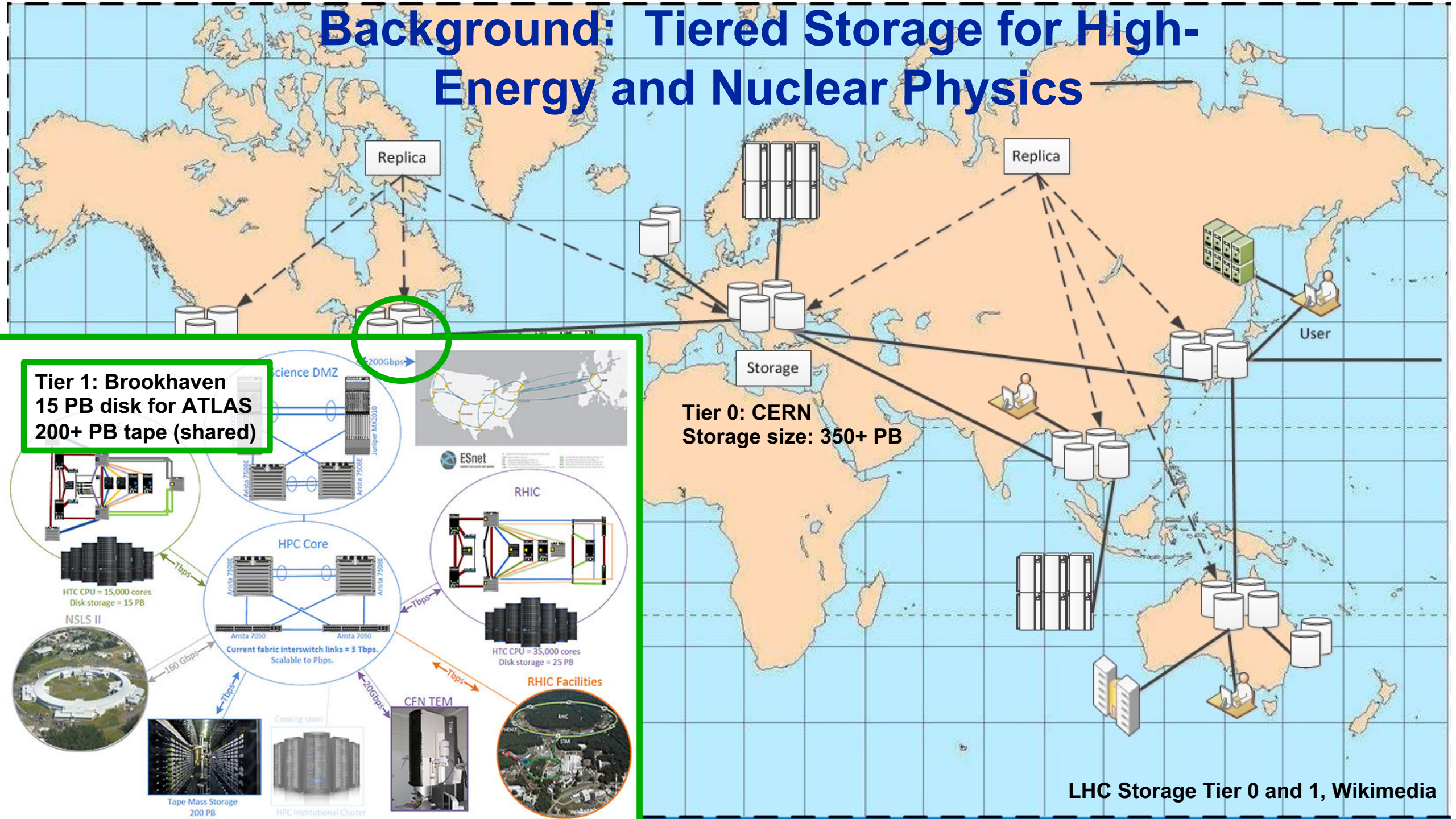
Access Patterns to Disk Cache for Large Scientific Archive

Yumeng Wang, University of California, Berkeley
K "John" Wu, Alex Sim, Lawrence Berkeley National Laboratory
Shinjae Yoo, Shigeki Misawa, Brookhaven National Laboratory

- Introduction/Motivation
- Overall file access statistics
- Tape access statistics
- Summary and Planned Work

Access Patterns to Disk Cache for Large Scientific Archive

Background: Tiered Storage for High-Energy and Nuclear Physics



Introduction: dCache Logs

- Brookhaven National Laboratory (BNL) hosts the Tier 1 site for A Toroidal LHC ApparatuS (ATLAS) experiment in United States
  - 76 PB and 80 million files for ATLAS on tape (HPSS) at BNL
  - 15 PB disk cache reserved for ATLAS data, managed by dCache
- Objective: study usage of HPSS files to find opportunities for improving data accesses, for example:
  - Are there opportunities for combining accesses to a tape?
  - Would pre-stage a group of files (aka, dataset) to disk reduce tape accesses?
- dCache log files available for the current study
  - January 2019 to March 2020
  - About 5PB of files accessed during this time period
  - 75.4 GB compression (gzipped) CSV files
- dCache is working well
  - Out of 50 requests to this dCache system, only one needs to touch the HPSS system
  - 70% of the files retrieved are used multiple times, often many times

- Introduction/Motivation
- Overall file access statistics
- Tape access statistics
- Summary and Planned Work

