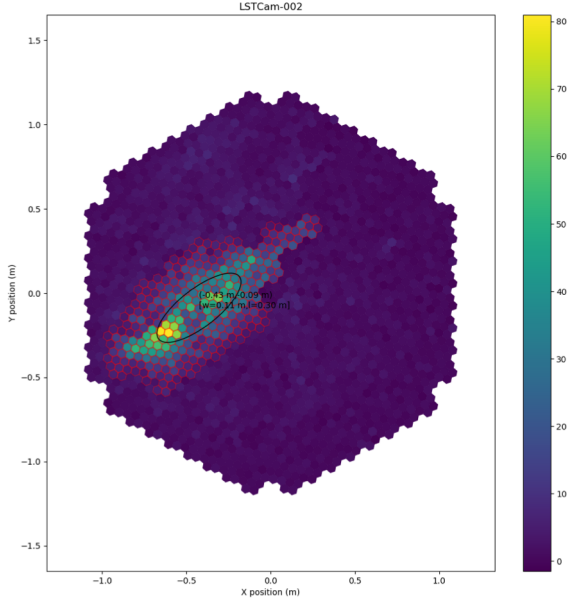
First events !



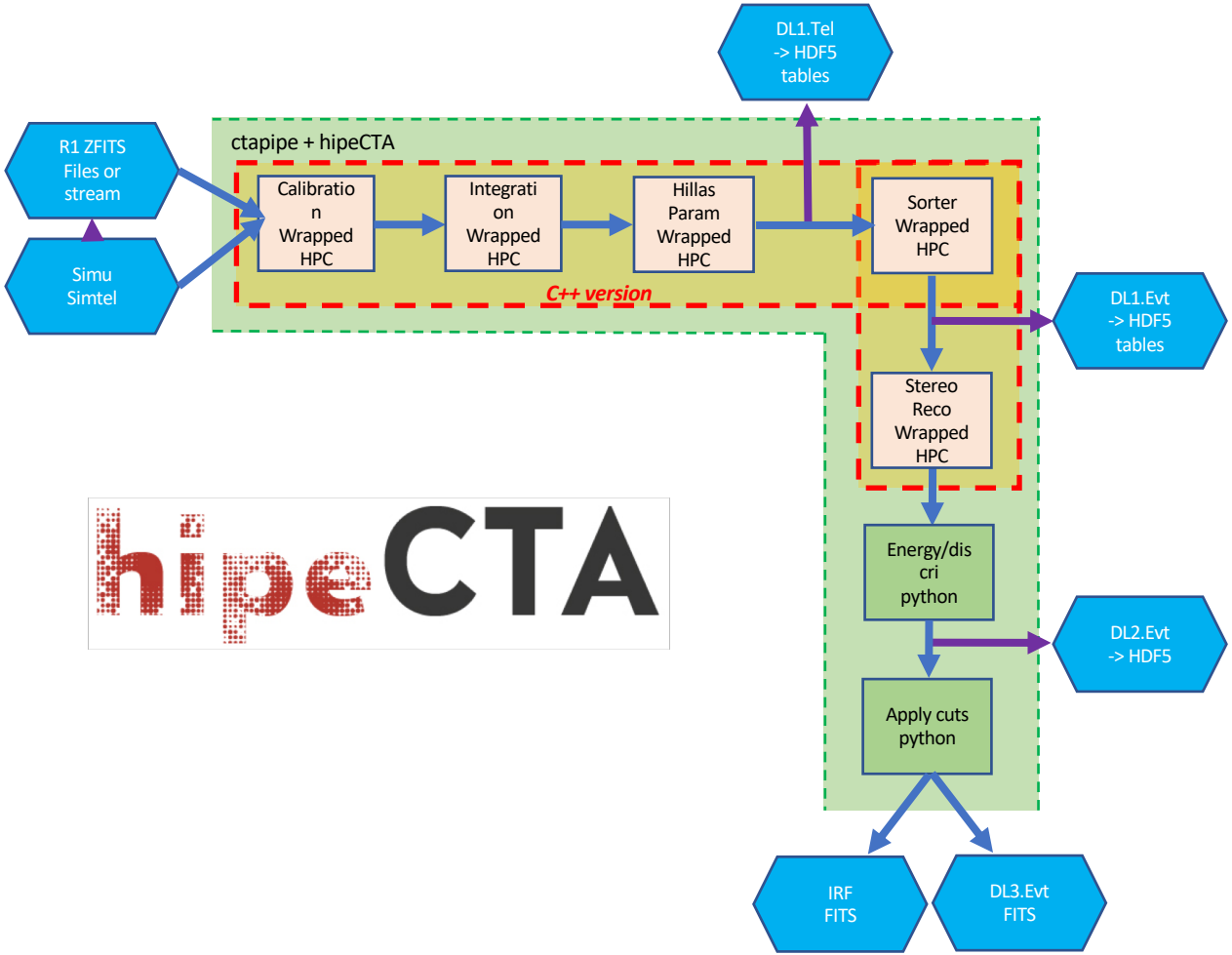


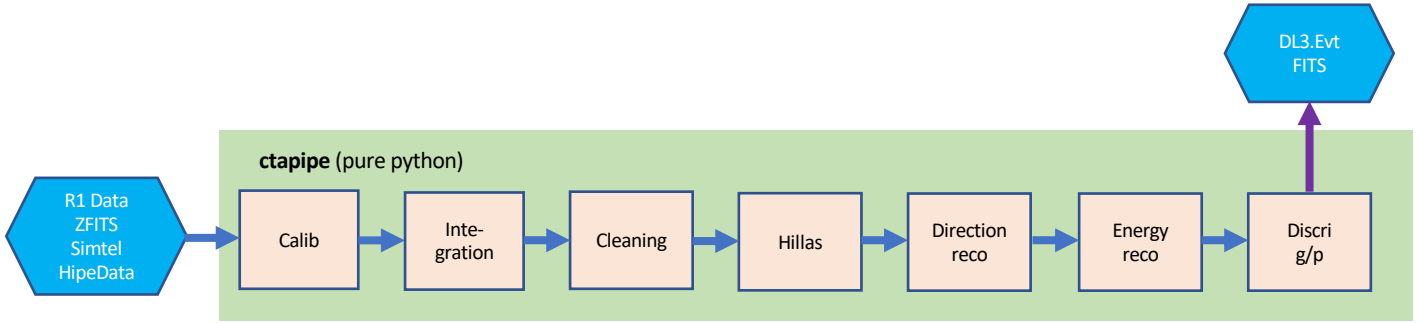
First events !

