

# Mating Multi-File `stdHep` Samples to `mcProduction`

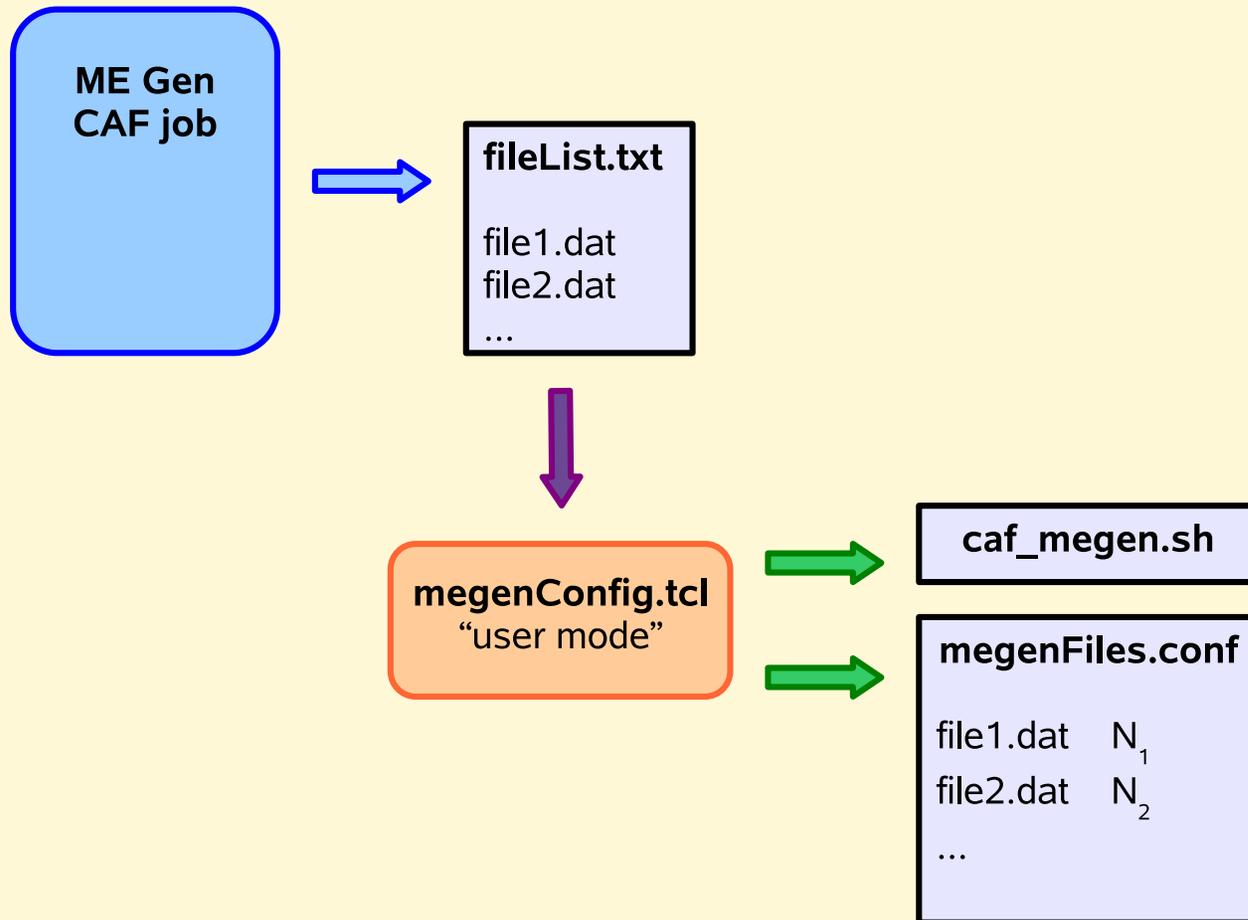
Jason Galyardt

Carnegie Mellon University

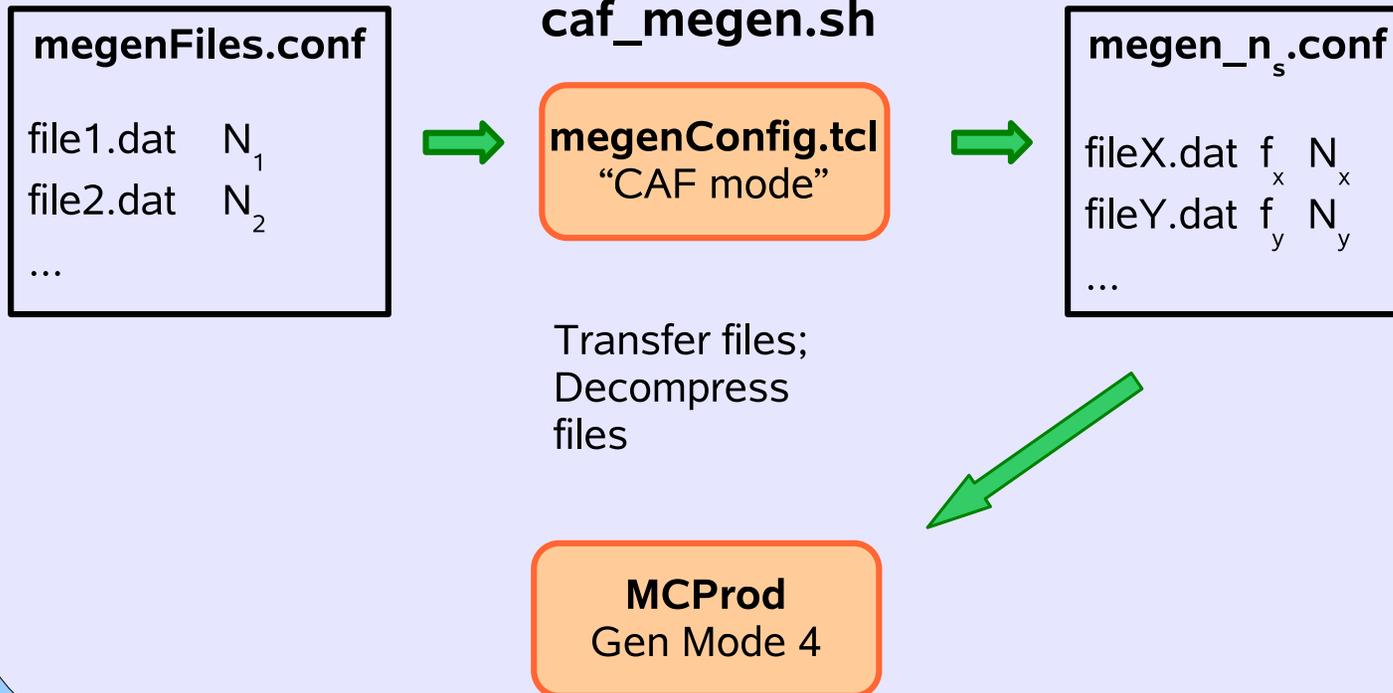
- The ability to generate a multi-file MC sample in `StdHep` format, is a necessity with some matrix element (ME) generators.
- Problems encountered when attempting to use such a sample with the current `mcProduction` machinery:
  - ◆ The number of `StdHep` files and their respective event counts are unlikely to match the event count distribution defined by the `mcProduction` machinery (`make_joblist.pl`)
  - ◆ `StdHep` to `HEPG` direct file conversion has problems (missing/incorrect banks, difficulty redecaying events)
  - ◆ `LesHouchesModule` can only read in one `StdHep` per `AC++` execution cycle (unlike `DHInput`)
- `megenConfig.tcl` was written to solve these problems.

- 2 modes (both modes run `make_joblist.pl`):
  - ◆ “user mode”  $\Rightarrow$  determine event counts for `StdHep` files.
  - ◆ “CAF mode”  $\Rightarrow$  do the accounting; for a particular job segment, determine which files are needed, their first events, and the number of events to read from each file.
- Assume the user has already run a CAF job which produced multiple `StdHep` output files (all with roughly the same event count)  $\Rightarrow$  all files are on one kerberized machine in one directory

# Step 1



## mcProduction CAF Job (segment $n_s$ )



# MCPProd Gen Mode 4

- New Generation Mode for MCPProd
