Communications from Physics Coordinator

Igor Denisenko
iden@jinr.ru

SPD Physics & MC Meeting
18.12.2024

The second test production

- Thanks to Artem, the the send test production using **SpdRoot 4.1.6.1** has been successfully finished in JINR and PNPI.
- In total there are **46112 reco files**, **4K events per file**, **~184.5 million events**.
- The first available data (**~40 million events**) could be quickly processed on lxui (see next slides), once we go to **1 billion** the **centralized data processing** may be needed.
- The simu-files will are being removed, the **reco-files will be kept** until the data production with the new SpdRoot version will run out of space (**persist at least until the end of January)**.
- The new production based on the SpdRoot development branch will start before the New Year for three scenarios (1-st stage physics, 2-nd stage DSSD and MAPS).
- Moving to centralized MC-production is an important **milestone** for us!

Some statisitcs

Some stats (4000 events per job)

Task ID	Task type	Jobs processed	Jobs successful	Start	Finish	Output dataset name	Output dataset size
265	reco	68914	31363	12:33 December 6, 2024	9:46 December 13, 2024	test.2024.MC.27GeV.test-minbias.00006.RECO.1.R	66.05 TB
262	simu	108447	37677	13:01 November 28, 2024	17:26 December 4, 2024	test.2024.MC.27GeV.test-minbias.00006.SIMUL.1.S	83.29 TB
266	reco	7277	4838	12:27 December 13, 2024	8:32 December 17, 2024	test.2024.MC.27GeV.test-minbias.00003.RECO.1.R	10.19 TB
259	simu	6271	4940	19:37 November 27, 2024	13:42 December 3, 2024	test.2024.MC.27GeV.test-minbias.00003.SIMUL.1.S	10.92 TB

Some statistics

Some stats (4000 events per job)

Task ID	Task type	Jobs processed	Jobs successful	Start	Finish	Output dataset name	Output dataset size
265	reco	68914	31363	12:33 December 6, 2024	9:46 December 13, 2024	test.2024.MC.27GeV.test-minbias.00006.RECO.1.R	66.05 TB
262	simu						83.29 TB
266	reco						10.19 TB
259	simu						10.92 TB

- Our simulation and reconstruction processes are **expensive for both CPUs and storage.**
- We can generate more than **100 million events / week**.
- Disk quota limits the amount of the stored data. Revision of the stored information for simulation and reconstruction is **strongly needed** (could somebody take this responsibility?)
- Job success ratio depends on both the production infrastructure and SpdRoot code. On the SpdRoot side there is still a fraction of jobs entering infinite loop with Nans appearing in GenFit.

