# Software Readiness for 2019 Data

Norman Graf (SLAC)

HPS Software Meeting

October 14, 2020

# Logistics

- We need good estimates of our CPU needs to process the full 2019 "good" data sample
  - ~50 Billion events
  - Goal is better than 10Hz with a memory footprint of less than 1 GB
- We need good estimates of the amount of computing power we can rely on.
  - will be competing with CLAS for processing resources
  - Now would be a good time to proceed. Are we ready?
- We need good estimates of our storage needs
  - ~600TB of evio data
  - Will tape access be an issue?
- Will inform the overall HPS data processing plan
- Proposal:
  - push to resolve the remaining "critical" issues
    - Ecal sampling fractions and timing
    - SVT hit/track timing, trigger phase, "monster events"
  - clean up current git branches
  - merge with master
  - select an SVT alignment
  - process "sample partitions"

# Sample Partitions

- Partitions sampled ('.*04[12]$') from the list of 282 "good" runs:
  - Roughly 150M events (142937602)
  - Roughly 3‰ of the 2019 "good" data (857 / 276339)
- Using:
  - EvioToLcio
  - pass1-dev_fix
  - PhysicsRun2019FullRecon.lcsim
  - HPS_TY_iter4
- Get the following statistics running on Jlab farm:
  - Job Count       : 857
  - CPU Count       : 857
  - CPU Days        : 134.7
  - Wall Days       : 136.7
  - CPU/Wall        : 0.986
  - MemUsed/Req     : 0.485
  - MemReq/Slot(GB) : 1.300
  - MemUsed/Job(GB) : 0.631
- Averaging ~13Hz
- 2.2TB of output.