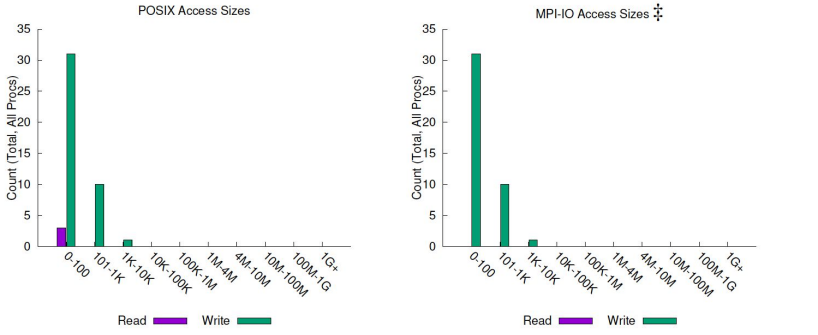
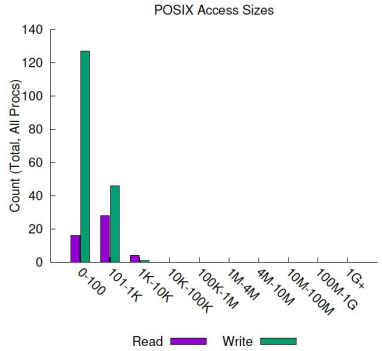
TMPI and Parallel HDF5

Test with reading few events and writing into HDF5 or ROOT.

- Both Parallel HDF5 and TMPIFile uses MPI for the parallel I/O.
- Test of doing I/O with 4 MPI processes.
  - Writing random numbers into a single HDF5/ROOT file.



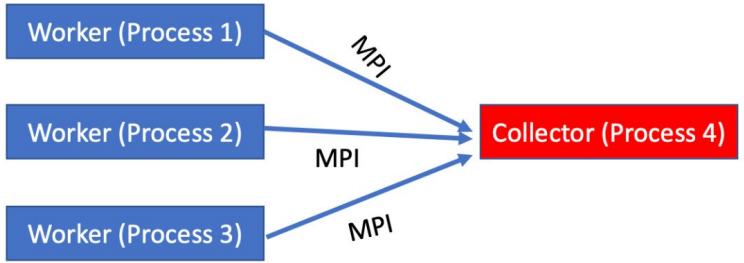
Parallel HDF5



TMPI

TMPIFile

- Derived from TMemFile
- Parallel I/O based on MPI
- Process data in parallel and write them into TFile



Communication is done via MPI functionalities

Reading/Writing into buffer is done using TMemFile functionalities

Each of the workers and collectors is one unique MPI Process or Rank.

Workers:
- Process Events (Populate TTrees or TH1D's)
- Send Processed Events to Collector Using MPI functionalities

Collectors:
- Receive Processed Events from Workers
- Merge them
- Write into disk

Comparison with Parallel HDF5

- Test was done interactively in LCRC machine.
- HDF5 (collective I/O): Write 1 Dataset
- ROOT (TMPIFile): Write 1 TBranch
- 4 ranks writing into 1 file
  - Same random number generator

IO bottleneck for TMPIFile when sync rate is small.

TMPI (sec)	PHDF5 (sec)	RANKS	EVENTS PER RANK		SYNC RATE
--	24	4	100000		1
--	2	4	100000		10
13	1.07	4	100000		100
4.5	1.17	4	100000		200
1.88	1.67	4	100000		500
1.29	2.72	4	100000		1000
*1.06	4.86	4	100000		2000

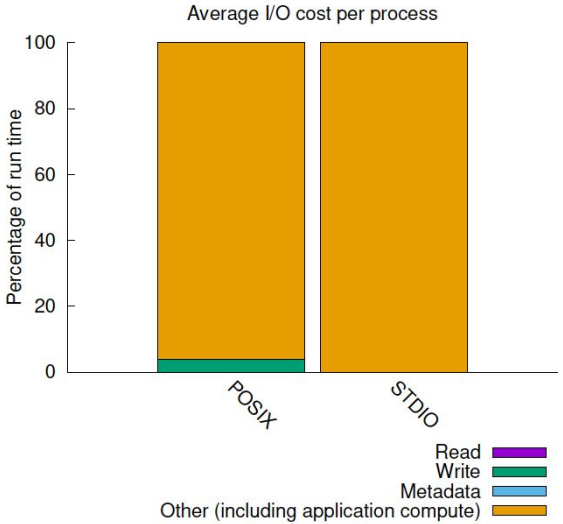
Sync rate is how often the events are written.
From 2000 sync rate, in the case of TMPI, the collector node dominates the total run time significantly.

