

# Enhancing TTreeCache Defaults

Brian Bockelman  
(Discussion Topic)

# Goals

- Make ROOT IO:
  - Work well over high-latency links
  - Work quickly on low-latency devices.
  - Optimize for analysis use cases (assuming experiments will pick reasonable defaults).

# Available Techniques

- TTreeCache on by default: **DONE (2016?)**.
- Prefetching (TFile.AsyncPrefetching): Read event clusters in separate thread prior to the first requested.
  - Not enabled by default (believed to deadlock CMSSW; issue not triaged).
- “Miss Cache”: When a cache miss occurs, allocate a buffer for the entire event cluster and prefill it with all active branches.
  - <https://github.com/root-project/root/pull/240> Stalled!

# Potential Pitfalls

- What can go wrong?
  - **Incorrect training is forever:** read patterns that differ after the training period are always un-optimized.
    - Miss cache “fixes” this because penalty for incorrect training is significantly decreased.
    - Now that we have the “prefill” mechanism, can we simply re-train every file?
  - Do we need to **change the “drop-behind” behavior**? Once we go beyond the current event cluster, its contents are dropped. Should we triple-buffer?
    - One buffer for the previous event cluster.
    - One buffer for the current event cluster(s).
    - One buffer for the event clusters currently being prefetched.

# Potential Pitfalls

- How do we detect a “random event access” use pattern? What should we do when it is detected?
  - Example policy: when more than 10 event cluster skips are detected per file, only use miss cache.
- What considerations should be made for multiple TTrees per file?
  - Should we really launch a prefetch thread per TTreeCache?
  - Should we optimize only the biggest TTree? Should we lock TTrees below a certain size into memory?
- Low-latency devices (NVMe, SSD): TTreeCache and friends are relatively computationally expensive (we think!) compared to cost of reads from an NVMe-class device. Should we detect this case and auto-disable:
  - Proposal: If the EMA of read operations is below 1ms, then disable TTreeCache at next prefetch event.