Access Patterns to Disk Cache for Large Scientific Archive

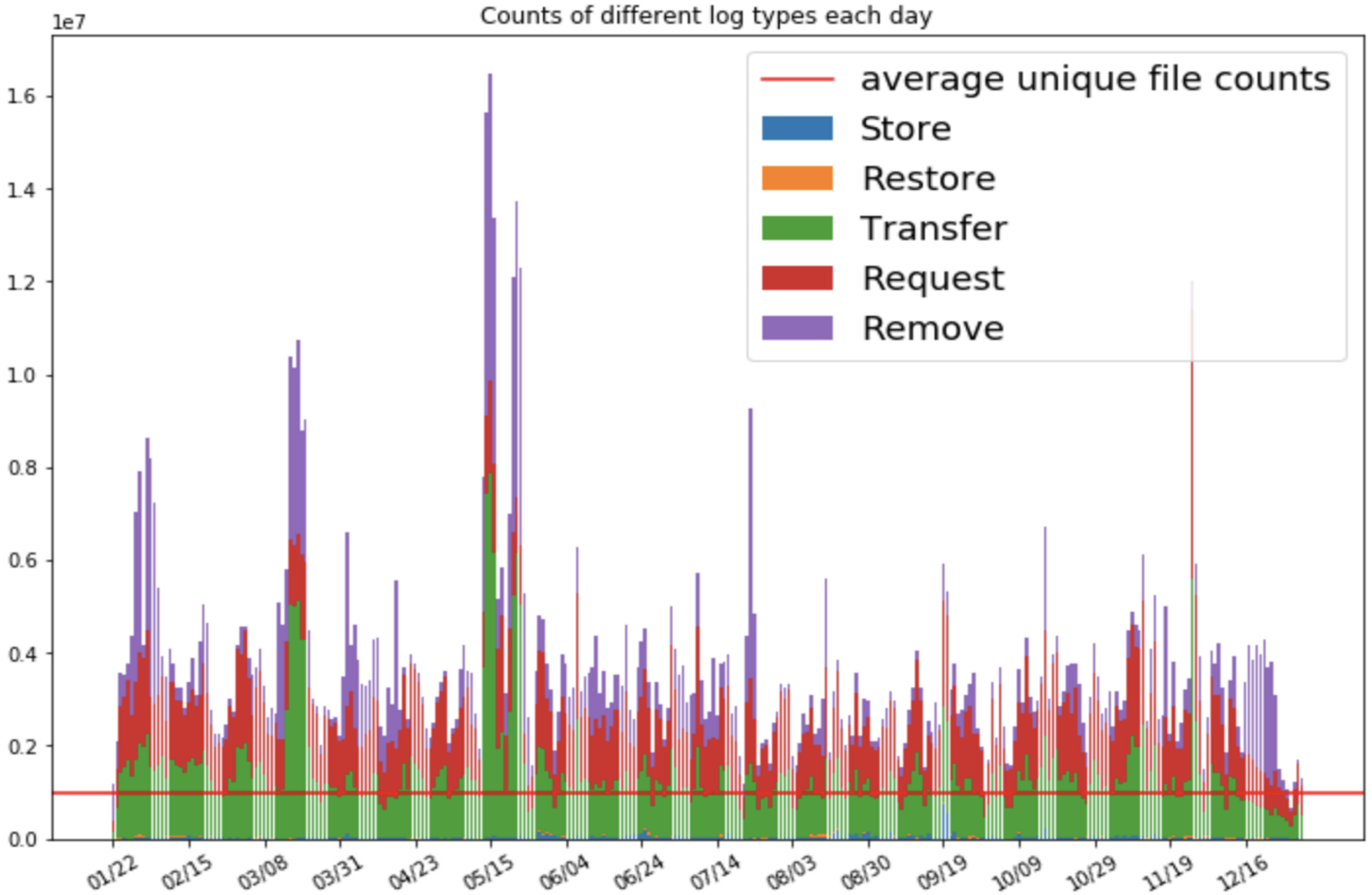
More than a million file requests per day in 2019

