RNTuple Test in CMSSW: AOD

Dr Christopher Jones

CCE-SOP

02 October 2024

AOD Format

- ## 2nd largest format that CMS produces
  - ~400kB/ event

- ## Stored to tap and kept on disk

- ## Characteristics
  - 552 top level branches
  - Fully split
  - LZMA level 4 compression

‡ Fermilab

RNTuple Storage of AOD

- Difference between fully split and partial split is negligible

File Sizes

File	Size	Relative %
TTree	4,000,822,391	100.00%
RNTuple	3,675,852,517	**91.88%**
RTuple No Split	4,175,693,891	104.37%
RNTuple Partial Split	3,644,613,716	91.10%

🟰 Fermilab

Thread Scaling

- Measurement of only time spent doing output related work
- Good thread scaling compared to TTree

OutputModule Throughput vs Threads

🪕 Fermilab
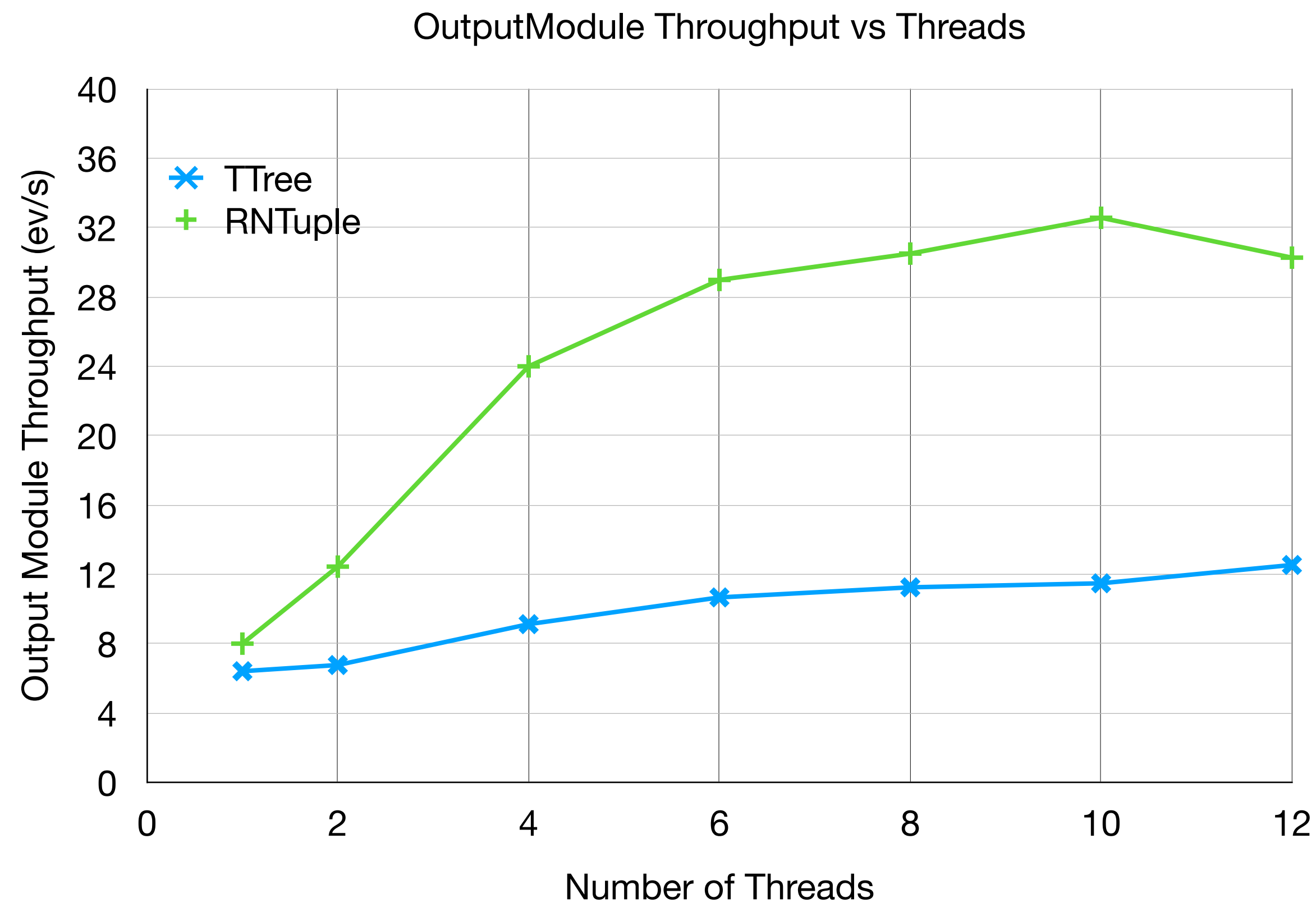
Memory Usage

- Both maximum total requested allocations and RSS are very poor

Max Total Allocation Requested

Job	Max Alloc (b	Overhead	Ratio
No Output	842,128,400	0	
TTree	1,365,851,992	523,723,592	1.00
RNTuple	6,648,308,776	5,806,180,376	**11.09**

