

# Strategies for Modeling Extreme Luminosities in the CMS Simulation

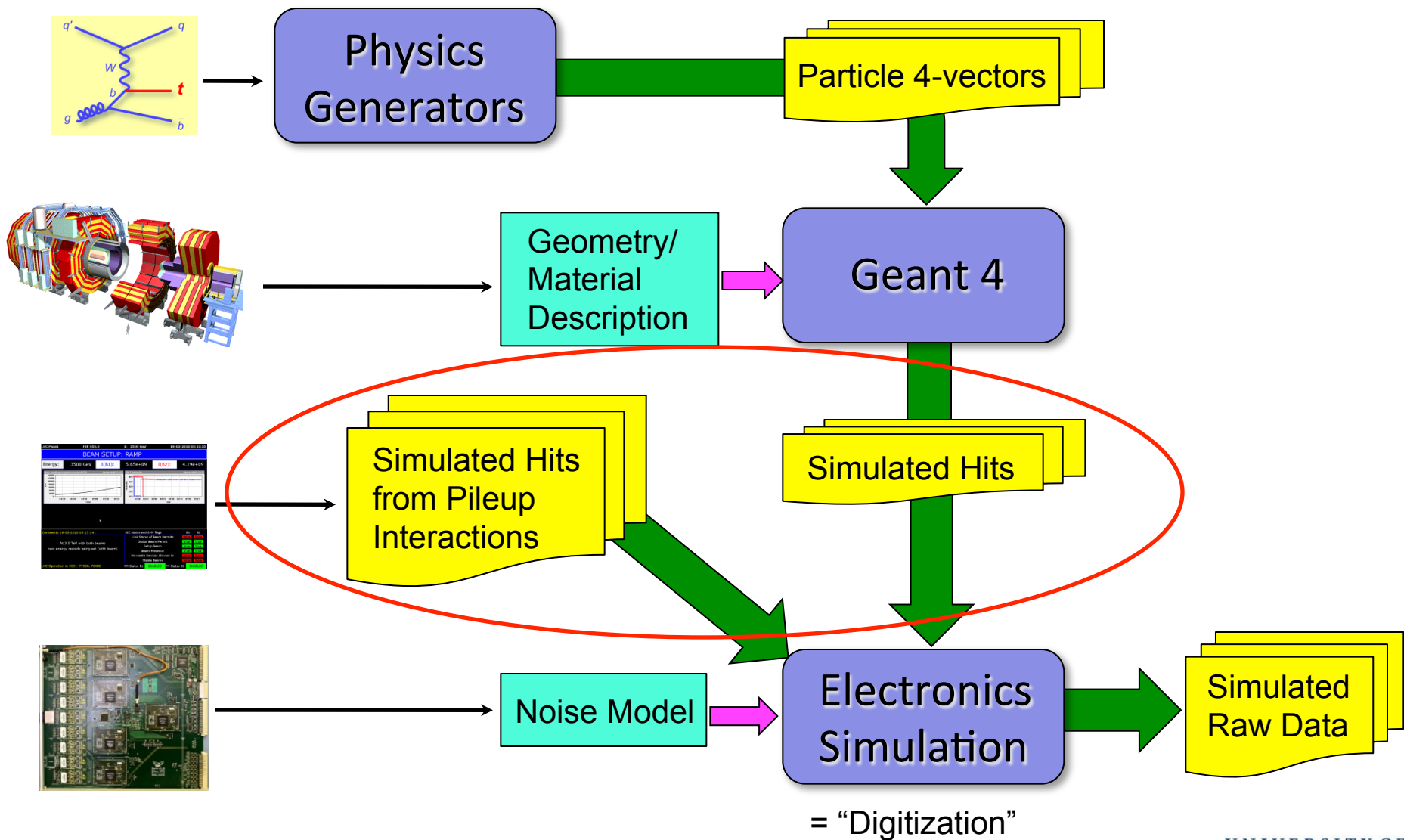
Mike Hildreth

Université de Notre Dame du Lac & Fermilab

*Representing the CMS Simulation Group*

- Overview of MinBias simulation in CMS
  - Machinery
  - cpu/memory performance
  - Out-of-time Pileup influence(s)
  
