A New Pileup Mixing Framework for CMS

Junghwan Goh[1], Mike Hildreth[2] on behalf of the CMS Collaboration

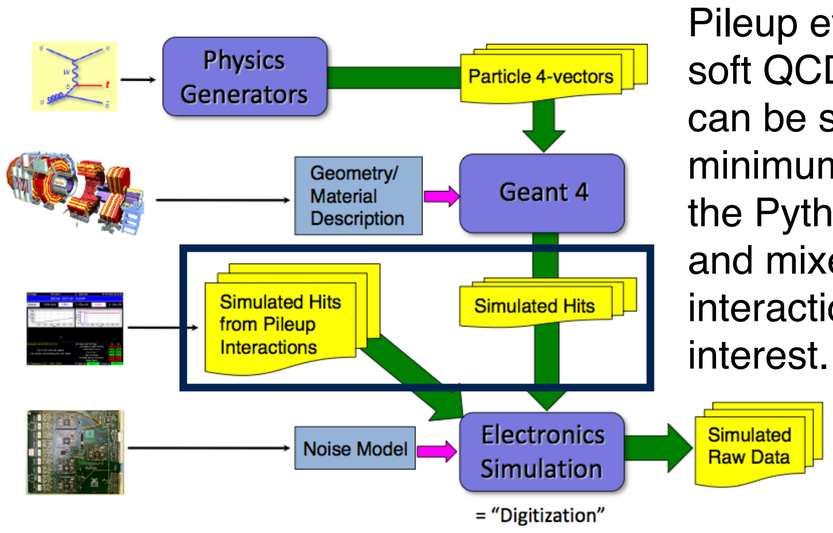
1) SungKyunKwan University, 2) Université de Notre Dame du Lac

The CMS Simulation uses minimum bias events created by a “standard” event generator (e.g., Pythia) to simulate the additional interactions due to peripheral proton-proton collisions in each bunch crossing at the LHC (also known as pileup). Due to the inherent time constants of the CMS front-end electronics, many bunch crossing before and after the central bunch crossing of interest must be included in the simulation, leading to hundreds of minimum bias events being used for each simulated hard-scatter event.

We report on the performance gains in I/O load and computational speed made possible by a new framework that allows the combination of pileup events in a “pre-mixing” step. This has been made possible by the development of software that allows single-channel information to be combined at the digitization level in each sub-detector, rather than accumulating simulated hits from GEANT. The logistics of large-scale production with pre-fabricated pileup distribution is described.

Pileup simulation in CMS

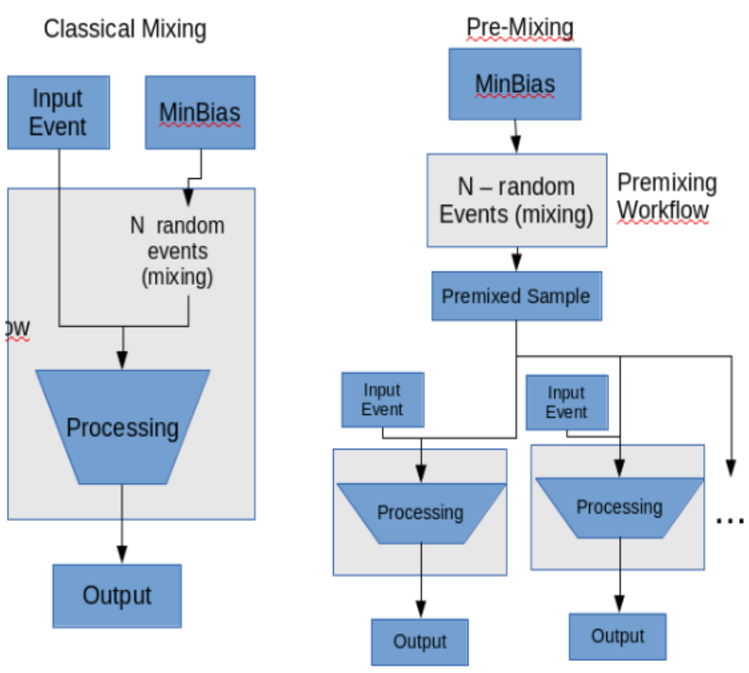
Multiple proton-proton collision events are produced in each bunch crossing at LHC, pileup events. Many bunch crossings before and after the central collision event have to be included due to the CMS front-end electronics.



Pileup events are mostly soft QCD processes. They can be simulated as minimum-bias events using the Pythia event generator and mixed with hard interaction of physics interest.

The average peak number of pileup events was ~33 during the 2012 running. It is expected to increase up to O(40) in Run II at 13TeV. At SLHC luminosity, 140 interactions per bunch crossing are expected. For the simulation, x16 bunch crossings are needed to include pileup events before and after the central collision event.

Premixing technique



In the classical pileup mixing, the Geant4 SimHits from N-random minimum bias events are mixed into the full simulation chain. Therefore this mixing process requires large computing resources such as frequent file transfers, CPU and memory consumption.

Premixing : Prepare samples only with mixing between pileup events with the desired pileup configuration.

These events are digitized and raw data is simulated in order to save disk space. Special care must be taken in order to preserve low-level signals in all subdetectors.