Max RSS

Job	Max RSS	Overhead	Ratio
No Output	1336922112	0	
TTree	2213576704	876654592	1.00
RNTuple	3650424832	2313502720	**2.64**

🎺 **Fermilab**
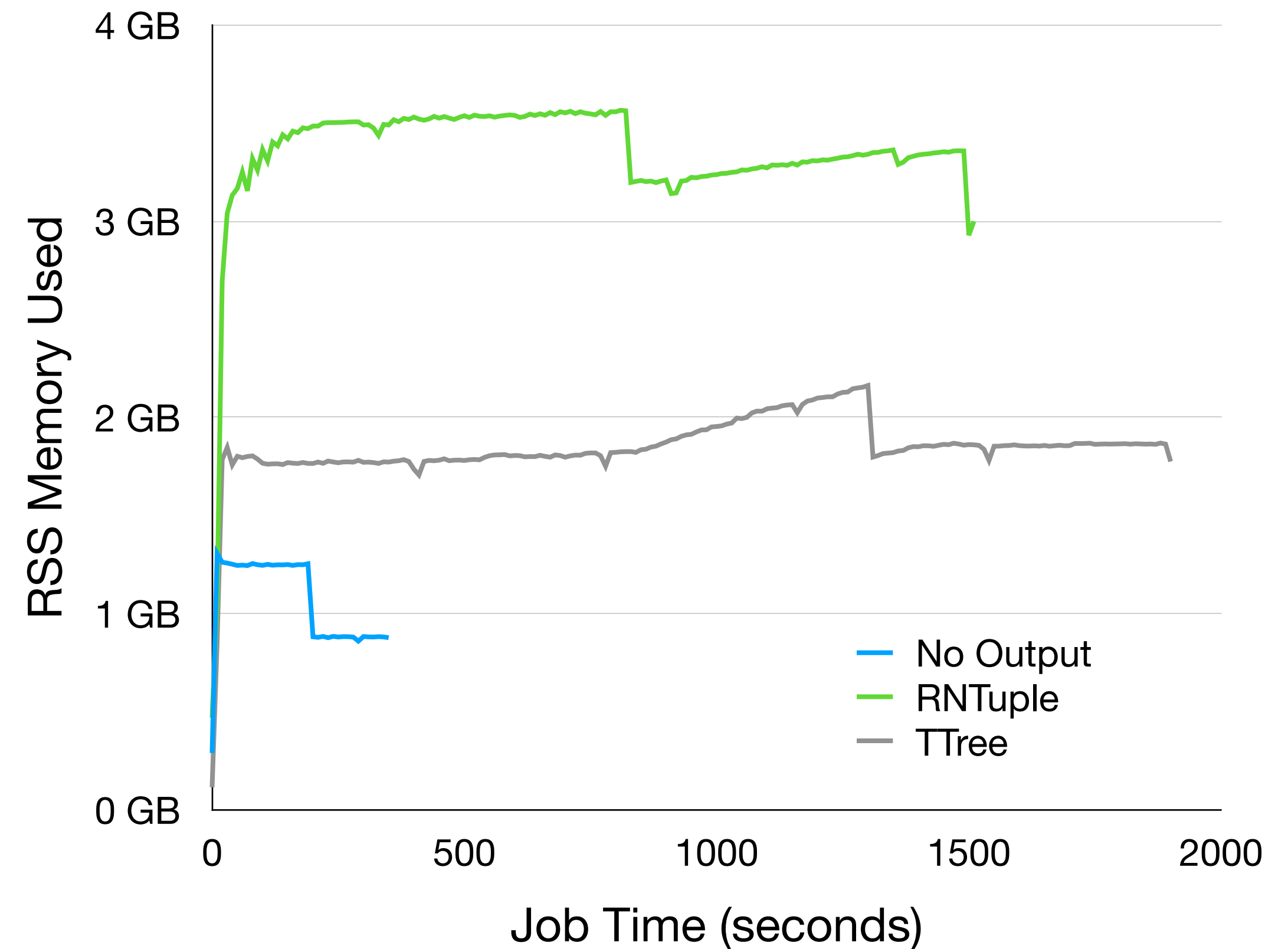
RNTuple Buffered Writing

- RNTuple has an option to turn off buffered writing

- Buffered writing aids in optimizing file layout
  - puts related pages together on the file
  - decreases file size by ~ 4%

File Sizes

File	Size	Relative %
TTree	4,000,822,391	100.00%
RNTuple	3,675,852,517	**91.88%**
RTuple No Buffered Write	3,849,309,650	**96.21%**