- Simulation Restructuring
  - Digitising with “accumulation”
  - Pre-Mixing
  
- Prospects

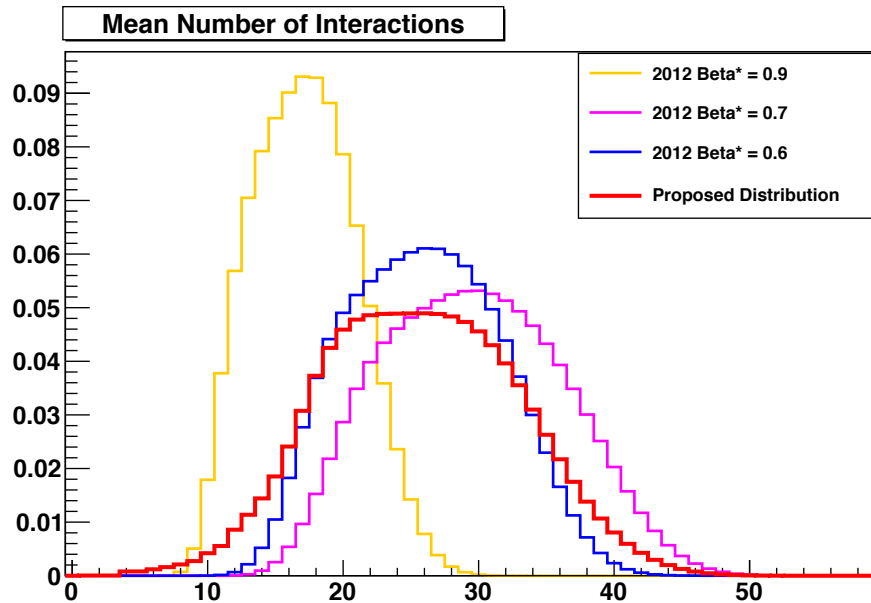
# Pileup Simulation Machinery



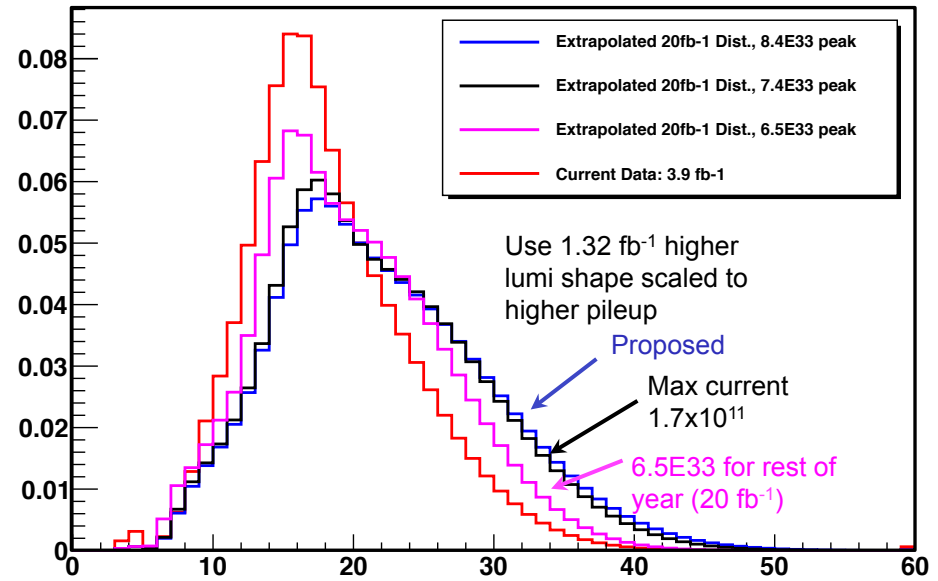
# Pileup Simulation Overview



- Pure MC inputs used to simulate pileup interactions
  - Pythia6 Tune Z2\* for Summer 12 production of minbias events
- Distribution of the number of interactions per beam crossing chosen in advance to simulate a desired luminosity profile:



2012 distribution chosen before data-taking. Extrapolations based on expected instantaneous luminosity and time evolution of accelerator performance.



2012 distribution correction mid-stream (after summer conferences) to create distribution based on observed luminosity, extrapolated performance. Note much lower mean value of pileup interaction multiplicity

# Pileup Simulation Overview (II)



- For each event, the instantaneous luminosity is chosen from the input distribution at random
  - The number of in- and out-of-time interactions to be overlaid are selected individually from a poisson distribution based on the chosen luminosity and the total inelastic cross section
    - (we have used  $\sigma_{\text{tot}} = 69.3\text{mb @ } 8 \text{ TeV}$ )
    - Pool of **pythia minbias (single interaction)** events used as input
  - Out-of-time interactions are simulated for each beam crossing that is “in scope” for a given production run
    - can do any arbitrary bunch configuration in 25ns steps
      - times of Geant SimHits are shifted to match bunch assignment
      - Digitization simulation considers hit times for pulse shapes
    - up to now, simulate  $\pm 50\text{ns}$  worth of bunch crossings
- Collection of Geant SimHits from all of the minbias events and hard-scatter “signal” event are merged, then processed by digitization/electronics simulation
  - no simulation of “double-hard-scatter” (yet)

