

Introduction to Computing

Elisabetta Pennacchio, IP2I
DUNE-France Analysis Workshop
APC, April 18th, 2023



This presentation is not a tutorial: a complete tutorial already exists!

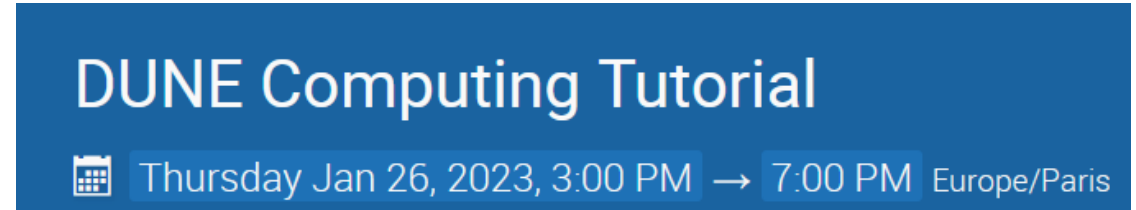
- A computing tutorial is systematically organized at every DUNE Collaboration meeting

- [Last tutorial](#) has been organized in January 2023

- Items discussed during the tutorial:

1. DUNE Computing Resources
2. LArSoft: Presentation of LarSoft (Presentation of the art framework and LArSoft, input files and most useful commands).
3. LArSoft: Demo of the analysis workflow. Hands-on to learn how to modify an existing analysis module.
4. LArSoft: Demo and hands-on
5. Grid: Submit, monitor and retrieve jobs

- For each item, a recorded video is also available



➤ The support material is updated regularly (January 2023: documentation about *jobsub_lite* and *tokens* added)

➤ checkout the the material here:

<https://dune.github.io/computing-training-basics-short/>

<https://dune.github.io/computing-training-basics-short/setup.html>

➤ This tutorial covers all the different aspects, and it is the right starting point:
ATTEND IT OR FOLLOW IT OFFLINE


This presentation aims to be:

- An overview of good practices → software, data processing
- A summary of available data (real and Monte Carlo): coldbox, FD2
- CCIN2P3: how to run analysis and develop software

The first FRANCE-DUNE workshop has been organized in December 2021

<https://indico.in2p3.fr/event/25730>

Discussion analyse DUNE-IN2P3

 lundi 6 déc. 2021, 14:00 → 17:00 Europe/Paris

Some good practices (see also [Heidi Schellmann talk](#) January DUNE CM)

Basic principles:

Science results have to be reproducible

- 1 → A tagged set of algorithms: code management, releases, configuration files, code visible by the collaboration
- 2 → known provenance of data: tracked processed chain, cataloged data
- safe delivery of data

Physics papers need to come from **accessible/versioned code** and **documented samples**

1 → A tagged set of algorithms: code management, releases, configuration files, code visible by the collaboration

1 Software

- **Art** is the official framework used by DUNE to reconstruct raw data, to run simulation campaigns and produce analysis results. Art has been chosen not only because of the features it provides, but also because it allows DUNE to use and share algorithms developed for other LArTPC experiments.
- It is very important that the collaboration can look at the prototypes/coldbox results with the standard analysis tools and be able to compare for instance with the horizontal drift data or look at the top and bottom drift CRPs, or more generally simulation results for FD HD and VD.
- The code is available and accessible to all DUNE members, versions are regularly tagged
CVMFS is used for distributing precompiled code:
 - DUNE software is in `/cvmfs/dune.opensciencegrid.org`
 - LArSoft code in `/cvmf/larsoft.opensciencegrid.org`

CVMFS is by its nature read-all so code is readable by anyone in the world with a CVMFS client : Fermilab , CERN, CCIN2P3, worker nodes....