- **Five access types: store, restore, transfer, request, remove**

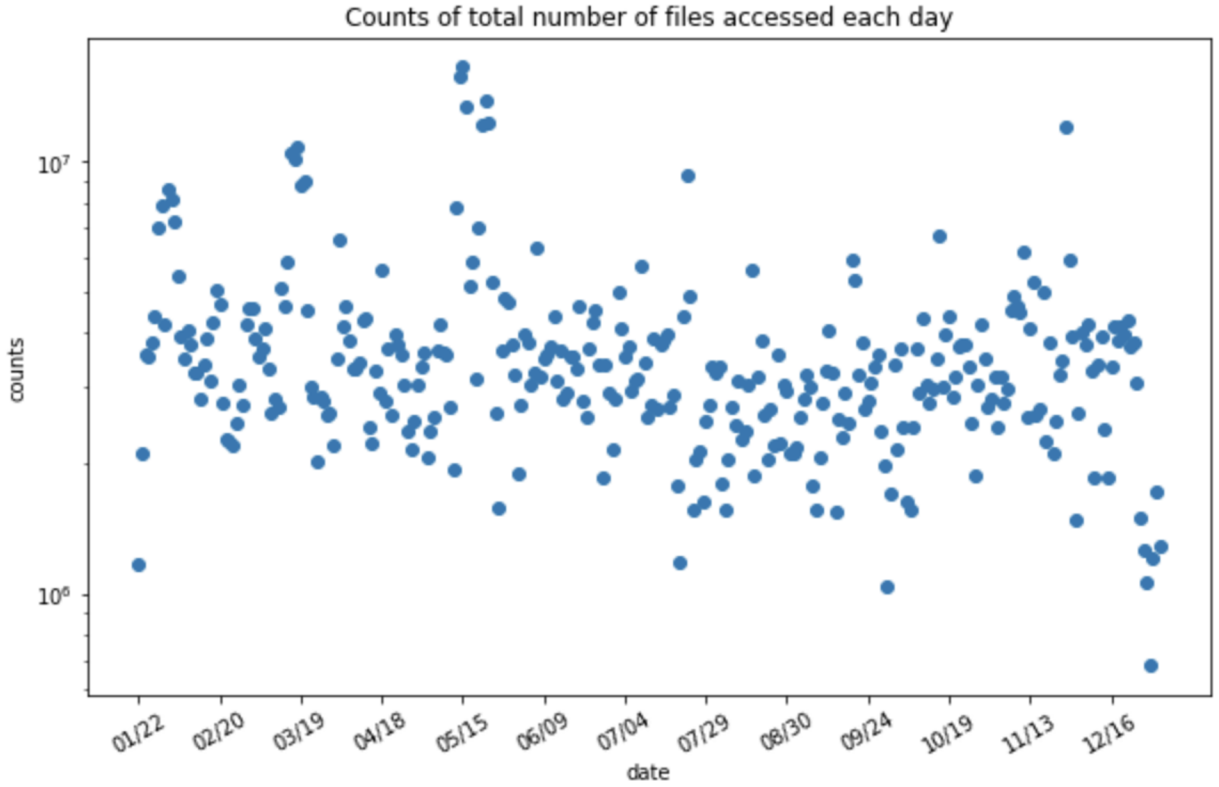
Log Count	Q1	Q2 (median)	Q3
Store	10370	20158	36530
Restore	1071	3470	8126
Transfer	905501	1164898	1464965
Request	1087876	1343616	1604839
Remove	217267	489294	992715

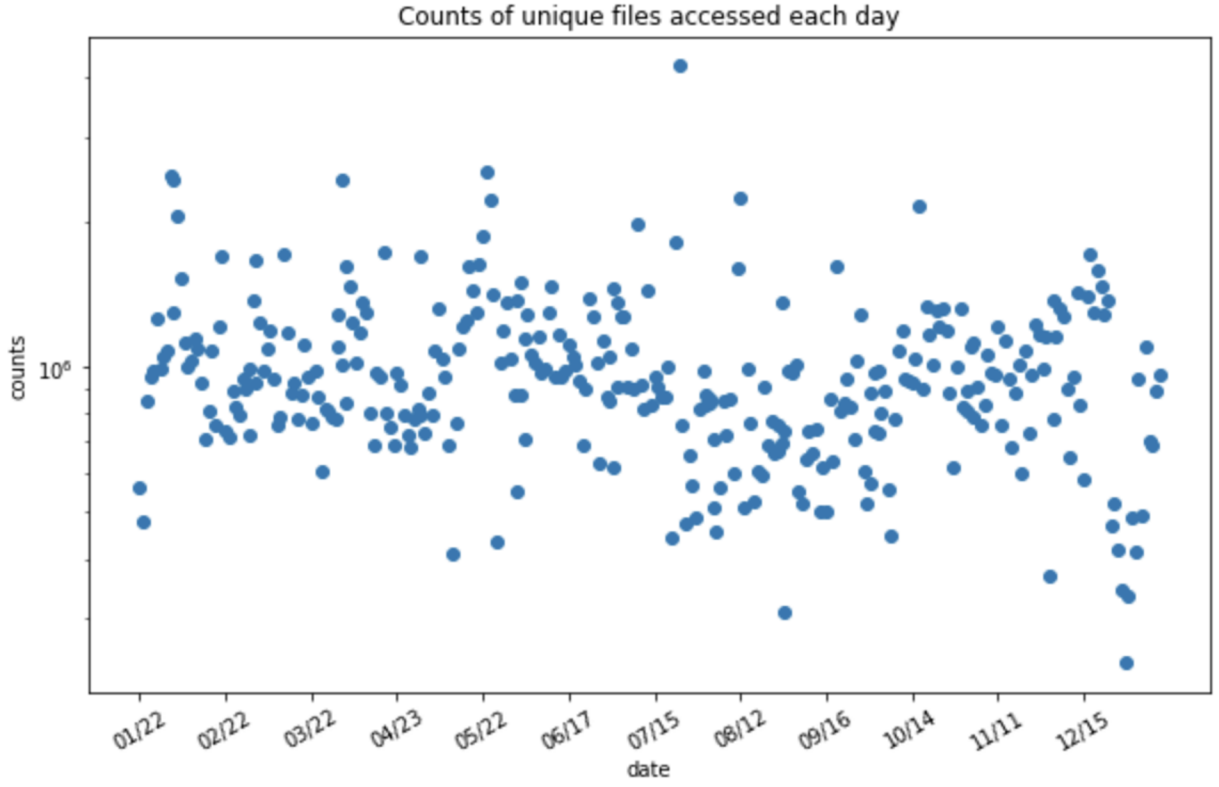
- **Majority of logs come from users transferring and requesting files**
- **Very few stores compared to transfers and requests**
  - **accessing local disk copies rather than reading from tape**