# Simulating Extreme Luminosities

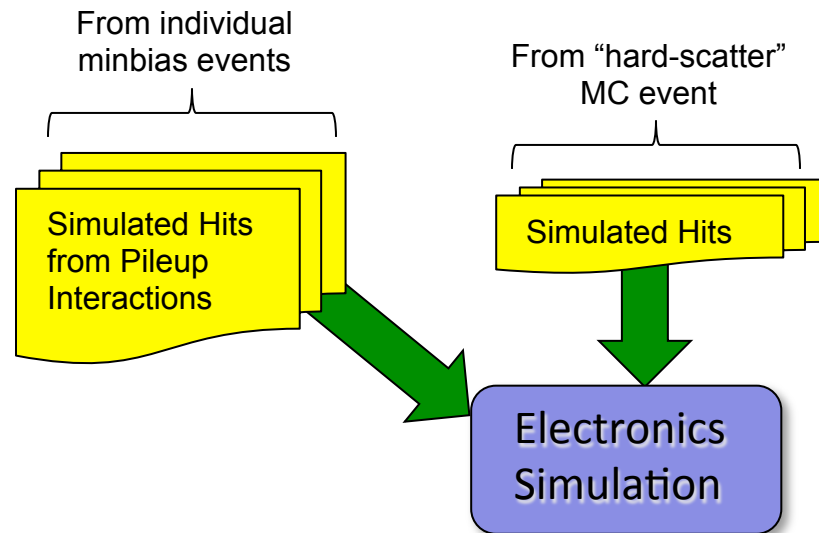


- Importance of Out-of-Time PileUp (OOTPU):
  - Because of signal shaping in readout electronics, CMS Ecal is sensitive to modeling of OOTPU up to 300ns before central BX
    - effects low end of the energy spectrum
    - important for modeling of trigger, etc.
  - Will require the simulation of **16BX** for each “hard-scatter” MC event
    - (At a bunch spacing of 25ns)
  - Studies ongoing to establish limits of sensitivity for other detectors, validate OOTPU simulation compared with Data
    - Ecal is likely to be the extreme
    - Some out-of-time hits seen in tracker simulation

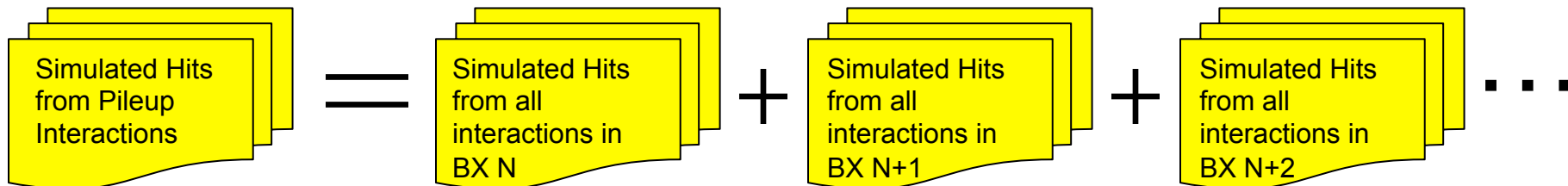
# Simulating Extreme Luminosities



- Details of Simulation Structure become important:
  - Reminder of the algorithmic structure for pileup simulation:



- In more detail:



Sum contains **all** interactions in **all** beam crossings – all in memory simultaneously!

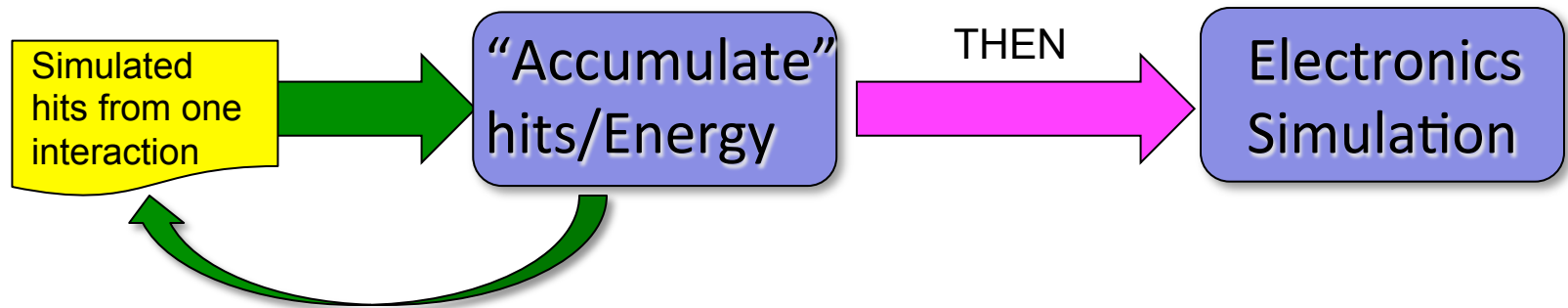
⇒ unsustainable at sLHC luminosities: ~140 interactions x 16 BXs = 2240 events in memory