How to run LArSoft

- LArSoft can be used at 3 different levels, each one corresponding to a different user profile (and a **different effort** required to get started)
 - 1) **B** Beginner: user only interested in reading already existing files.
 - 2) **I** Intermediate: a minimal knowledge is required
 - 3) **D** Developer: users that can modify and build the code



1) Beginner: you do not need to build code, use the DUNE one

→ Login (CCIN2P3, CERN, FERMILAB) and type:

```
source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh
```

```
export DUNEVERSION=v09_72_00d00  
export DUNEQUALIFIER="e20:prof"
```

```
setup dunesw $DUNEVERSION -q $DUNEQUALIFIER
```

The *source* line sets up versions of the software products and the environment needed to run the DUNE-specific code using LArSoft

The *setup* line says to use version 09_72_02d00 of the dunesw software

→ This done, it is possible to run larsoft: ***lar -c myfcl.fcl artrootfile.root***

lar line runs the art framework using a DUNE 'fcl' file as input. This file defines what the software is supposed to do

2) Intermediate: modify the configuration files, by changing one or more parameters

- Copy the fcl file in your working dir, edit, and modify it. A basic knowledge of fcl structure is needed

→ Modifying a fcl file allows to better understand the processing chain

3) Developer : in cases where configuration changes will not be sufficient you will need to modify, build, then run code

LArSoft: Demo and hands-on

tutorial

Presentation of the art framework and LArSoft, input files and most useful commands. Demo of the analysis workflow. Hands-on to learn how to modify an existing analysis module.

- Create a new working area from a fresh login + DUNE set-up
- Set up local products and development environment
- Check out the repository to be modified
- Make changes to the code and build it
- Run the code you just built

Whatever your level/time availabilities are :

- **DUNE TUTORIAL** it contains ALL needed information !

- A very complete presentation by Tom Junk (FIFE Summer school, June 2021):
[*Introduction to art, with LArSoft examples*](#)
(and links inside...)



Introduction to *art*, with LArSoft examples

Thomas R. Junk
FIFE Summer School
June 17, 2021

Useful tips

How to find fcl files?

By setting up DUNE environment, the variable `$FHICL_FILE_PATH` is defined.

This variable points to all directories where fcl files can be found. To look for a particular fcl file, check Tom presentation, page 46 :

A very short perl script for searching search paths

```
#!/usr/bin/perl
# usage: searchls.pl <environment_variable_name> <filename>

$envlist = $ENV{$ARGV[0]};
$name = $ARGV[1];
@plist = split(/:/,$envlist);
foreach $ptest (@plist) {
    $ftname = $ptest."/".$name;
    print `ls $ftname 2>/dev/null`;
}
}
```



```
-bash-4.2$ search.pl FHICL_FILE_PATH standard_reco_dunevd10kt.fcl
./standard_reco_dunevd10kt.fcl
/cvmfs/dune.opensciencegrid.org/products/dune/dunetpc/v09_37_00/job/standard_reco_dunevd10kt.fcl
./standard_reco_dunevd10kt.fcl
```

And also (page 47)



Looking for a fcl file that defines or mentions something

art utilities are great if your fcl file already includes the right thing. But what if you need to find something you haven't yet included?

```
#!/usr/bin/perl

# look through *.fcl in FHICL_FILE_PATH for a fcl file that contains a string
# usage: fcllookup.pl <string>

$envlist = $ENV{"FHICL_FILE_PATH"};
$symname = $ARGV[0];
@plist = split(/:/,$envlist);
foreach $ptest (@plist) {
    @files=<$ptest/*.fcl>;
    foreach $file (@files) {
        $output = `fgrep $symname $file`;
        if ($output ne "") {
            print $file,"\n";
        }
    }
}
}
```

How to inspect a fcl file

- *fhicl-dump* crp3cb_data_oct2022_reco.fcl
- *fhicl-expand* crp3cb_data_oct2022_reco.fcl

Processes fcl files using FHICL_FILE_PATH to look up #included files.