**🎇 Fermilab**

Memory Usage: No Buffered Write

- RNTuple has an option to turn off buffered writes
  - buffered writing helps with
- Memory usage improved, but still not very poor

Max Total Allocation Requested

Job	Max Alloc (b	Overhead	Ratio
No Output	842,128,400	0	
TTree	1,365,851,992	523,723,592	1.00
RNTuple	6,648,308,776	5,806,180,376	11.09
RNTuple noBufWrite	3,652,409,152	2,810,280,752	**5.37**

Max RSS

Job	Max RSS	Overhead	Ratio
No Output	1,336,922,112	0	
TTree	2,213,576,704	876,654,592	1.00
RNTuple	3,650,424,832	2,313,502,720	2.64
RNTuple noBufWrite	3,090,898,944	1,753,976,832	**2.00**
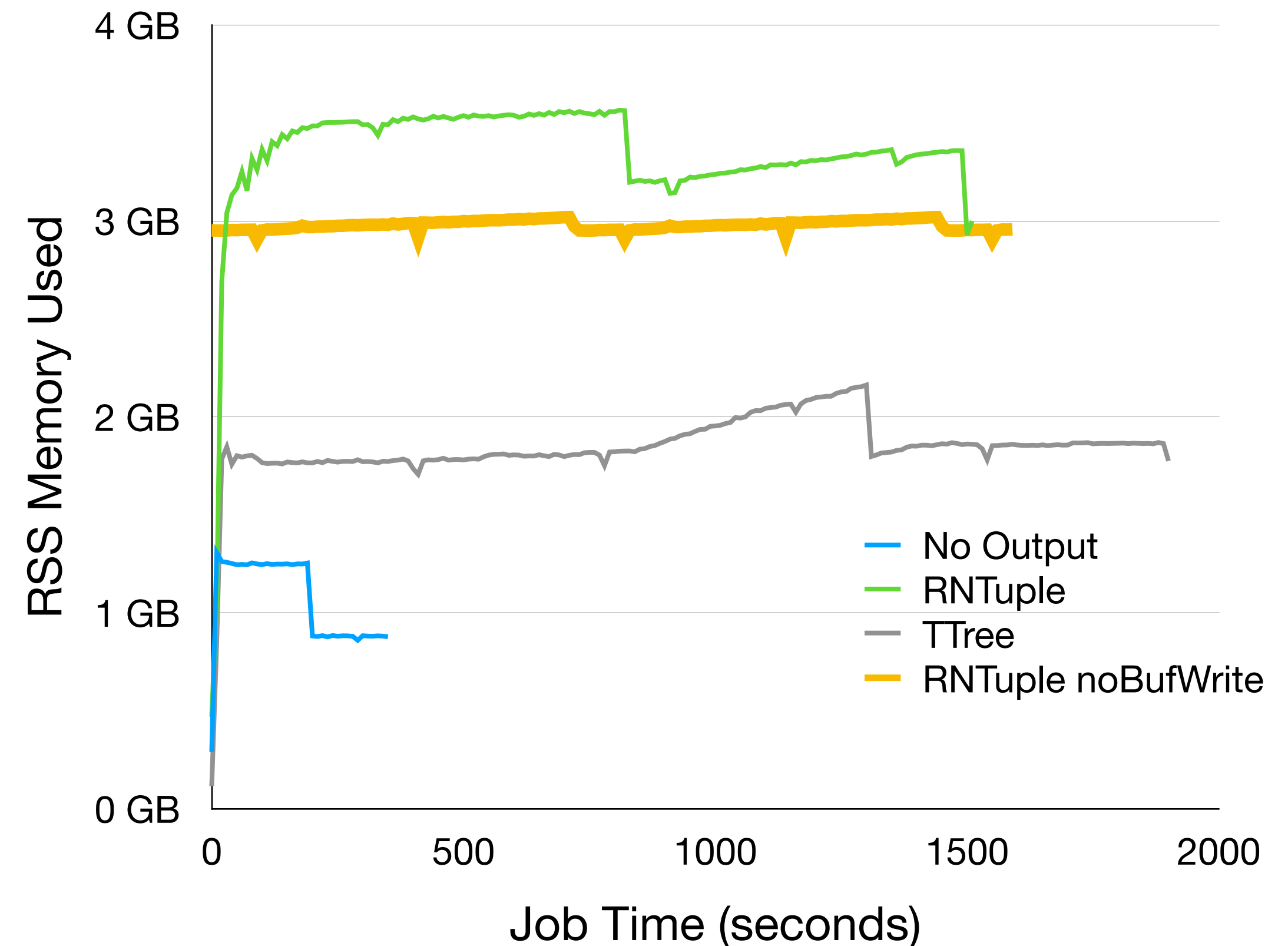
Thread Scaling : No Buffered Write

- Turning off buffered write turns of threading support

OutputModule Throughput vs Threads

🐝 **Fermilab**