Up to 17 Million File Accesses per Day



Average File is Accessed 3 Times a Day



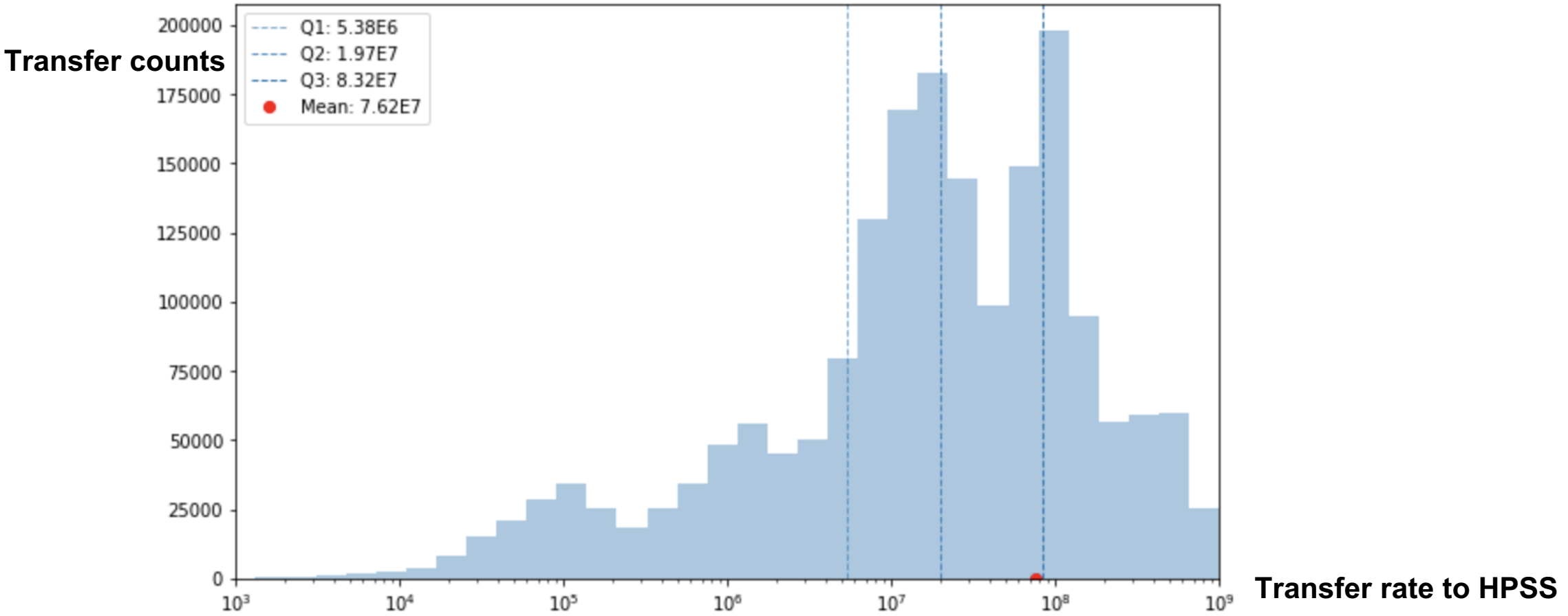


- Daily numbers of unique files: 3.5×10^6
- Daily total numbers of accessed files: 10^6
- Daily file-accessing repetition rate
  - Total number of access/number of unique files accessed)
  - First quartile as 3.02, median as 3.35, third quartile as 3.75

- Introduction/Motivation
- Overall file access statistics
- Tape access statistics
- Summary and Planned Work