# Modifications to Simulation Structure



- Re-factorization of code to process all interactions sequentially
  - Required substantial rewrite of Digitization code, re-organization of internal event processing
    - digitization is now internal to the pileup addition process
  - New structure: (available in current CMS SW release)



repeat until all interactions are processed, including “hard-scatter”

- content of individual interaction events is dropped once they have been processed:
  - only 1 event in memory at any given time
  - substantial performance gains (numbers later)
  - makes sLHC simulations possible within CMS infrastructure



# CPU/Memory Performance



- Some timing/performance results:

- Single neutrino events with 8 TeV Pythia minbias events simulating individual interactions
- Events processed merely through pileup addition and digitization

Scenario	CPU/ev (s)	RSS (MB, 100 ev)
Summer12 [-2,2], 50 ns, <PU> = 21	5.2	976  (3 BX only)
[-12,2], 25 ns <PU> = 20	12.6	1186  (16 BX)
[-12,2], 25 ns <PU> = 40	27.4	1518  (16 BX)

- Memory Reduction for a sample with 100 interactions/crossing:

Pileup Configuration	VSIZ (MB)	RSS (MB)
Old Mixing Software	2520	2020
New Mixing Software	1283	933

# Simulation meets Computation

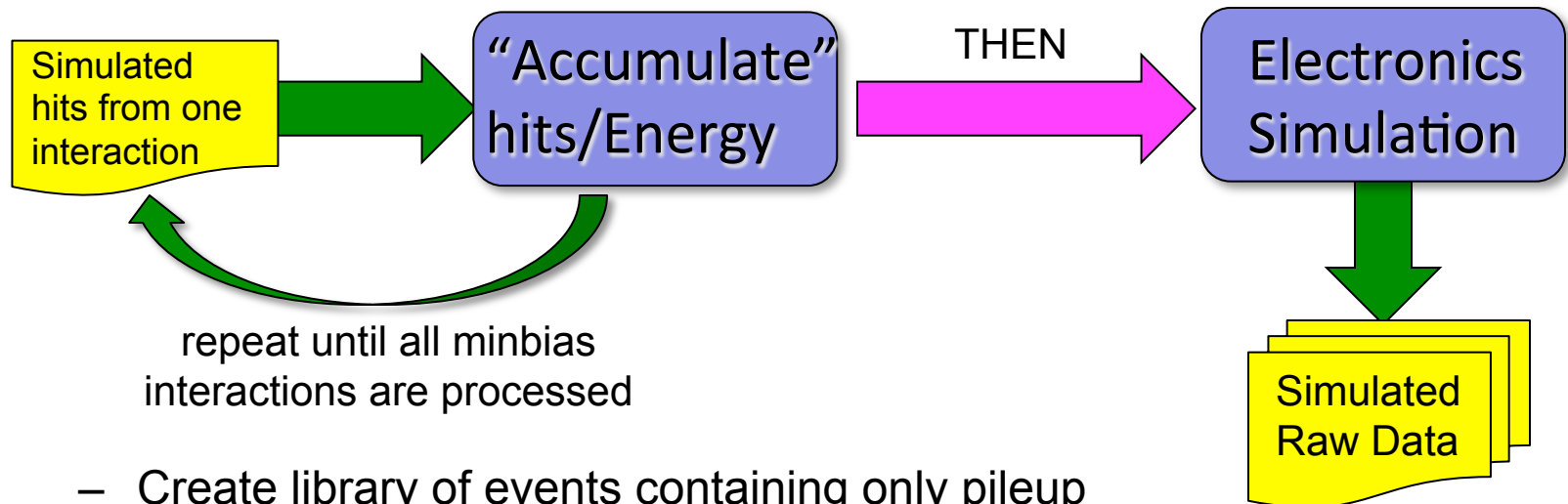


Even if the events are read sequentially, it still will require more than 2000 minbias events to produce a single MC event with appropriate pileup at sLHC luminosities

- nightmare for computing infrastructure if huge minbias event files have to be made available to each compute node for MC production