How to inspect an artroot file

- *config_dumper* -P 1727_62_d_cb_reco_67267222_0_2023-04-11T095153Z.root
dumps configuration information stored in an artroot file, output is in the form of a FCL file
- *count_events* 1727_62_d_cb_reco_67267222_0_2023-04-11T095153Z.root
- *product_sizes_dumper* -f 0 1727_62_d_cb_reco_67267222_0_2023-04-11T095153Z.root
inspects an artroot file and lists the data products inside, along with their sizes in the file
- *sam_metadata_dumper* 1727_62_d_cb_reco_67267222_0_2023-04-11T095153Z.root
metadata allow to understand data samples

Data Products (1)

LArSoft Data Products

- See Tingjun Yang's talk at the January 2018 ProtoDUNE analysis workshop

<https://indico.fnal.gov/event/19133/contributions/50492/attachments/31462/38611/dataproducts.pdf>

- Raw Digits
- recob::Wire
- recob::Hit
- recob::Cluster
- recob::Track
- recob::Vertex
- recob::SpacePoint
- recob::Shower
- recob::PFParticle
- simb::MCParticle
- simb::MCTruth
- anab::Calorimetry
- sim::SimChannel (legacy)

Data Products (2)

LArSoft Data Products

A very good introduction to data products such as raw digits, calibrated waveforms, hits and tracks, that are created and used by LArSoft modules and usable by analyzers was given by Tingjun Yang at the [2019 ProtoDUNE analysis workshop](#) (larsoft-data-products).

There are a number of data product dumper fcl files. A non-exhaustive list of useful examples is given below:

Code

```
dump_mctruth.fcl
dump_mcparticles.fcl
dump_simenergydeposits.fcl
dump_simchannels.fcl
dump_simphotons.fcl
dump_rawdigits.fcl
dump_wires.fcl
dump_hits.fcl
dump_clusters.fcl
dump_tracks.fcl
dump_pfparticles.fcl
eventdump.fcl
dump_lartpcdetector_channelmap.fcl
dump_lartpcdetector_geometry.fcl
```

Data product dumper fcl files

Some of these may require some configuration of input module labels so they can find the data products of interest

! Key Points

- Art provides the tools physicists in a large collaboration need in order to contribute software to a large, shared effort without getting in each others' way.
- Art helps us keep track of our data and job configuration, reducing the chances of producing mystery data that no one knows where it came from.
- LArSoft is a set of simulation and reconstruction tools shared among the liquid-argon TPC collaborations.

Some good practices (see also [Heidi Schellmann talk](#) January DUNE CM)

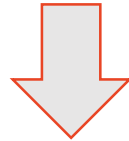
Basic principles:

Science results have to be reproducible

1 → A tagged set of algorithms: code management, releases, configuration files, code visible by the collaboration

2 → known provenance of data: tracked processed chain, cataloged data

→ safe delivery of data



Intermediate or developer levels: once a fcl file has been modified, or a some new code has been prepared, ad hoc tests are needed !!!

- Always test your code on a “small sample” (for cold box some few raw data input files) before moving larger processing
- What is a “waste of CPU time (and/or resources)”?

Running ~1000 jobs without carefully testing the code before, repeat this operation several times and **DO NOT SHARE output files** with colleagues.

- **If you plan/need to run simulations or reconstruct a large sample of data:**

1. Discuss within the relevant physics group
2. Contact the production group, so that very large data sample (MC or raw data) are properly processed, described, stored. These 2 last steps are sometimes hidden to “standard users” but are essential to insure data access to everyone in a reproducible way (*hidden: metadata generation and handling, where to store*)

The point is not to prevent people from working, but to define some basic rules to work together in a collaborative way

Data availability

Coldbox

The updated summary of coldbox reconstructed datasets is available here:

https://wiki.dunescience.org/wiki/ProtoDUNE_Vertical_Drift_operation_status

ProtoDUNE Vertical Drift operation status

Getting started [\[edit\]](#)

- CERN elog [\[1\]](#)
- Slack channel : # protodune-vd-sim-reco
- Weekly analysis meeting (before September 10th, 2022) [\[2\]](#)
- Weekly analysis meeting [\[3\]](#)

raw data	reconstructed data
vd-coldbox-top_cycle1_cosmics_raw	vd_coldbox_top_Nov2021_reco_v094900d00
vd-coldbox-top_cycle2_cosmics_raw	vd_coldbox_top_Dec2021_reco_v094900d00
vd-coldbox-top_crp1b_cosmics_raw	vd_coldbox_top_crp1b_reco_v094900d00