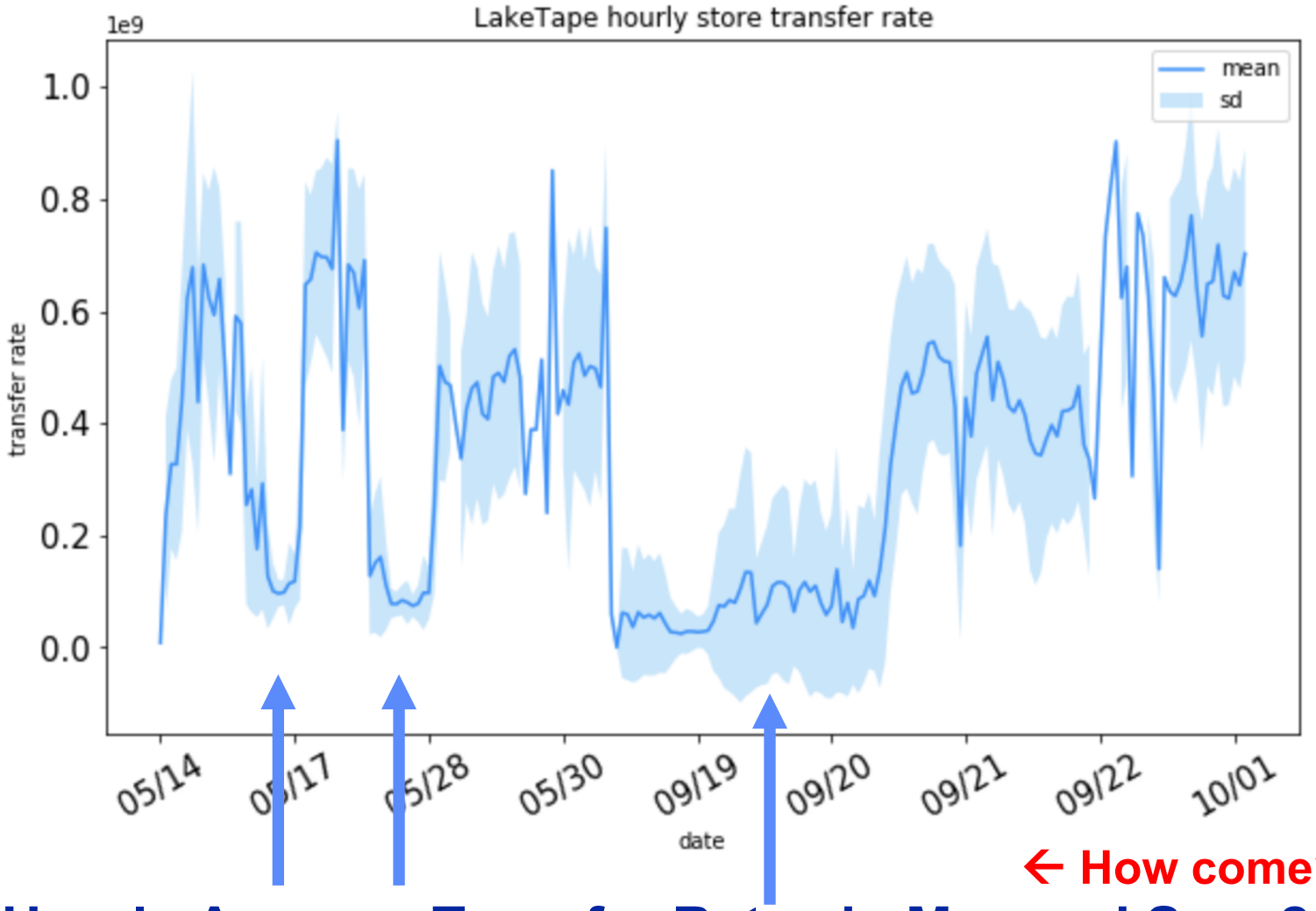
Access Patterns to Disk Cache for Large Scientific Archive

Transfer Could Reach 1GB/s, 20MB/s Typical



- First quantile: 5.4 MB/s
- Median: 19 MB/s
- Third quartile: 83 MB/s
- Mean: 76 MB/s

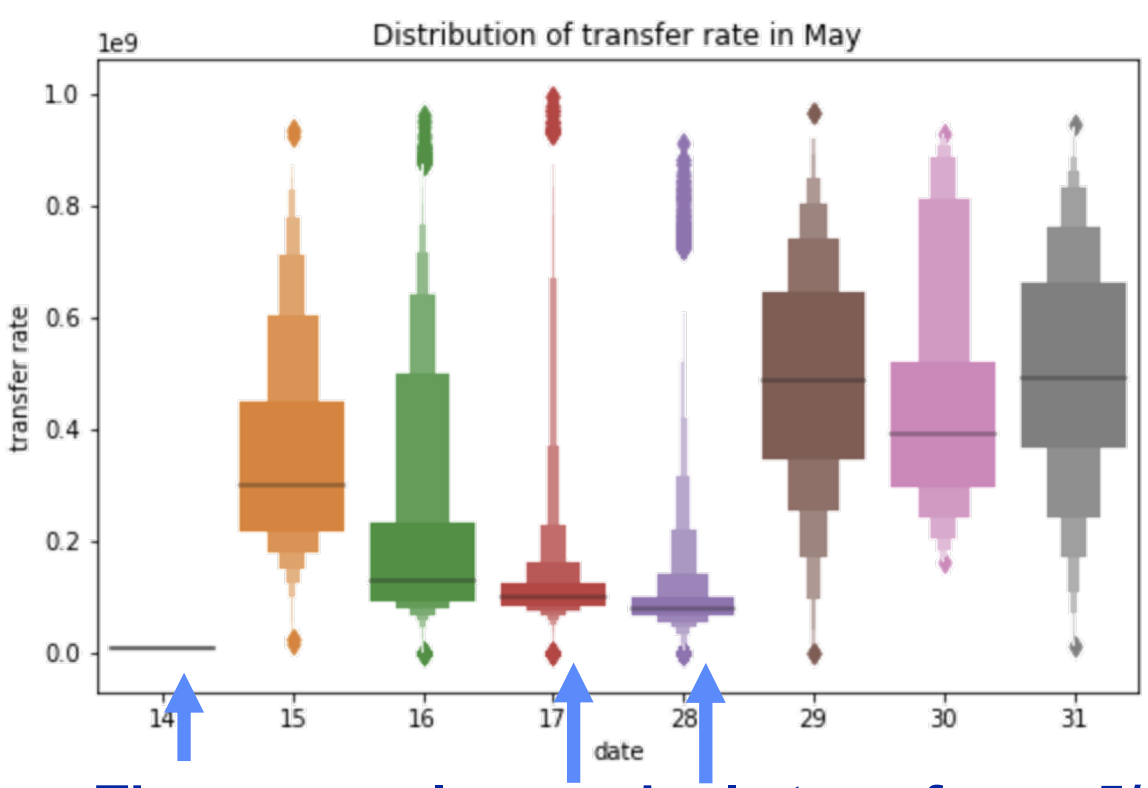
On some days, Transfers are Persistently Slow