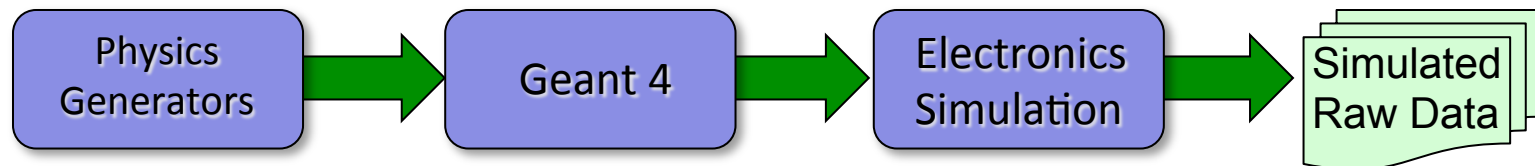
- Potential Solution: “Pre-Mixing”

- For the pure minbias pileup simulation,

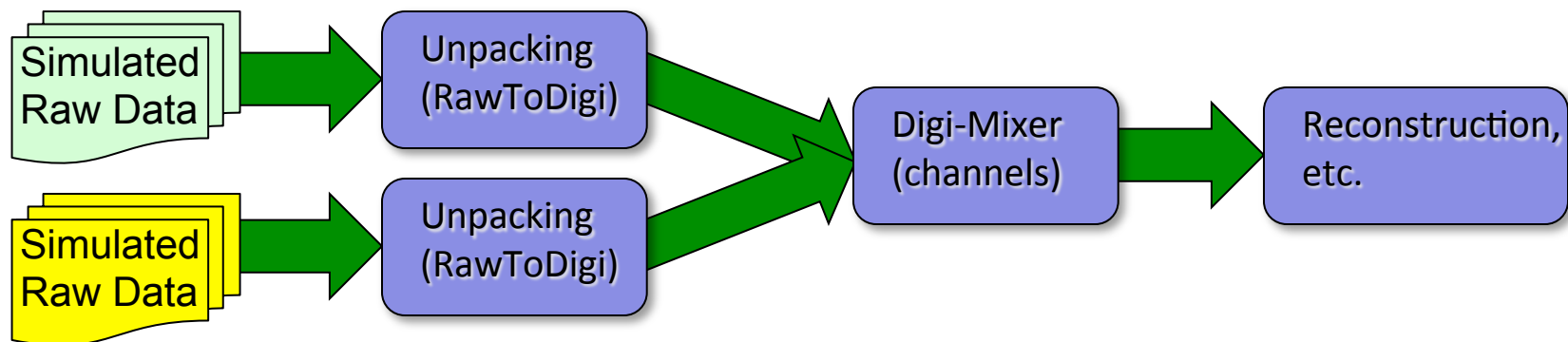


- Create library of events containing only pileup contributions, following pre-determined luminosity profile to calculate how many interactions to include

- Next, “hard-scatter” sample is created and processed through the digitization step with no pileup, convert to Raw data to save space



- Finally, streams are merged



- employs “DigiMixing” scheme where Digis are combined at the individual channel level
- **only 1 pileup event** is needed for each “hard scatter” MC event
  - much simpler for computing infrastructure

# Pre-Mixing Status



- Infrastructure exists
  - (see next slide)
  - currently under validation
  - computing workflows tested and functional
- Caveats
  - In order to correctly combine signals that would have been just below zero-suppression thresholds, events should be produced without zero-suppression
    - impractical for tracking systems
      - impact of this omission under study
    - not a standard production workflow for calorimeters
      - extensive testing/tweaking required
- planned for deployment in time for large scale production of MC for 13 TeV running

- CMS has developed a framework to combine events at single-channel level (Digitization)
  - or at higher level: Reconstructed objects
  - Data or MC
- Has been used for
  - Trigger studies (Data-on-Data overlay)
  - Tracking & Vertexing efficiency studies (Track embedding)
  - Calorimeter noise studies (Data-on-MC overlay)
- Planned for deployment as a tool for 13 TeV running
  - If zerobias data events used, automatically gives a “post facto” correct description of most pileup effects, including beam backgrounds, noise, etc.

- Pileup Simulation is not an easy problem
  - generator issues
    - (not discussed here)
    - do our generators actually match the physics?
  - CPU/Memory consumption will continue to be problematic
    - especially for sLHC
    - constant vigilance required to keep this under control
      - may require simplification of simulations
    - out-of-time pileup is difficult to study/quantify
- Current implementation very successful
  - major reworking of infrastructure has been necessary to confront the challenges of high(est) luminosity simulation
  - with modifications, ready for 13-14 TeV and even sLHC
    - more optimisation possible