

FAQ

Does the `small_data.event()` call automatically save or do I need the save call at the end?

You need to call save at the end, like the example.

How does it order the events?

They are saved in time-order.

Can I get averages for summary data or somehow count the total number of events across my threads?

If you want, say, an average acqiris waveform, you can compute the sum over events, then save the sum as shown in the example. Similarly you can also save the sum of the number of events and use the two to compute an average.

How big can the per event data be?

Wherever possible, we recommend to keep the data small (e.g. factor of 100). Performance (both writing and reading the small file) will degrade significantly. But you can make it big (see `gather_interval` info below).

What does `gather_interval` mean in the example?

`gather_interval` controls how often the data is gathered from all the cores and is written to the file. You need to set it small if you're saving large data to avoid using up all the machine memory

Is there a way in psana general to determine the number of events in a run so I can preallocate arrays?

Since `MPIDataSource` can be run in real-time while data is being taken, there is no well-defined method to return the number of events. We tend to use per-event lists instead of arrays, since they are more dynamic. There is another mode called "idx" where you can learn the number of events, but that mode doesn't work until after the run is completed. See the bottom of this example:
<https://confluence.slac.stanford.edu/display/PSDM/Jump+Quickly+to+Events+Using+Timestamps>