RNTuple Testing Using CMSSW

Dr Christopher Jones
CCE SOP

18 September 2024

Testing

- Use a standard ROOT MiniAOD input file
  - 84,000 events
- Have prototype components that can read/write RNTuple TFiles
  - have various options to control performance
- Testing procedure
  - Read the MiniAOD file
  - Write either TTree or RNTuple based file containing full content of the input

File Sizes

Format	File	Size	Relative Size
TTree	**Standard**	4.69GB	100.0%
	Fully Split	4.8GB	102.4%
RNTuple	**Tail Optimization**	4.74GB	101.2%
	Fully Split (standard)	4.75GB	**101.3%**
	Fully Unsplit	5.77GB	**123.2%**
	Partially Split	4.54GB	**96.8%**
	Partially Split & Tail Optimization	4.53GB	96.7%

Partially Split

- Wrote two RNTuple files
  - Fully split (standard)
  - Fully unsplit

- Compared size on disk for each top level Fields

- Force those which are smaller unsplit to be written unsplit

- One Field contributed to 90+% of the size reduction
  - LHEEventProduct
    - standard : 447 MB
    - unsplit: 237 MB

LHEEventProduct Size

- For split:1 sub field takes vast majority of the size

```
LHEEventProduct_externalLHEProducer__GEN.obj.weights_._0.wgt [#0]  --  SplitReal64{id:321}
  # Elements:          90720000
  # Pages:             11123
  Avg elements / page: 8156
  Avg page size:       36129 B
  Size on storage:     401872315 B
  Compression:         1.81
```

- For unsplit

```
LHEEventProduct_externalLHEProducer__GEN [#1]  --  Byte{id:248}
  # Elements:          1319471280
  # Pages:             20181
  Avg elements / page: 65381
  Avg page size:       11744 B
  Size on storage:     237021573 B
  Compression:         5.57
```

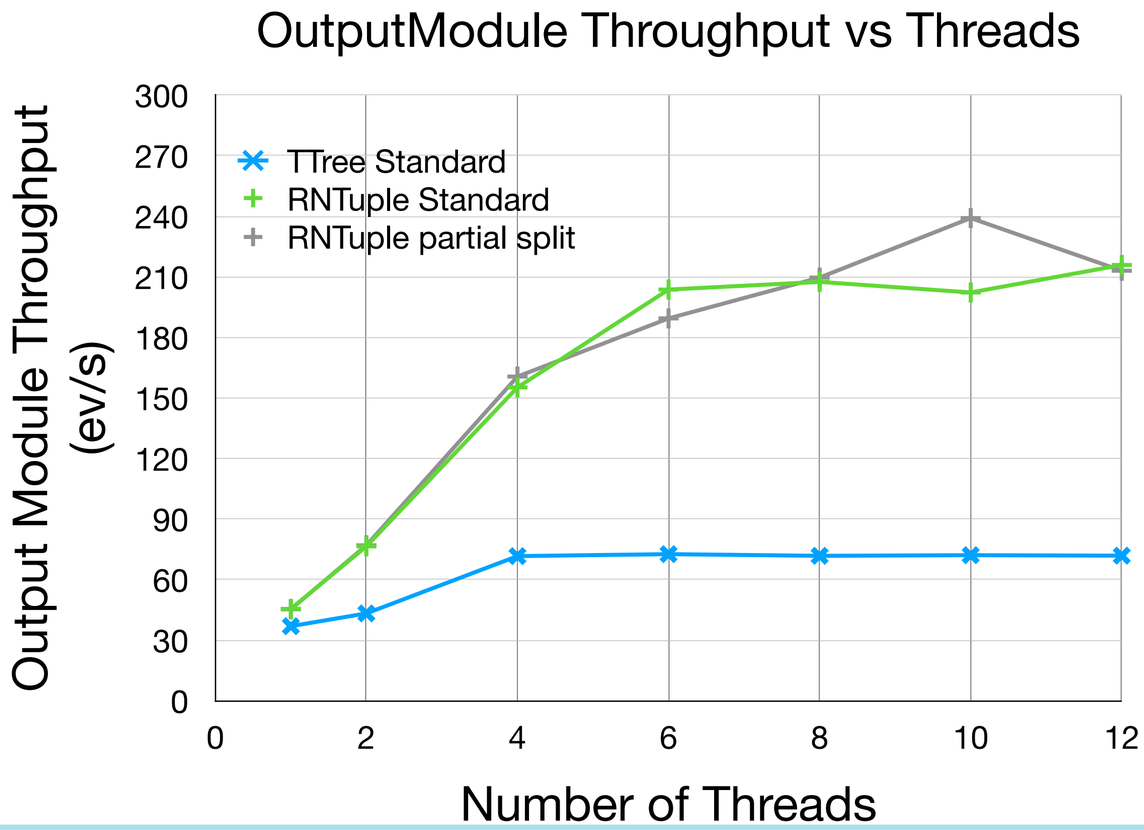
- For split TTree

```
*Br   41 :LHEEventProduct_externalLHEProducer__GEN.obj.weights_.wgt :        *
*         | Double_t wgt[LHEEventProduct_externalLHEProducer__GEN.obj.weights__]*
*Entries :    84000 : Total  Size=  726277947 bytes  File Size  =  155775319 *
*Baskets :     1177 : Basket Size=    2958848 bytes  Compression=   4.66     *
```