March 17th 2023 configuration of coherent noise removal fixed, 1727 has be reprocessed

- run 1727
- dunesw: v09_69_01d00
- fcls: crp3cb_data_oct2022_reco.fcl, standard_anatree_crpcb_data.fcl
- dataset: vd-coldbox-top_runset_1727_reco_v096901d00 (~30K events)
- ntuple file: 1727_reco_v096901d00_anatree_v096901d00.root

April 12th, 2023

misconfiguration (Wirecell) in the sampling rate fixed, correct value of Efield implemented (https://dunescience.slack.com/files/U02KGE7HLHF/F051J5Q0W66/230403-dqdx.pdf?origin_team=T03RN7KU3&origin_channel=C043QB33NV8)

- run 1727
- dunesw: v09_72_00d00
- fcls: crp3cb_data_oct2022_reco.fcl,
- dataset: vd-coldbox-top_runset_1727_reco_v097200d00 (~30K events)

For each data set, a description of the workflow and a reference to the relevant presentations at the analysis meeting is provided

FD2 Monte Carlo production

https://wiki.dunescience.org/wiki/Vertical_Drift_FD_simulation_and_reconstruction

Vertical Drift FD simulation and reconstruction

Contents [hide]

- 1 Geometry
- 2 Workflow
- 3 Productions
 - 3.1 FHC
 - 3.2 RHC
- 4 July 2022
 - 4.1 FHC (30 deg)
 - 4.2 RHC (30 deg)
 - 4.3 FHC (48 deg)
 - 4.4 RHC (48 deg)

dunesw v09_37_00, 48 deg January 2022 ~6M events

dunesw v09_53_00d02, August 2022 ~24M events

Workflow [edit]

At time of writing, the GENIE (with LBNF beam flux) and MARLEY generators are currently 'plumbed' into the workflow. The particle tracking simulation solely uses the refactored larg4. The detector response simulation and signal processing are provided by wirecell. Pandora provides the pattern recognition and 3D reconstruction. The CVN (WIP) provides neutrino flavour tagging at the event-level.

The relevant fcls to run the workflow are:

1. **generation:** [prodgenie_nu_dunevd10kt_1x8x6_3view.fcl](#) (other generator fcls can be found [here](#))
2. **Particle tracking simulation:** [standard_g4_dunevd10kt_1x8x6_3view.fcl](#)
3. **Detector response simulation:** [standard_detsim_dunevd10kt_1x8x6_3view.fcl](#)
4. **Reconstruction:** [standard_reco_dunevd10kt_1x8x6_3view.fcl](#)
5. **Analysis tree:** [standard_anatree_dunevd10kt_1x8x6_3view.fcl](#)

Production 1 (48 deg)

<https://indico.fnal.gov/event/52011/contributions/228798/attachments/149639/192706/2021-11-22%20wire-cell-sim-sigproc.pdf>

https://indico.fnal.gov/event/52011/contributions/228800/attachments/149638/192726/VDReconstruction_221121_DBrailsford.pdf

https://indico.fnal.gov/event/53402/contributions/235587/attachments/152510/197523/VDProductionSamDefs_280222_DBrailsford.pdf

<https://indico.fnal.gov/event/53402/contributions/235500/attachments/152515/197529/2022-02-28-2d-sigproc.pdf>

Production 2 (30+48 deg)

<https://indico.fnal.gov/event/54041/contributions/238815/attachments/153961/199916/FDSimRecoMeet11Apr22.pdf>

<https://indico.fnal.gov/event/54239/contributions/239838/attachments/154356/200530/2022-04-25%20parital%208x14%2C%20combined%20workflow.pdf>

https://indico.fnal.gov/event/54239/contributions/239837/attachments/154355/200529/VD30DegProduction_250422_DBrailsford.pdf

<https://indico.fnal.gov/event/54472/contributions/240959/attachments/154877/201563/2022-05-09%20DNN-ROI.pdf>