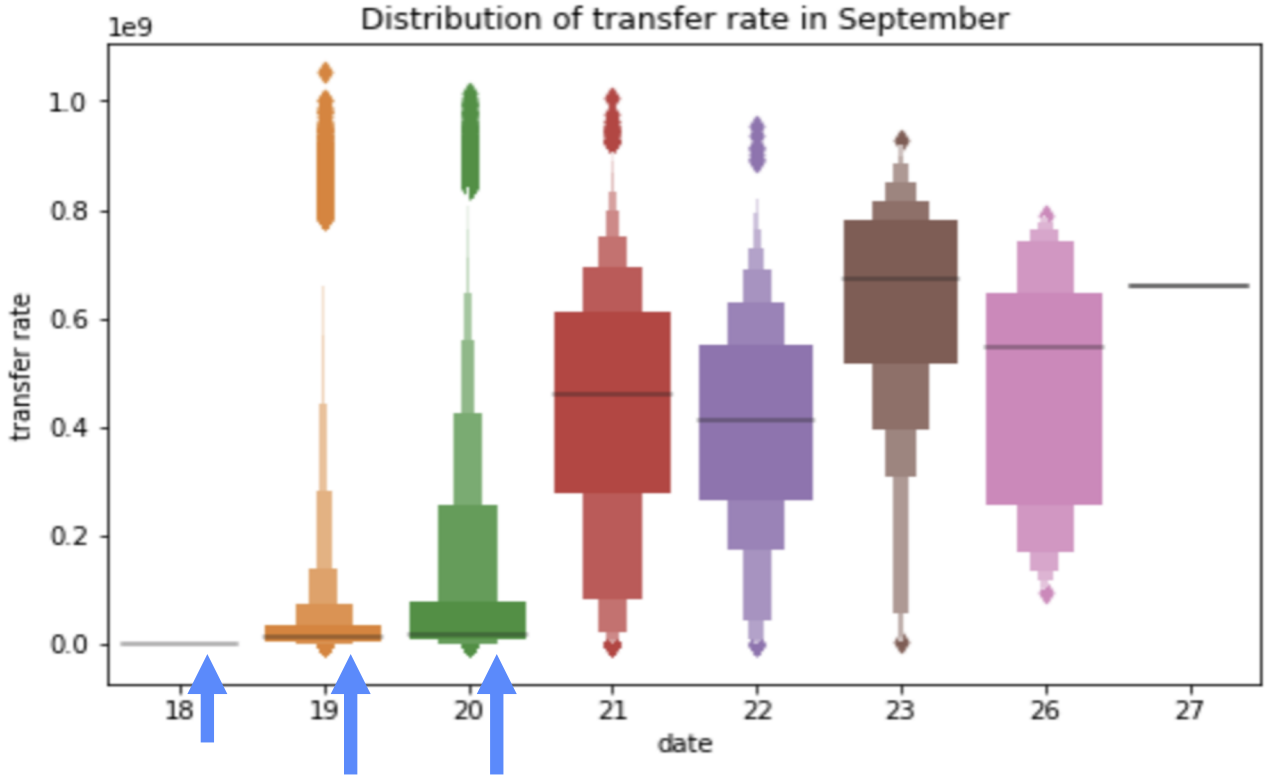


← How come?

Hourly Average Transfer Rates in May and Sep. 2019

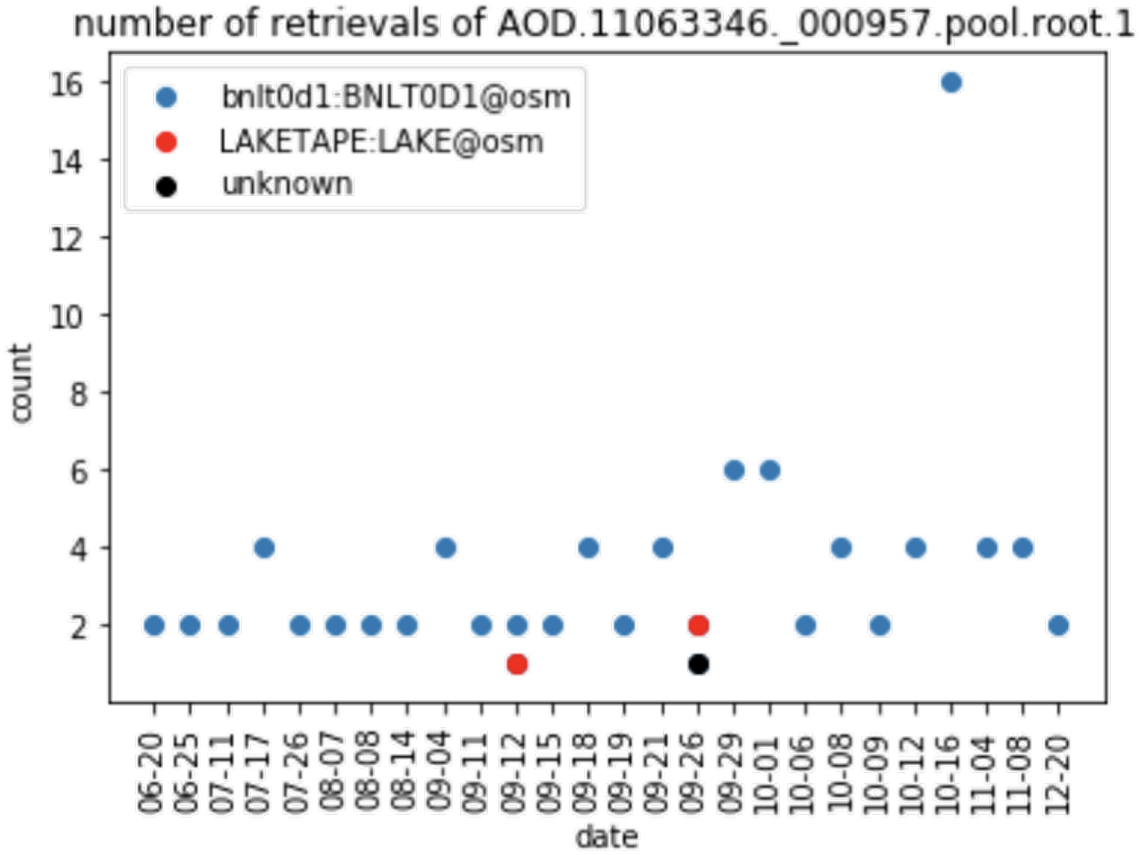
Looking Into Transfer Rates Each Day in May and Sep. 2019