Fermilab

Threading Performance

- Ran the copy jobs at different thread counts
  - Number of concurrent events always kept at 1
  - Use only time spent in the code that interacts with output file



Fermilab

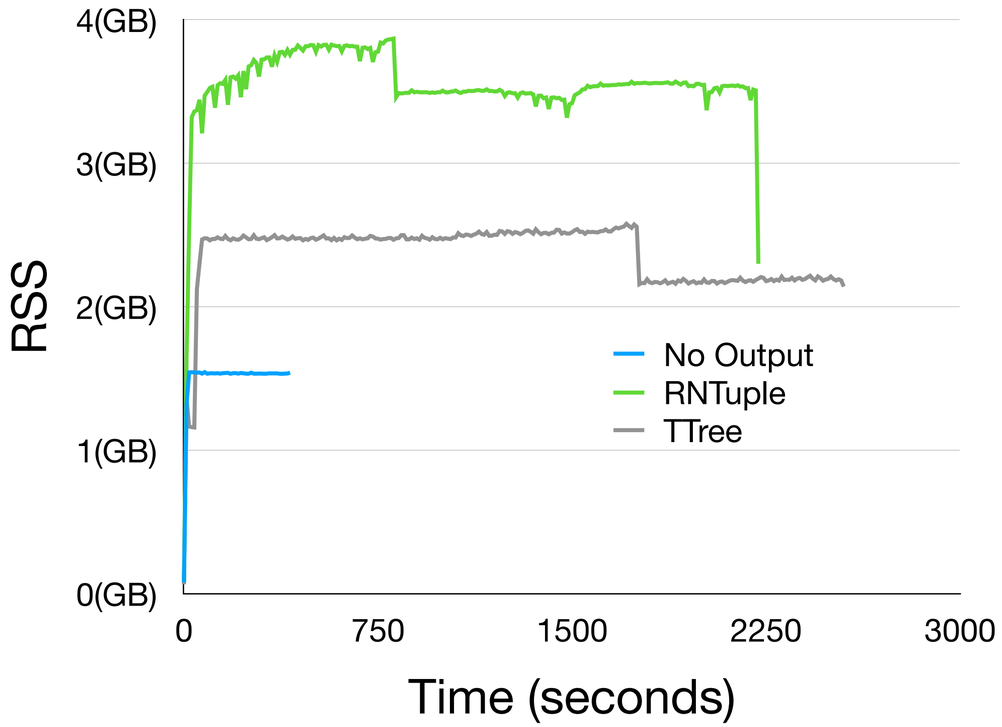
Max Memory Usage

- Monitor calls to new and delete
  - track amount of memory presently allocated and record the max for the job
- Run job with a dummy outputer to get baseline memory
- RNTuple output uses the partial split setting

Job	Max Memory	Output Overhead	Relative Overhead
Dummy	1.34GB	0GB	
TTree standard	2.33GB	0.99GB	1.00
RNTuple tail opt	6.89GB	5.55GB	5.59
RNTuple	4.27GB	2.94GB	2.96

RSS Usage

- Have an external process which periodically samples RSS of running job



Job	Max RSS	Overhead
No Output	1.58GB	0GB
TTree	2.64GB	1.06GB
RNTuple	3.96GB	2.38GB

Fermilab