All talks here: [FD simulation and reconstruction meeting \(27 juin 2022\) · INDICO-FNAL \(Indico\)](#)

https://indico.fnal.gov/event/55417/contributions/246443/attachments/157279/205666/VD30Deg_180722_DBrailsford.pdf

https://indico.fnal.gov/event/55865/contributions/248548/attachments/158867/208661/VD30DegAndTDR_150822_DBrailsford.pdf

https://indico.fnal.gov/event/55865/contributions/248547/attachments/158868/208662/Vdtests_simReco_15August.pdf

Complete overview of geometry and software [here](#)

L. Paulucci

The '1x8x6' workspace geometry



What's next?

Coldbox

- Setting up and testing of reconstruction of CRP2-CRP3 raw data : run 1727 is the reference run: it is processed every time the reconstruction software is updated

Goal: converge on a stable reconstruction code, to be used to process more CRP2-CRP3 runs.
Important also in view of Module-0

- Preparation of CRP4 and CRP5 reconstruction code is ongoing

Module-0

- Simulation campaign to be defined (Friday meeting)

Future FD MC productions (1)

New simulation campaign for both FD1-HD and FD2-VD

- Simulation differences make it difficult to quantitatively compare detector performances
- The group's primary focus is now updating the FD1-HD samples to use an identical simulation to FD2-VD
 - Same neutrino interaction model
 - More realistic detector response modelling using Wirecell

- A new LBL production is in preparation, simulating ~20 million neutrinos split between both detectors

Timeline

- Software preparation: **1 month**
- 1st phase of production campaign (simulation and hit reconstruction): **1 month**
- CVN, Pandora, energy reconstruction tuning: **2 months**
- 2nd phase of production campaign (high level reconstruction): **1 month**

→ *It will not start before May CM*

Biweekly meeting

<https://indico.fnal.gov/category/497/>

Slack channel #fd-sim-reco

Future FD MC productions (2)

New simulation campaign for both FD1-HD and FD2-VD

Low Energy Physics

Generate a new fast light simulation with updated geometry for the next LE events MC production

Analysis to be updated for supernova and solar neutrino studies with the VD and HD geometries with upgraded simulation framework

Working on:

- More complete background model with radioactivity from cavern and cryostat
- Low energy physics reconstruction with Pandora
- VD light simulation with Xe light

Main sensitivity studies to be updated with these new samples:

- **Solar neutrinos:** performed on HD geometry, on-going for VD
- **Low energy resolution and performance** update HD and VD TDRs
- **Triggering efficiency with TPC and PDS** (as a function of the PE threshold)



Working at CCIN2P3

- Local work at CCIN2P3 is meant for **analysis or software development**, **NOT FOR MASSIVE DATA PROCESSING**
- **Software development:**
it is possible to run, to check out, modify, and build DUNE code
(see example log here */sps/lbno/workshop/log.out*)
- **Data access:** a DUNE VO certificate is needed. Transition to token authentication is ongoing

Advantages of working locally:

- 1) editing, root sessions, graphical windows → easier and faster
- 2) Batch system (SLURM) user-friendly, straightforward access to output logs
- 3) Temporary areas where to store files available
- 4) All data are accessible

Storage areas

1) Personal storage area \$HOME

Daily backup

```
% echo $HOME
```

```
/pbs/home/<u>/<user>
```

2) \$THRONG directory: common area
where users may share with each
other important data/code

