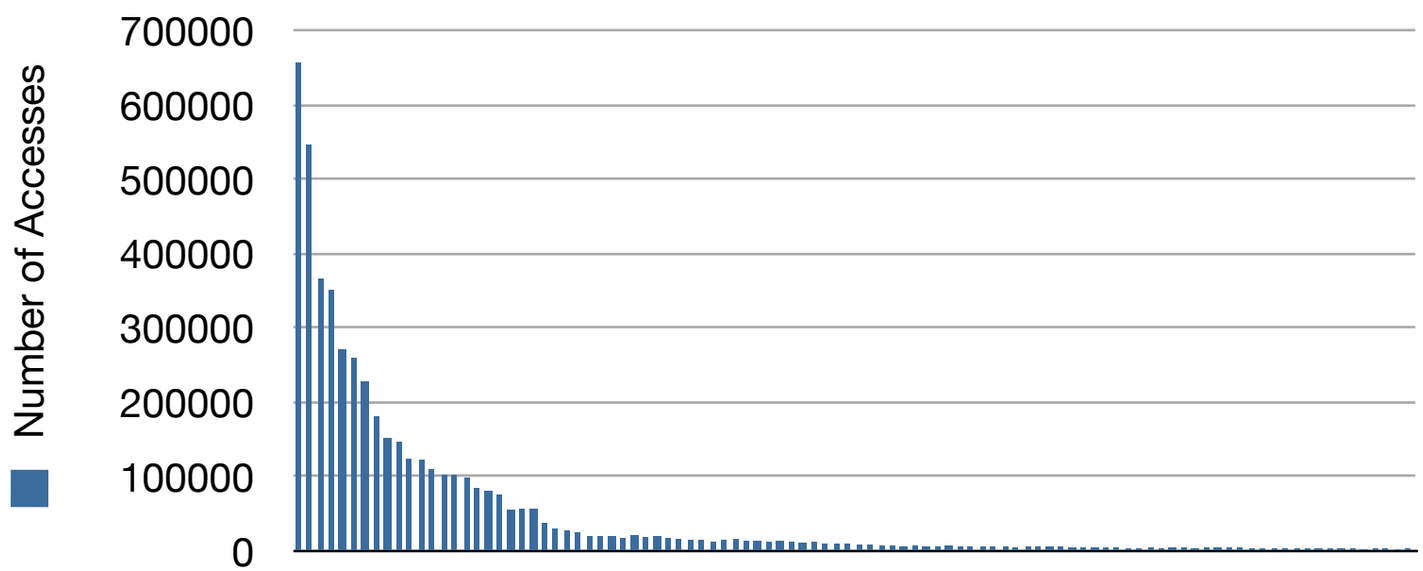
ATLAS Demonstrator: PanDA Dynamic Data Placement

Kaushik De, Tadashi Maeno, Torre Wenaus, Alexei Klimentov, Rodney Walker, Graeme Stewart



The Problem

- At the moment jobs go to pre-placed data
- However, pesky users don't use the data we carefully place for them
- So we need a more responsive system:
  - Don't pre-place data
  - Trigger on demand replication to Tier-2s
    - But queue jobs where the data is **now**, albeit with rapid rebrokering
  - Clean up when T2 storage gets full (LRU - DDM popularity service)
- Do this with the tools we have in the toolbox today

What we want to learn

- Why is this interesting for anyone else?
  - Learn about the algorithms for triggering filling a cache - initially we will try 'queued jobs for this data'
    - bestSite(jobQueueTime, ~~dataMovementTime~~)
    - See what the cache turnover is with real ATLAS user activity
    - Metrics: Usage of T2 resources, global user job waiting times, global event rate, network usage
- This demonstrator can evolve as long as it talks to a system which can
  - Put data into a cache
  - Run jobs close to that cache
- So we can try to couple in to other demonstration projects