Quantitative Analysis of Data Caching for the HL-LHC DataLake

Irvin Umaña Chacon

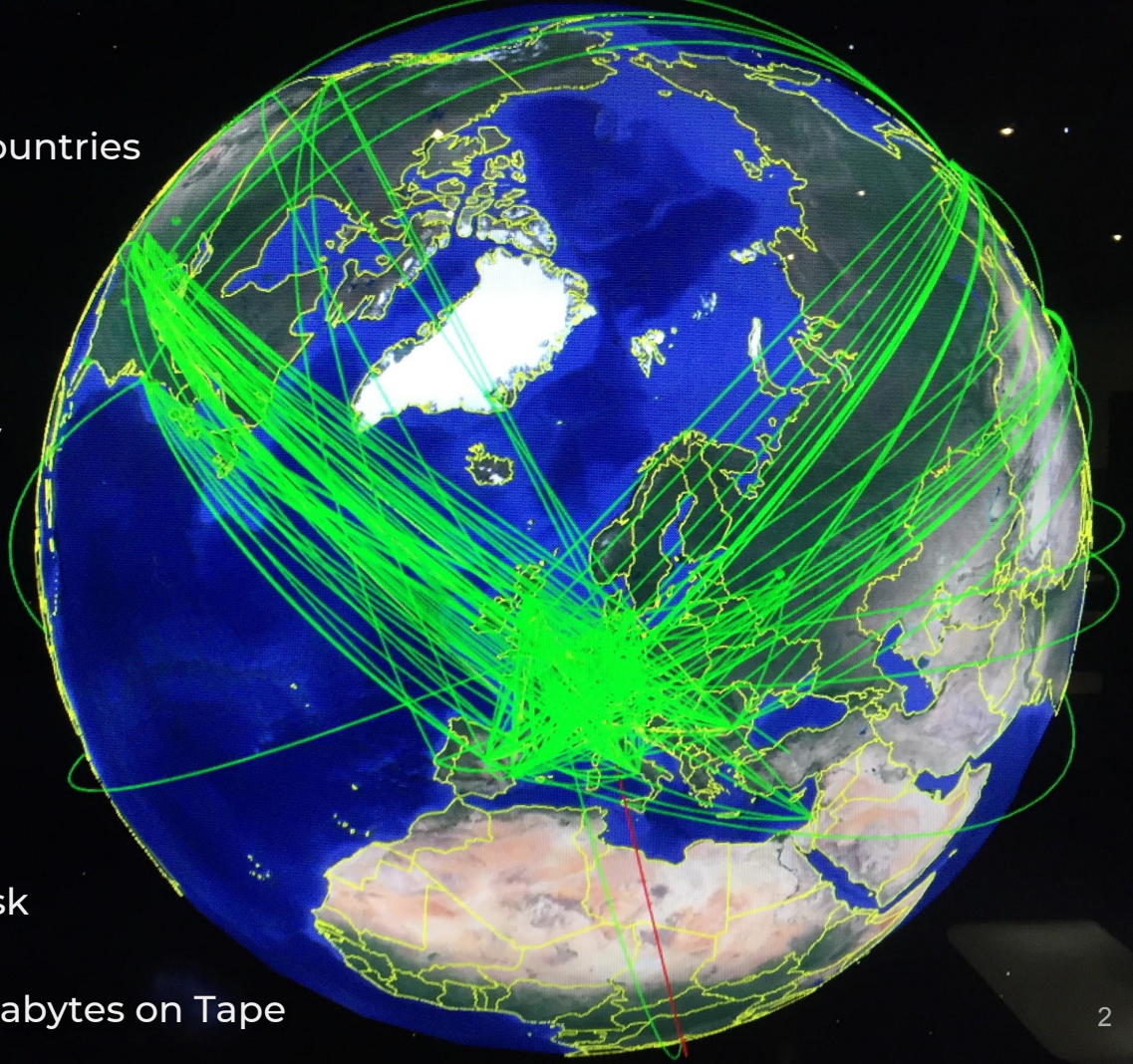
Supervisor: David Smith

IT-DI-LCG Department

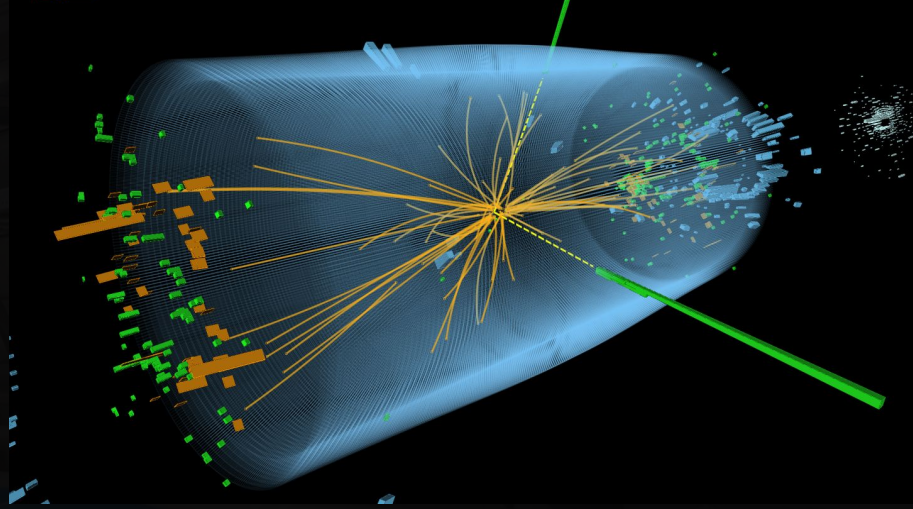
WLCG Project

Worldwide LHC Computing Grid

- 42 Countries
- 2 Million tasks per day
- 800,000 computer cores
- >170 Computing centers
- ~500 Petabytes on Disk
- ~400 Petabytes on Tape



High Luminosity LHC



Improvements and upgrades to the LHC.

Increases the number of collisions/pile up.

More data going into the grid.

HL-LHC DATALAKE R&D PROJECT

Goal

Understand the effects of having small sites in the grid processing data rather than storing it.

Seeks to understand job performance when making these architectural changes.

Considers the use of proxy servers as caches.

Motivation

Reduce the costs associated with storing data in the small sites.

For example, costs related to personnel, software, data tracking work, etc.

There are already some small sites working with the Atlas experiment that provide processing rather than storage services.

Understanding Caching in the HL-LHC DataLake

Relevant in job performance:

- Latency on requests
- Cache size
- Reduction in WAN usage
- Averaging out bandwidth needs

Tools:

XRootD

↓

XCACHE

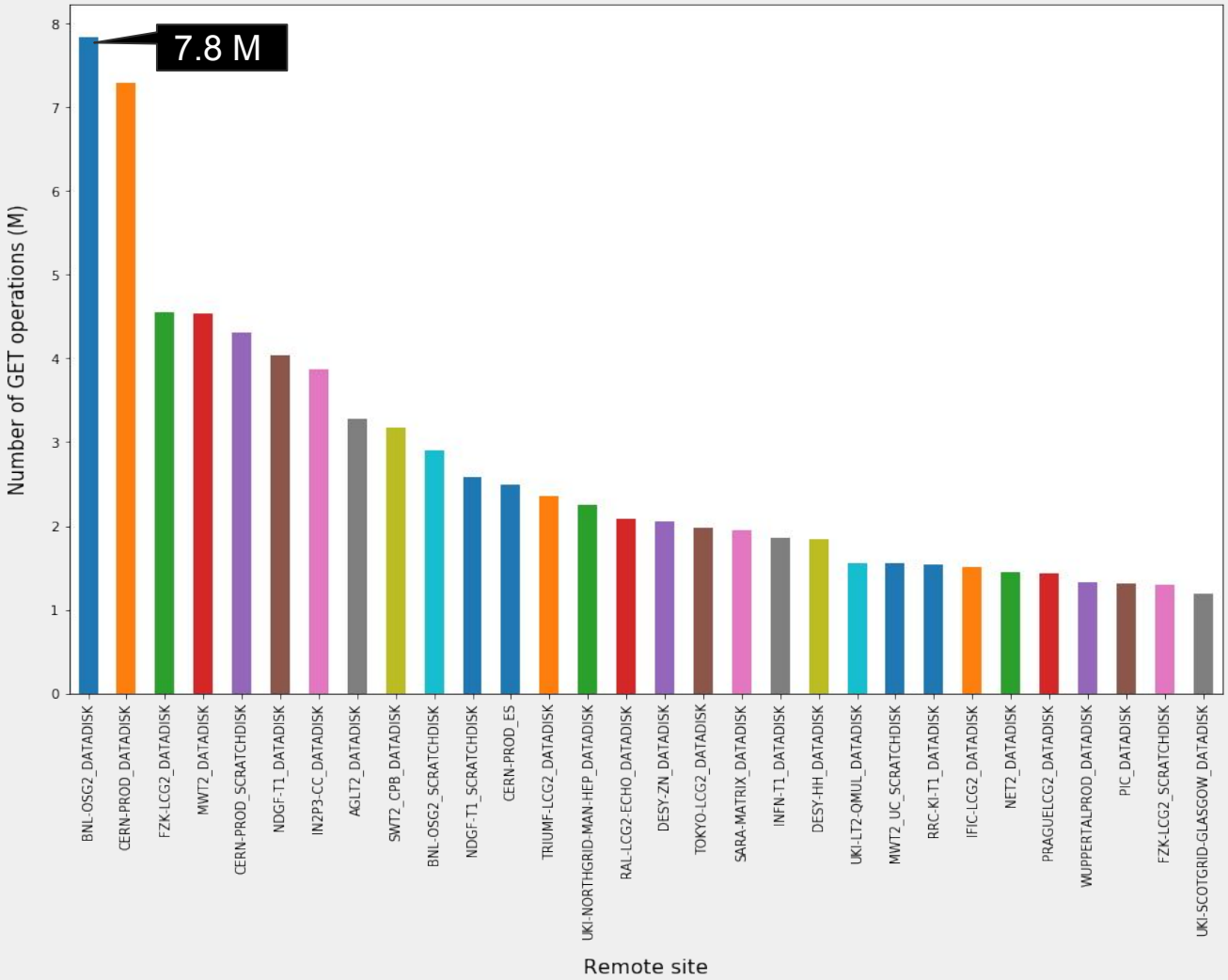
Quantitative Analysis: data movement re-enactment and cache simulation

Simulate a cache working for a particular site to estimate the relationship between hit rate and cache size

Understand the distribution of the number of requests (GET operations) and the amount of data accessed by the different sites that appear on the Rucio files