Daily backup

```
% echo $THRONG_DIR
```

```
/pbs/throng/<group>
```

3) Group directory

NO BACKUP

```
/sps/lbno
```

Setting up a session

```
>more setup.sh
```

```
source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh
```

```
export DUNEVERSION=v09_72_00d00
```

```
export DUNEQUALIFIER="e20:prof"
```

```
setup dunesw $DUNEVERSION -q $DUNEQUALIFIER
```

```
setup sam_web_client
```

```
> source setup.sh
```

```
Setting up larsoft UPS area... /cvmfs/larsoft.opensciencegrid.org
```

```
Setting up DUNE UPS area... /cvmfs/dune.opensciencegrid.org/products/dune/
```

```
> samweb get-file-access-url 1727_19_b_cb_reco_57388421_0_2023-02-25T203614Z.root --schema=root
```

```
root://fndca1.fnal.gov:1094/pnfs/fnal.gov/usr/dune/tape_backed/dunepro/vd-coldbox-top/full-  
reconstructed/2023/detector/test/VD_coldbox_CRP2_CRP3_2022/00/00/17/27/1727_19_b_cb_reco_57388421_0_2023-02-  
25T203614Z.root
```

```
> lar -c standard_anatree_crpcb_data.fcl -n 1 root://fndca1.fnal.gov:1094/pnfs/fnal.gov/usr/dune/tape_backed/dunepro/vd-coldbox-  
top/full reconstructed/2023/detector/test/VD_coldbox_CRP2_CRP3_2022/00/00/17/27/1727_19_b_cb_reco_57388421_0_2023-02-  
25T203614Z.root
```

To stream raw data files some preliminary work is needed (LD_PRELOAD env variable has to be defined)

if the same files has to be read several times, it is more efficient to have a local copy:

```
-bash-4.2$ ls -rtl
total 0
-bash-4.2$
1. locate the file
-bash-4.2$ samweb locate-file 1727_19_b_cb_reco_57388421_0_2023-02-25T203614Z.root
enstore:/pnfs/dune/tape_backed/dunepro/vd-coldbox-top/full-reconstructed/2023/detector/test/VD_coldbox_CRP2_CRP3_2022/00/00/17/27(11219@fb063619)
-bash-4.2$
-bash-4.2$
-bash-4.2$
-bash-4.2$
2. Copy the file
-bash-4.2$ ifdh cp /pnfs/dune/tape_backed/dunepro/vd-coldbox-top/full-reconstructed/2023/detector/test/VD_coldbox_CRP2_CRP3_2022/00/00/17/27/1727_19_b_cb_reco_57388421_0_2023-02-25T203614Z.root .
Copying 1324057152 bytes https://fndcador.fnal.gov:2880/dune/tape_backed/dunepro/vd-coldbox-top/full-reconstructed/2023/detector/test/VD_coldbox_CRP2_CRP3_2022/00/00/17/27/1727_19_b_cb_reco_57388421_0_2023-02-25T203614Z.root => file:///sps/lbno/elisabetta/CCIN2P3/2023/workshop/testcp/1727_19_b_cb_reco_57388421_0_2023-02-25T203614Z.root
-bash-4.2$ ls -rtl
total 1556138
-rwxr-xr-x 1 pennacc lbno 1324057152 Apr 13 12:07 1727_19_b_cb_reco_57388421_0_2023-02-25T203614Z.root
-bash-4.2$
```


CCIN2P3 disposes also of dCache storage space (disk and tape)

Data available on dCache at CCIN2P3

❖ Reconstructed data:

- **np02-DP** np02_hitrecon_2019
protodune-dp_runset_1415_cosmics_reco_NP02_keepup_March2021_v0
protodune-dp_runset_1407_cosmics_reco_NP02_keepup_March2021_v0
- **np02 6m tracks** np02_6m_2022_fullreco_v096301d0
- **COLDBOX** vd_coldbox_top_Nov2021_reco_v094900d00
vd_coldbox_top_Dec2021_reco_v094900d00
vd_coldbox_top_crp1b_reco_v094900d00
vd_coldbox_top_ntuple_2021_v095400d00
vd-coldbox-top_runset_1727_reco_v096700d00
vd-coldbox-top_runset_1727_reco_v096300d00
vd-coldbox-top_runset_1727_reco_v096901d00
vd-coldbox-top_runset_1727_reco_v097200d00

❖ All raw data from prototyping activities at EHN1

- More data to come in the very next future: CRP reconstruction results, ProtoDUNE simulations....
- The implementation of token authentication is finished, but some more tests are needed so for the moment the access to this storage is not yet open to users for analysis
- Once the tests are finished, during the DUNE-IN2P3 Wednesday meeting I'll circulate instructions and example on how to access these data
- **Analysis root files are available** on /sps/lbno/coldbox : **accessible in read mode to all users in lbno group** (no need of certificate or token)

- **Some questions:**

- Are you using these ntuple files?