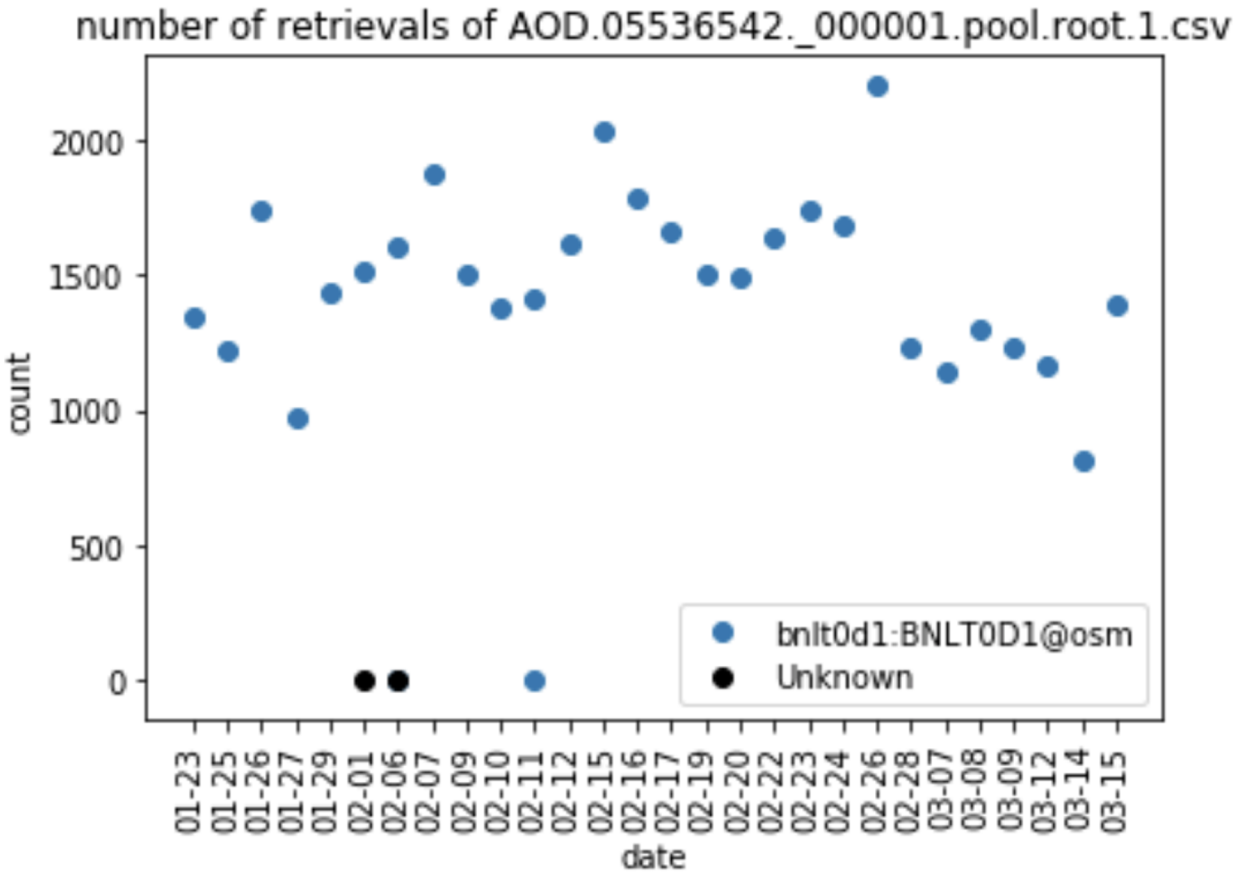
- There are only one single transfer on 5/14 and 9/18, and they happened to be slow
  - The single transfer could be fast as on 9/27
- On other four days, 5/17, 5/18, 9/19 and 9/20, there are a mix of slow transfers with fast transfers
  - Many more slow ones than fast one
  - Need additional information to further understand the slow transfers