The premixed sample is then used by multiple workflows. A hard scattering event and premixed event are used in each workflows to do digitization step.

Computing performance gain

Computing performance was tested in production workflows in 2014. 25M ttbar events were processed in high pileup scenario (40PU, 25ns bunch spacing).

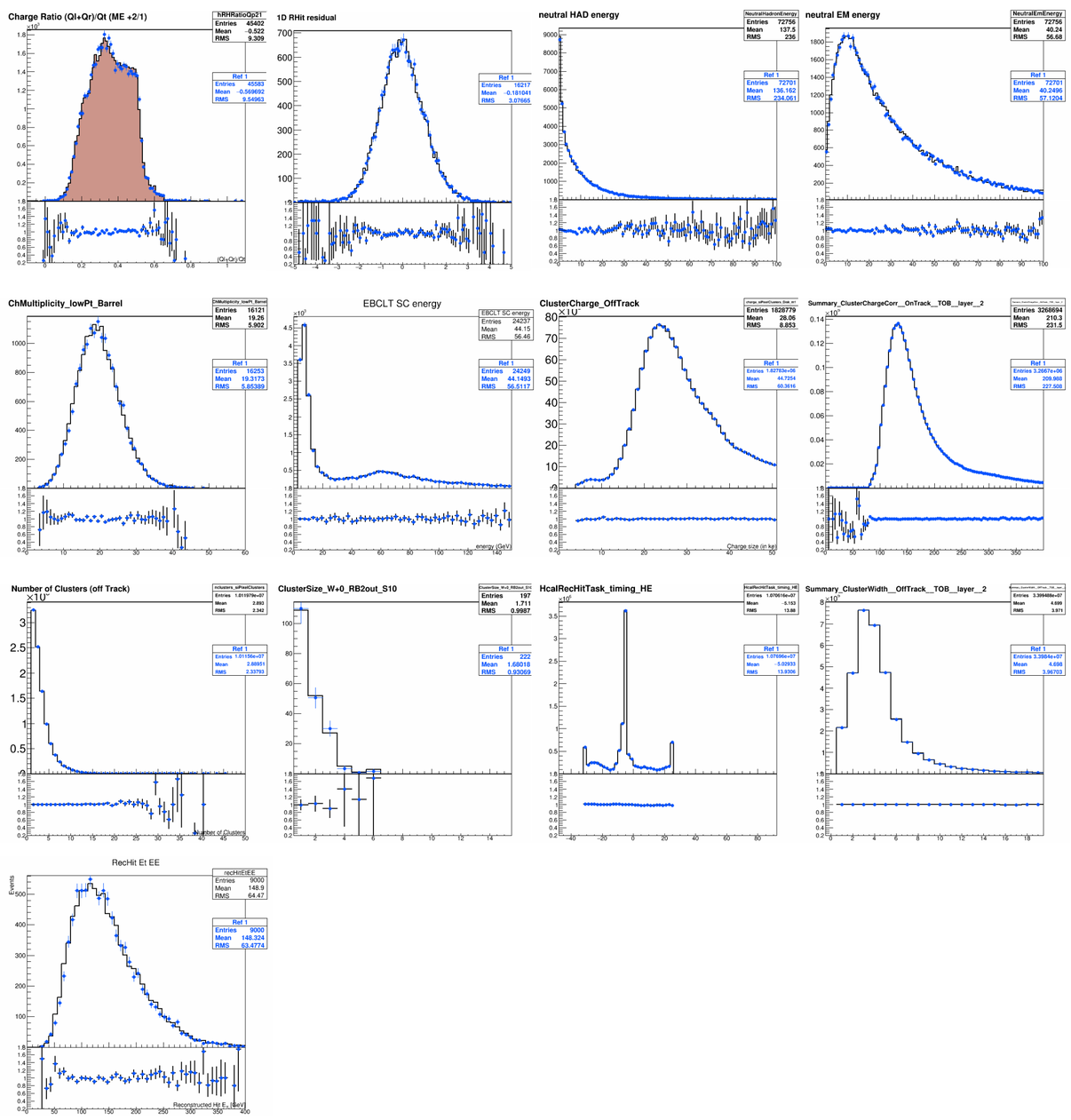
Metric	Std	Max	Min	Avg
Classic DIGI step				
AvgEventTime	8.841	84.603	28.340	49.296
ReadMBPerSecond (*)	0.510	4.768	1.595	2.833
Premixing DIGI step				
AvgEventTime	2.404	23.860	9.708	14.012
ReadMBPerSecond (*)	0.033	0.258	0.052	0.181

3 times faster,
10-15 times lower I/O rate

Factor of three-five performance gain is observed, mostly due to lower I/O requirements. This will allow production samples of any pileup distribution to be produced at Tier2 sites, and with a substantially higher rate.

Physics performance

Validation of pre-mixing is intensively performed in CMS. Detailed checks are done by every detector performance simulation and physics groups. Physics results are unchanged by premixing.



Conclusion

A new pileup mixing technique, premixing, has been developed and tested in CMS. This technique reduces CPU load and I/O rate by significant factors which allows simulation in very large pileup scenarios at Tier 2 sites.