- Very probably the answer is no → why?

- 1) I was not aware files are available, 2) the variables in the files are not useful, 3) important variables are missing.....*

- in case 2 and 3 are the right answers, we should identify which variables are needed

Batch system: Slurm

<https://doc.cc.in2p3.fr/fr/Computing/computing-introduction.html#batch>

Actuellement le logiciel d'ordonnancement de la ferme locale du CC-IN2P3 est **Slurm**.

Pour aller plus loin, vous pouvez aussi vous référer aux documents utilisés lors des **formations dispensées** par le CC-IN2P3.

https://gitlab.in2p3.fr/ccin2p3-support/formations/batch/2022.02/tuto_batch/-/blob/master/tutorial02.md
https://gitlab.in2p3.fr/ccin2p3-support/formations/batch/2022.02/tuto_batch/-/blob/master/tutorial01.md

- **Soumettre un job**
 - Exemples de soumission
 - Principales options de sbatch
 - Environnement et limites
 - Ressources de stockage et logicielles
- **Suivi des jobs**
 - Etat du service de la ferme
 - Statut de soumission d'un job
 - Efficacité d'un job
 - Suspension et altération d'un job
 - Annulation d'un job
 - Statut de fin de job
 - Profilage des jobs
- **Configuration**
 - Partitions
 - Qualité de service
 - Nœuds

Example scripts in / sps/lbno/workshop/slurm

workshop_lartest.sh

```
#!/bin/sh

#SBATCH --job-name=my_job_test      # Job name
#SBATCH --output=my_job_test_%j.log  # Standard output and error log

#SBATCH --partition=htc             # Partition choice (htc by default)

#SBATCH --ntasks=1                 # Run a single task
#SBATCH --mem=2000                  # Memory in MB per default
#SBATCH --time=0-02:00:00           # Max time limit = 7 days, in this example 2 hours

#SBATCH --mail-user=myname@mail     # Where to send mail
#SBATCH --mail-type=BEGIN,END,FAIL   # Mail events (NONE, BEGIN, END, FAIL, ALL)

source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh

export DUNEVERSION=v09_72_00d00
export DUNEQUALIFIER="e20:prof"
setup dunesw $DUNEVERSION -q $DUNEQUALIFIER

^lar -c crp3cb_data_oct2022_reco.fcl 1727_1_a_cb.test
```

values to be optimized

- It is possible to receive a mail during job execution
- Can be useful, but the mailbox can be fully saturated!

Input file and output files in the launch directory

workshop_lartest_copy.sh

```
#!/bin/sh

#SBATCH --job-name=my_job_test      # Job name
#SBATCH --output=my_job_test_%j.log # Standard output and error log

#SBATCH --partition=htc            # Partition choice (htc by default)

#SBATCH --ntasks=1                 # Run a single task
#SBATCH --mem=2000                  # Memory in MB per default
#SBATCH --time=0-02:00:00          # Max time limit = 7 days, in this example 2 hours

source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh

export DUNEVERSION=v09_72_00d00
export DUNEQUALIFIER="e20:prof"
setup dunesw $DUNEVERSION -q $DUNEQUALIFIER

cp input_1727.test $TMPDIR/input_1727.test

# how to write outfile on $TMPDIR
lar -c crp3cb_data_oct2022_reco.fcl $TMPDIR/input_1727.test -o $TMPDIR/output.root

echo "copy back output file"

cp $TMPDIR/output.root myoutput.root
```

input and output files
in \$TMPDIR

better for large files

\$TMPDIR

disk space temporarily available on the execution node during the execution of the batch job.