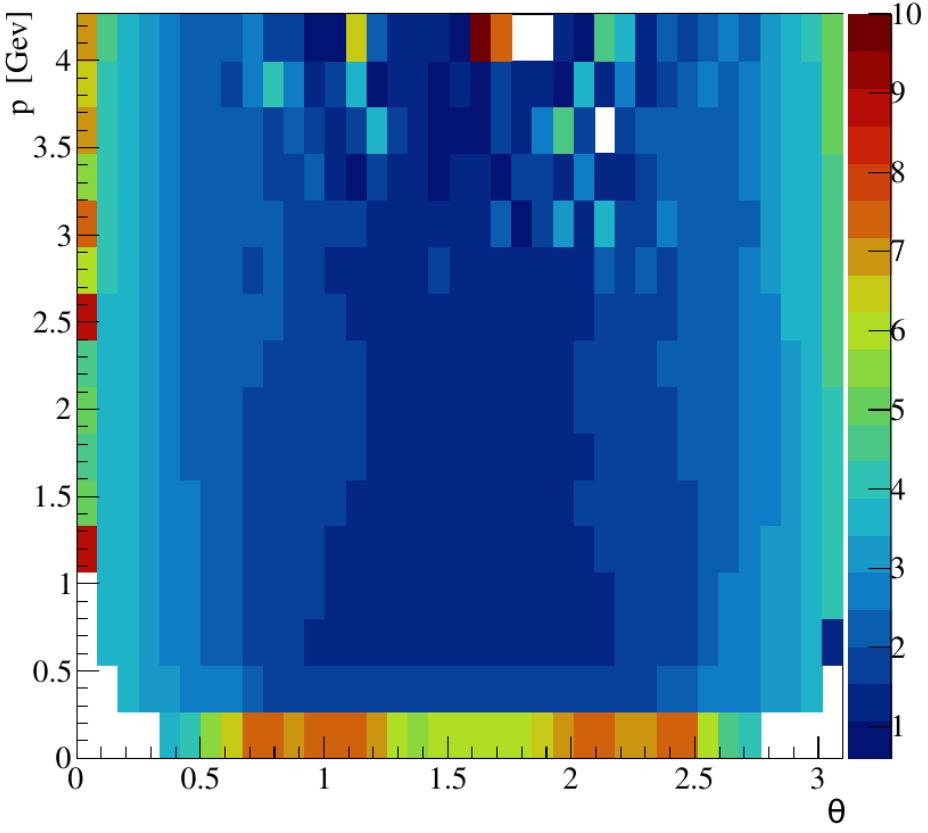
Running user analysis on the production data

Name	Last commit	Last update
REAMDE.md	README update	1 week ago
pilot.py	Initial commit	1 week ago
slurm_scr.sh	Added commnets	4 days ago

- There is a batch script **template** at gitlab (modify it for you needs!): https://git.jinr.ru/idenisenko/batch_scripts
- Analysis script is of the form *script.C(input_file.root, output_file.root)*
- Lightweight jobs run on 40 million events in 24 hours.
- The update list of the production files is at /eos/nica/spd/users/iden/production/test-minbias/reco_file_list_full.txt

```
#!/bin/sh

# ================================================================
# ==== Check number of batch jobs here!
# ==== You have to calculate number of jobs manually as
# ==== njobs = nfiles / files-in-job [+1]
# ==== Provide the correct array arg here.
# ==== By default there will be two jobs, each processes 50 files
#SBATCH --array=1-2
# ================================================================

#SBATCH --job-name=MB-tests # Job name
#SBATCH --mail-type=FAIL # Mail events (NONE, BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=iden@jinr.ru # Where to send mail
#SBATCH --no-requeue # do not auto requeue on errors
#SBATCH --ntasks=1 # Run on a single CPU
#SBATCH --cpus-per-task=1
#SBATCH --mem=2gb # Job memory request
#SBATCH --time=2-00:00:00 # Time limit days-hrs:min:sec
#SBATCH --tmp=10G # $TMPDIR space



cd ${TMPDIR}
# ============== Check absolute paths here! ===============
# ============== Check number of reco files in job (--files-in-job) ===============
python /afs/jinr.ru/user/i/iden/slurm/ana2/pilot.py \
    --container-path=/cvmfs/spd.jinr.ru/images/spdroot-4.1.6.sif \
    --ana-script=/afs/jinr.ru/user/i/iden/slurm/ana/scripts/ana2.C \
    --reco-file-list=/eos/nica/spd/users/iden/production/test-minbias/reco_file_list.txt \
    --par-file-list=/eos/nica/spd/users/iden/production/test-minbias/par_file_list.txt \
    --files-in-job=50 \
    --output-file-dir=/eos/nica/spd/users/iden/production/test-minbias/out2 \
    --job-id=${SLURM_ARRAY_TASK_ID}

```

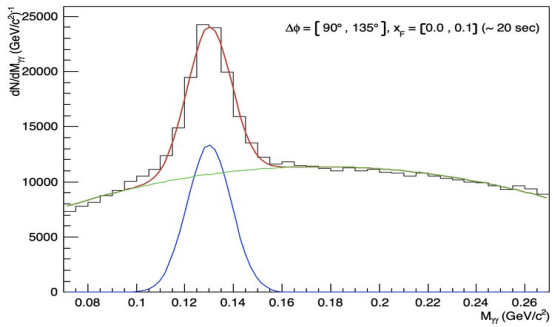
Some tests of the production (momentum resolution)