- Runs `cdfSim` once for each input file
- Feeds multiple input files to `DHInput` for the next AC++ exe (either `TRGSim++` or `ProductionExe`, whichever is requested).
- Still get one output file in the end of the expected size, name

- `megenConfig.tcl` works in both “user mode”
- “CAF mode” keeps failing for the last couple of days (cause should be tracked down in the next day).
- Some additional feature requests to incorporate yet.
- Can be made compatible with Rob’s “on-the-fly concat” scheme, but work needed.
- Design may change once people decide how to support long-term storage of ME gen samples (see next slide)

# Long Term Storage Issues

- The current design has the drawback that no one can guarantee that an ME gen sample will be available in the future
- HEPG files can be stored reliably, but this is not an ideal output format because
  - ◆ Run & event numbers must be either encoded in the file or overwritten later (complicating things for users).
  - ◆ Rerunning a decay package may be required, which complicates things for users.
- May be able to store `StdHep` files in FNAL's Patriot storage system; *should* have SAM support, but it's not clear at the moment.

# Conclusion / To Do

- Developed a tool to mate multi-file `StdHep` MC samples to the `mcProduction` machinery.
- In the last stage of testing for the current design
- Design may float a bit depending on decisions concerning supported storage methods.
- Should work fine with Rob's "on-the-fly concat" scheme, but needs work