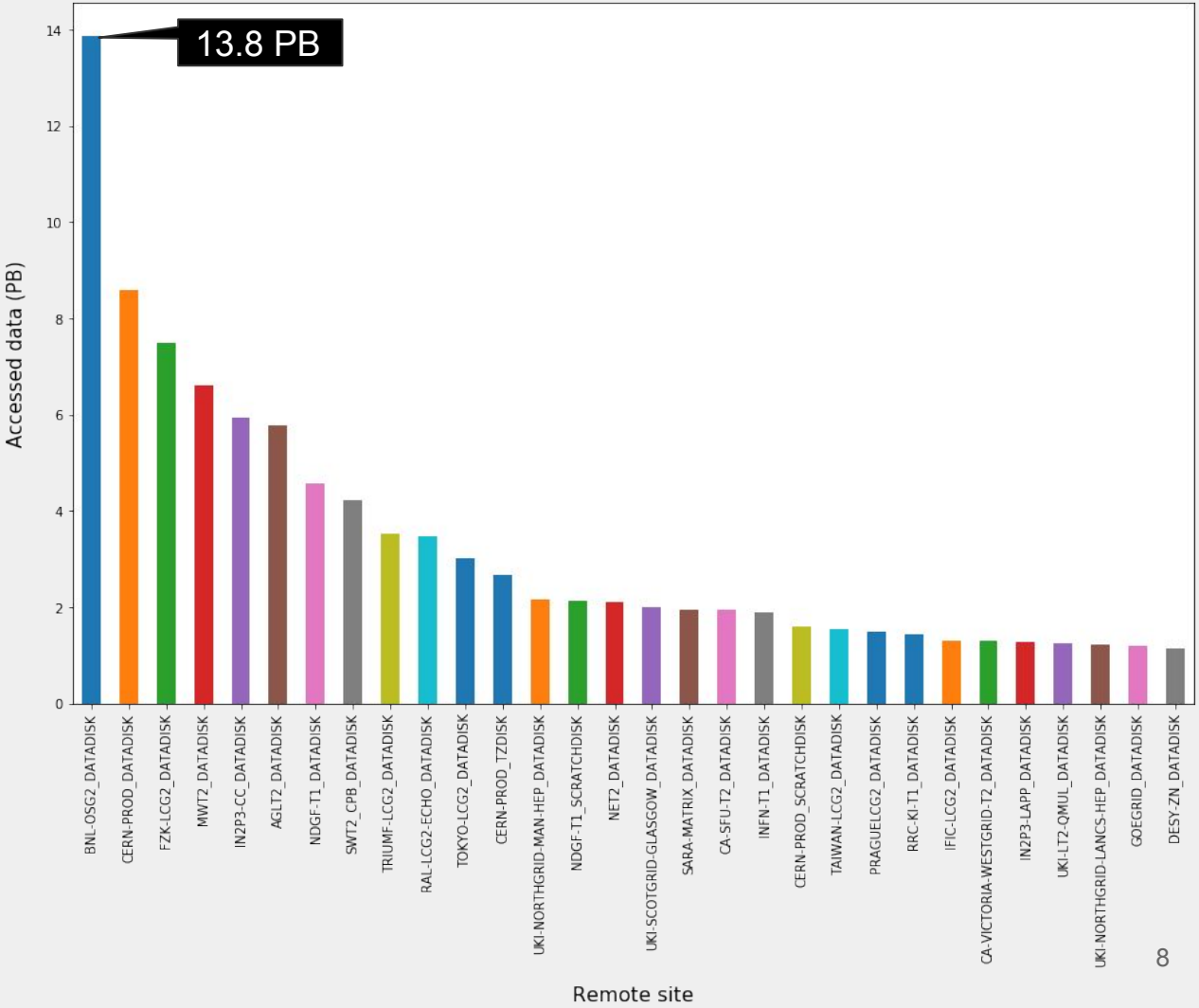
Rucio Data Management collects information of the operations made by the different sites in the grid

Top 30 sites with the largest number of GET operations between 31-05-2018 and 25-06-2018



- 26 days worth of Rucio files for a single experiment.
- 262 GB of data processed.
- 306 remote sites identified.

Top 30 sites with the most data accessed between 31-05-2018 and 25-06-2018



- 26 days worth of Rucio files for a single experiment.
- 262 GB of data processed.
- 306 remote sites identified.

PRAGUELCG2

Tier 2 Site

Total Online Storage: 6.3 PB

Physical CPUs: 7,028

Logical CPUs: 31,992

Number of operations: 1.4 M

Total data accessed: 1.5 PB

SIMULATION

XCache replacement algorithm:

Purges files last accessed furthest in the past.

Rucio files indicate:

Requests made for a particular file, its size and the time it was requested.