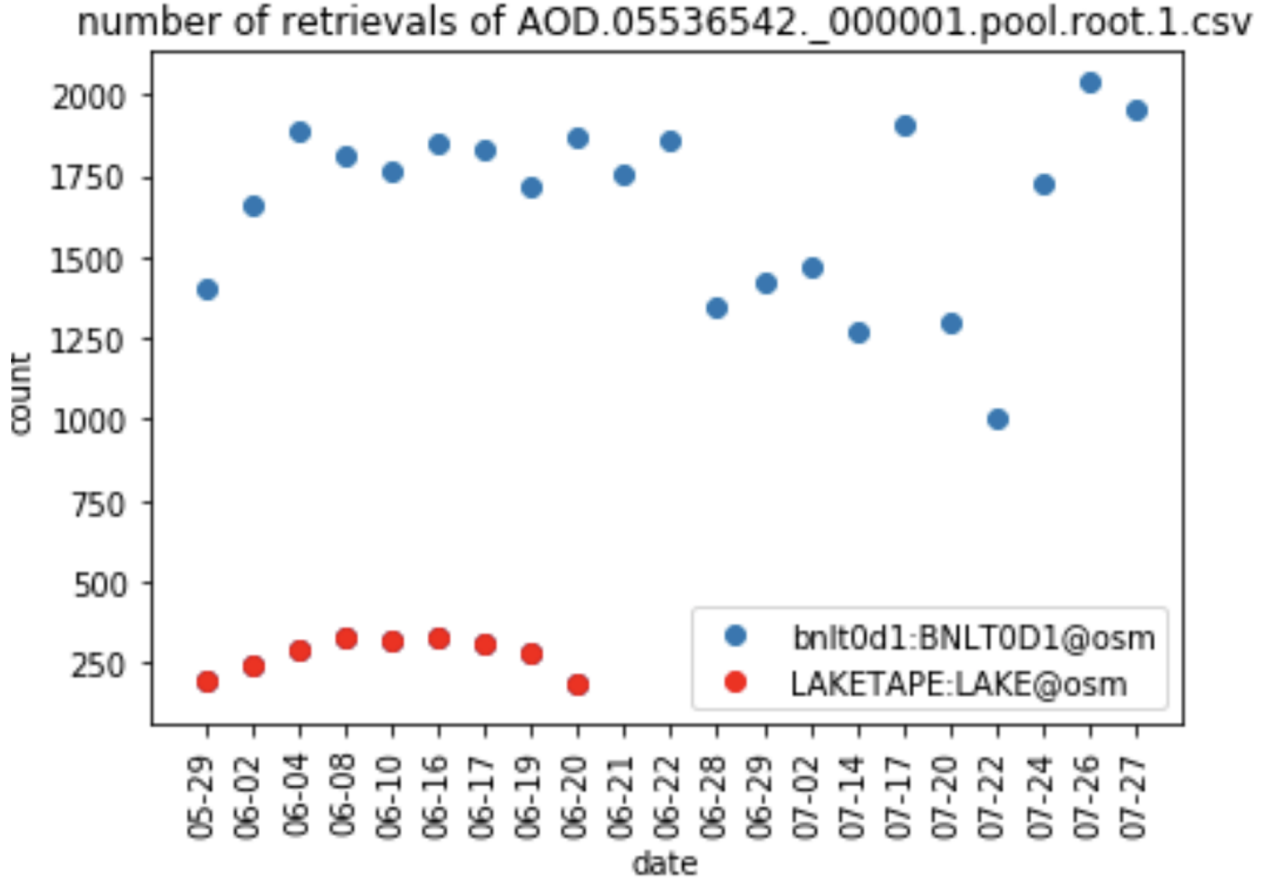
Following a Single LakeTape file



- Accessed 191 times to AOD.11063346._000957.pool.root.1
- From 1/2019 to 5/2019, always accessed from the disk copy
- On 5/28/2019, retrieved from LakeTape 3 times
- Requested again for 34 times before it is read again from LakeTape on 9/12/2019
- Don't see "remove", but retrieved again on 9/12 and 9/26. How come?

Similar Mystery with Retrievals of Files





- AOD.05536542._000001.pool.root.1.CSV and AOD.05536543._000001.pool.root.1.csv
- Repeatedly accessed from LakeTape instead of the disk copy of the file
- The reasons for this access pattern could not be determined from the current logs

- Introduction/Motivation
- Overall file access statistics
- Tape access statistics
- Summary and Planned Work

Access Patterns to Disk Cache for Large Scientific Archive

Summary

- On an average day, there are about 3.5 million file accesses to about 1 million unique files.
- Only 1 in 50 file accesses needs to go to HPSS, ~98% cache hit rate!
- Roughly 70% of files retrieved from HPSS are used multiple times. Still have potential to improve cache hit rate?
- Need additional information to understand:
  - Multiple re-readings from the tape instead of using the disk file copies
  - On some days, the transfer rates are consistently lower than other days

Additional Investigations

- **Forecasting popularity of dataset based on the data access frequency and the last data access**
  - Data categorization/temperature metric (HOT, WARM, COLD, FROZEN)
  - Proportion of never touched data would be interesting to quantify

- **Refine forecasting further**
  - data temperature to become colder or warmer
  - with workload on the disk pool, which disk pool a file reside
  - with the network traffic and performance
  - forecast for the later time window

On-Going Tasks

- **Beyond individual files**
  1. Explore information that maps a file to a dataset, then forecast access probability in the next time window
  2. Refine the prediction model, and forecast far into the later time

- **New data until 1/31/2021 for catalog dump, billing dump**
  - filename, datasetname (DSN), creation time, accesstime
  - filename-datasetname-mapping
  - Time is unixtime in ms
  - The file is in csv format

- **dCache at Nordic Data Grid Facility (NDGF)**
  - Similar dCache configuration with a larger distribution in geography
  - Learning on the BNL data, and validate with the NDGF data

Thanks!

Any Question?