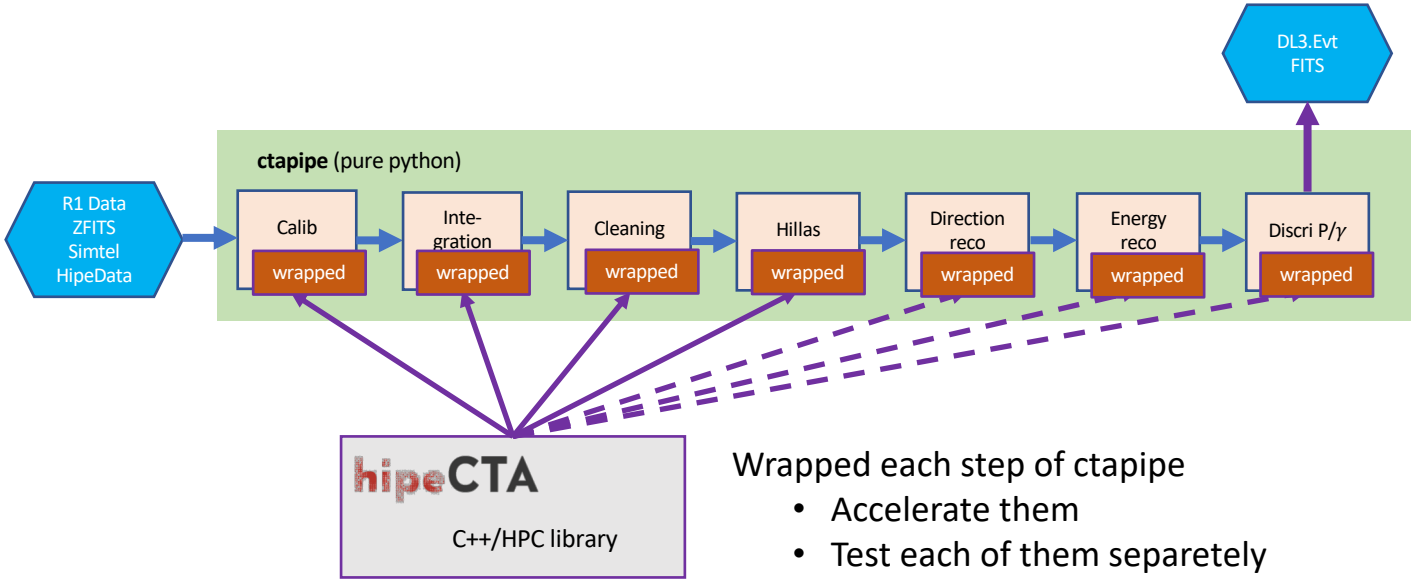




Back-up

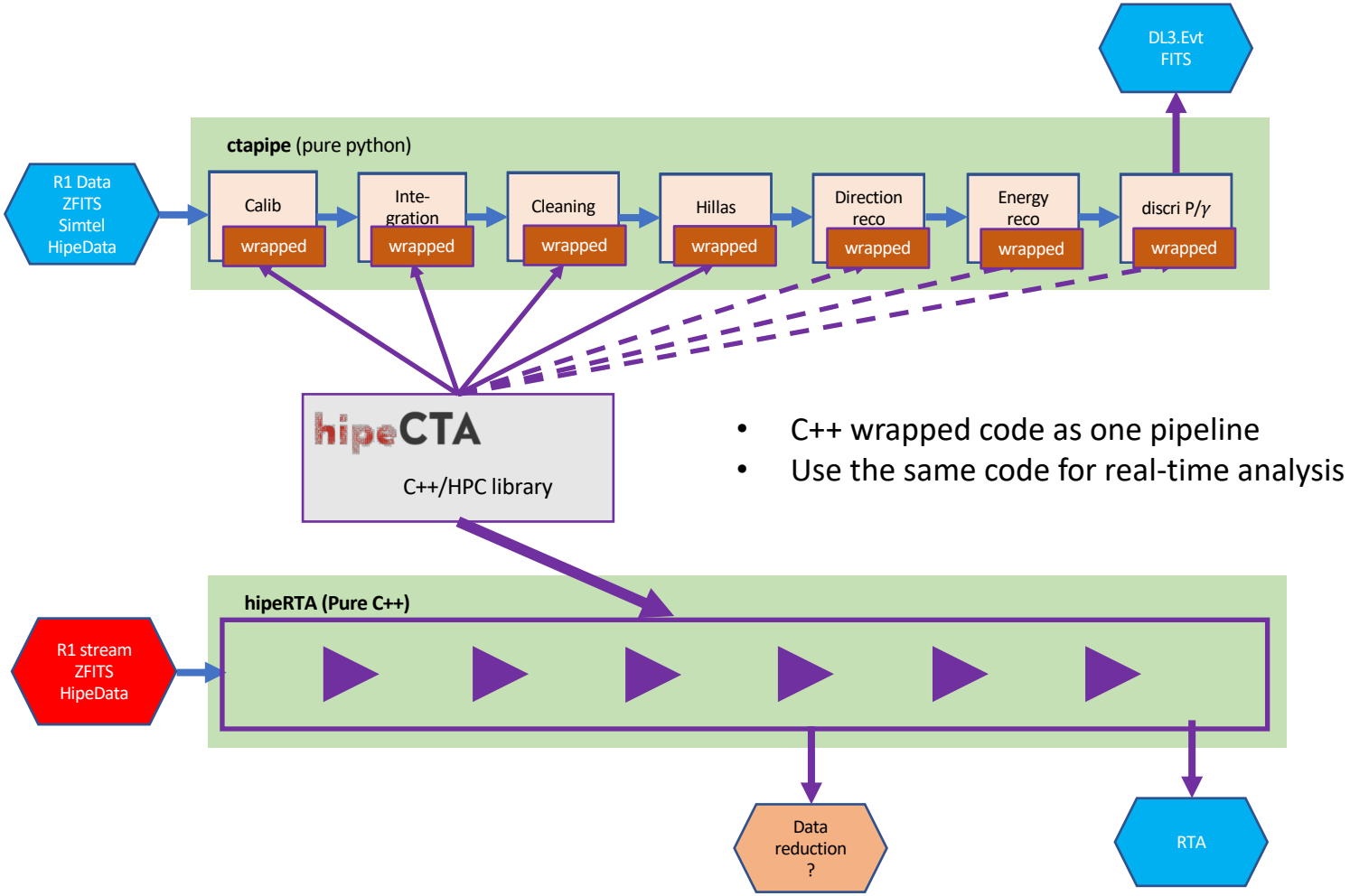

DL1.Tel
-> HDF5
tables
R1 ZFITS
Files or
stream
Simu
Simtel
ctapipe + hipeCTA
Calibratio
n
Wrapped
HPC
Integrati
on
Wrapped
HPC
Hillas
Param
Wrapped
HPC
Sorter
Wrapped
HPC
C++ version
DL1.Evt
-> HDF5
tables
Stereo
Reco
Wrapped
HPC
hipeCTA
Energy/dis
cri
python
DL2.Evt
-> HDF5
Apply cuts
python
IRF
FITS
DL3.Evt
FITS


DL3.Evt
FITS
ctapipe (pure python)
R1 Data
ZFITS
Simtel
HipeData
Calib
Inte-
gration
Cleaning
Hillas
Direction
reco
Energy
reco
Discri
g/p



Wrapped each step of ctapipe

- Accelerate them
- Test each of them separetely
- Reduce bottlenecks



- C++ wrapped code as one pipeline
- Use the same code for real-time analysis