Job Monitoring

To launch a job

```
-bash-4.2$ sbatch workshop_lartest.sh
sbatch: INFO: Account: lbno
sbatch: INFO: Submission node: cca003
sbatch: INFO: Partition set to: htc
sbatch: INFO: Partition limited to one node per job.
sbatch: INFO: Time limit set to: 0-02:00 (120 minutes)
Submitted batch job 26515228
```

To follow its execution

```
-bash-4.2$ squeue
          JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
          26515228      htc myjob te pennacc  R        12:23         1 ccwslurm0040
```

Once finished

```
-bash-4.2$ seff 26515228
Job ID: 26515228
Cluster: ccslurmlocal
User/Group: pennacc/lbno
State: COMPLETED (exit code 0)
Cores: 1
CPU Utilized: 00:20:04
CPU Efficiency: 94.14% of 00:21:19 core-walltime
Job Wall-clock time: 00:21:19
Memory Utilized: 1.47 GB
Memory Efficiency: 75.50% of 1.95 GB
-bash-4.2$
```

Check memory usage: do not ask for more memory of what is needed...

Testing the code/the workflow

When you login at CCIN2P3, you connect to an interactive server, which is foreseen for code development and very short tests.

Longer tests, such as run *lar* on several events in a raw data file, can be run an [“interactive job”](#)

1. Open a session with slurm option
(example in */sps/lbno/workshop/inter.sh*)

2. Resources are allocated

3. Run interactively the job

```
-bash-4.2$ srun -t 0-08:00 -n 8 --mem 2G --pty bash -i
srun: INFO: Account: lbno
srun: INFO: Submission node: cca002
srun: INFO: Partition set to: htc_interactive
srun: INFO: Partition limited to one node per job.
srun: INFO: #####
srun: INFO: #   You are about to be logged on an interactive worker   #
srun: INFO: #           type 'exit' or ctrl-d to end this session       #
srun: INFO: #####
srun: INFO: Time limit set to: 0-08:00 (480 minutes)
srun: job 26547139 queued and waiting for resources
srun: job 26547139 has been allocated resources
bash-4.2$ source setup.sh
Setting up larsoft UPS area... /cvmfs/larsoft.opensciencegrid.org
Setting up DUNE UPS area... /cvmfs/dune.opensciencegrid.org/products/dune/
bash-4.2$ lar -c crp3cb_data_oct2022_reco.fcl -n 1 input_1727.test
DuneDeconvolutionService::ctor: DuneDeconvolutionService:
DuneDeconvolutionService::ctor:                               LogLevel: 1
DuneRoiBuildingService::ctor: DuneRoiBuildingService:
DuneRoiBuildingService::ctor:                               LogLevel: 1
DuneRoiBuildingService::ctor:   NSigmaStart: 3
DuneRoiBuildingService::ctor:   NSigmaEnd: 1
DuneRoiBuildingService::ctor:   PadLow: 50
DuneRoiBuildingService::ctor:   PadHigh: 50
StandardAdcWireBuildingService::ctor: StandardAdcWireBuildingService:
```

Conclusions

Follow

- the computing tutorial
- DUNE analysis meetings
 1. **coldbox**: Friday afternoon meeting, subscribe to the e-group (<https://e-groups.cern.ch/>): *cenf-vd-integration-analysis*
 2. **Far detector sim/reco** : *Biweekly meetings, Monday afternoon*
subscribe dune-reco mailing list
- During these meetings **LArSoft configuration aspects** are discussed and advertised (fcl files, code development, algorithms status,...).
- Analysis results are discussed as well.

It is necessary that the IN2P3 collaborators integrate in this global context, provide feedback and contributions, and also profit of help which **may come from LarSoft, Pandora, wirecell, ... experts**

To be discussed:

1

■ Some questions:

→ Are you using these ntuple files?

→ Very probably the answer is no → why?

1) I was not aware files are available, 2) the variables in the files are not useful, 3) important variables are missing.....

→ in case 2 and 3 are the right answers, we should identify which variables are needed

2

CCINP23: are there particular needs? Is there something you would need that is missing?

3

If needed, a dCache *scratch area* can be setup (similar to Fermilab)

Answers may come during the workshop....