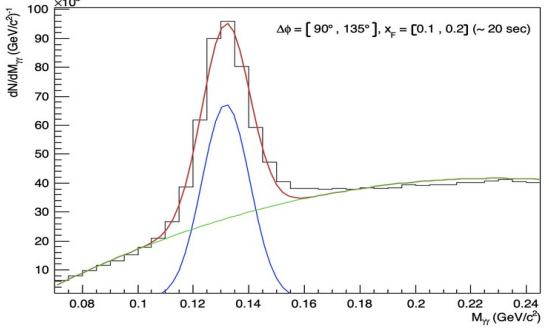
Momentum resolution as a function of azimuthal angle and momentum (40 million MB events, double Gaussian variance)

Some tests of the production (π^0 from photon clusters)

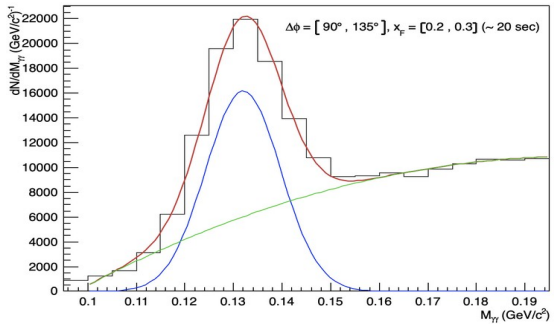
$x_F = [0.0 - 0.1]$



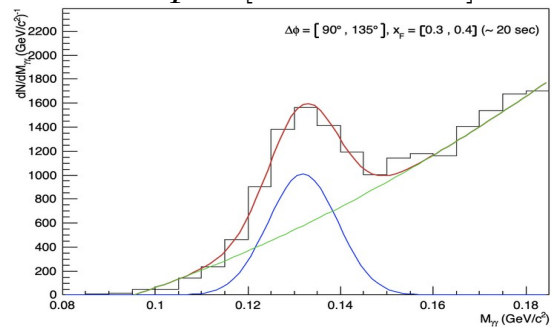
$x_F = [0.1 - 0.2]$



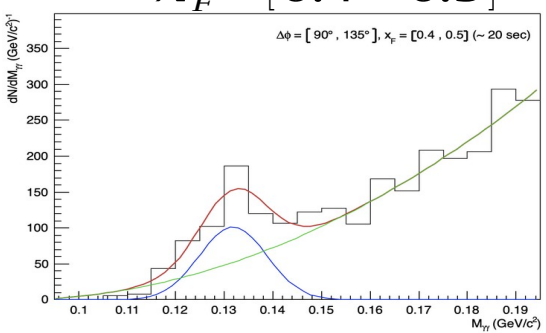
$x_F = [0.2 - 0.3]$



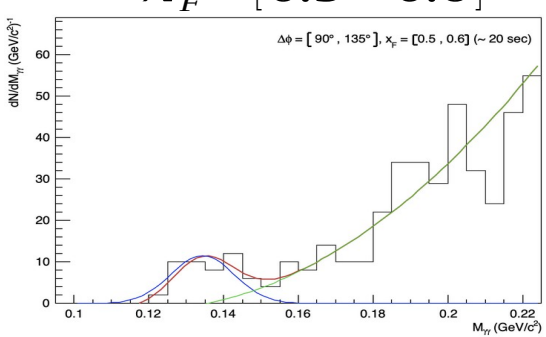
$x_F = [0.3 - 0.4]$



$x_F = [0.4 - 0.5]$



$x_F = [0.5 - 0.6]$



π^0 in online calorimetry (figures prepared by Katherin)

Thank you!