TMPI and PHDF5 Comparison

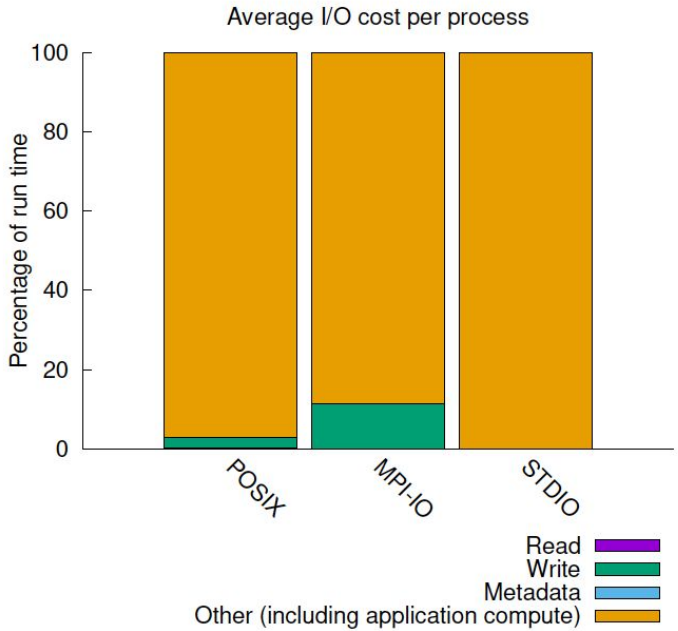
TMPI (sec)	PHDF5 (sec)	RANKS	EVENTS PER RANK		SYNC RATE
--	24	4	100000		1
--	2	4	100000		10
13	1.07	4	100000		100
4.5	1.17	4	100000		200
1.88	1.67	4	100000		500
1.29	2.72	4	100000		1000
1.06	4.86	4	100000		2000

In the case of TMPIFile, the performance depends upon distribution of Workers per collector and sync rate.

DARSHAN Logs



TMPI



PHDF5

For now Darshan cannot capture MPI related I/O calls in TMPI. Working on this.

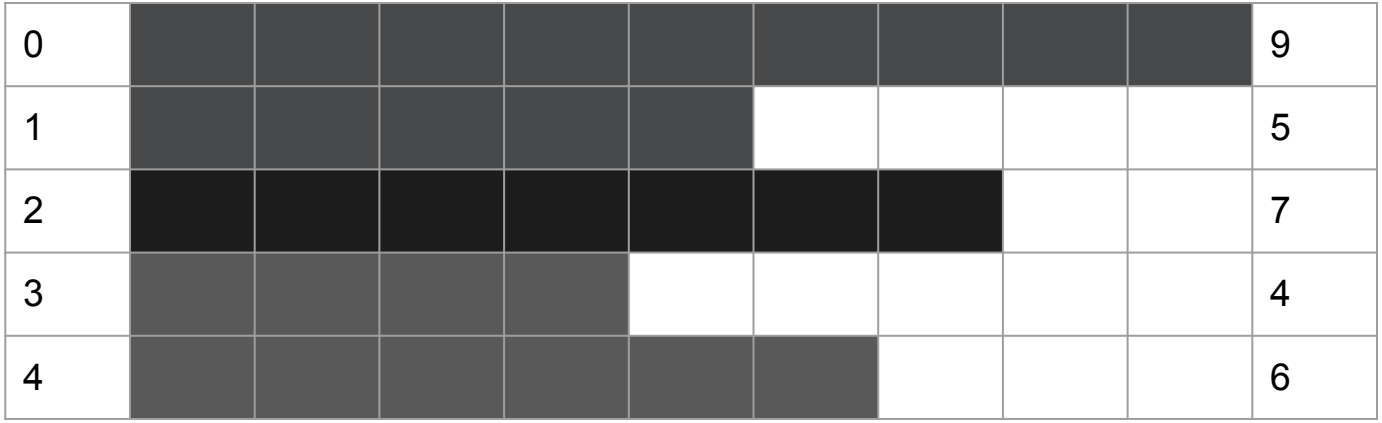
2D array for the GPUs

0									
1									
2									
3									
4									

2D matrix.

Can be used both for Matrix like calculation and also 2D arrays like structure.
Based on link

2 D Arrays for GPU



Basically a Matrix with some maximum number of columns (size of arrays)
Number of rows (number of 1 D arrays)
An additional array to store the number of elements in each row.

Can be used both for Matrix like calculation and also 2D arrays like structure.
Based on link