Simulation program:

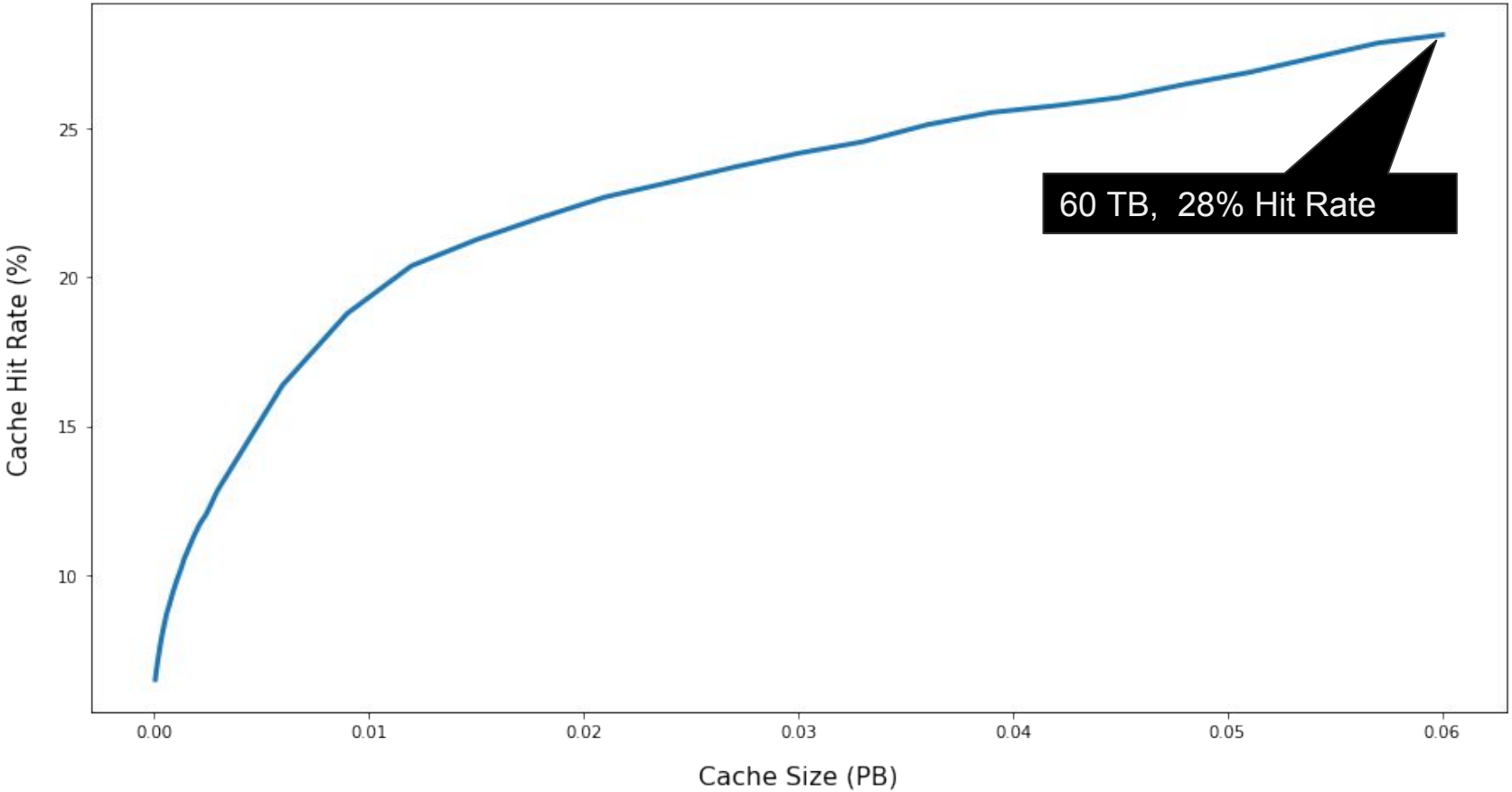
Works on a fixed size for one iteration over all the relevant Rucio data.

Executes a similar replacement algorithm as XCache.

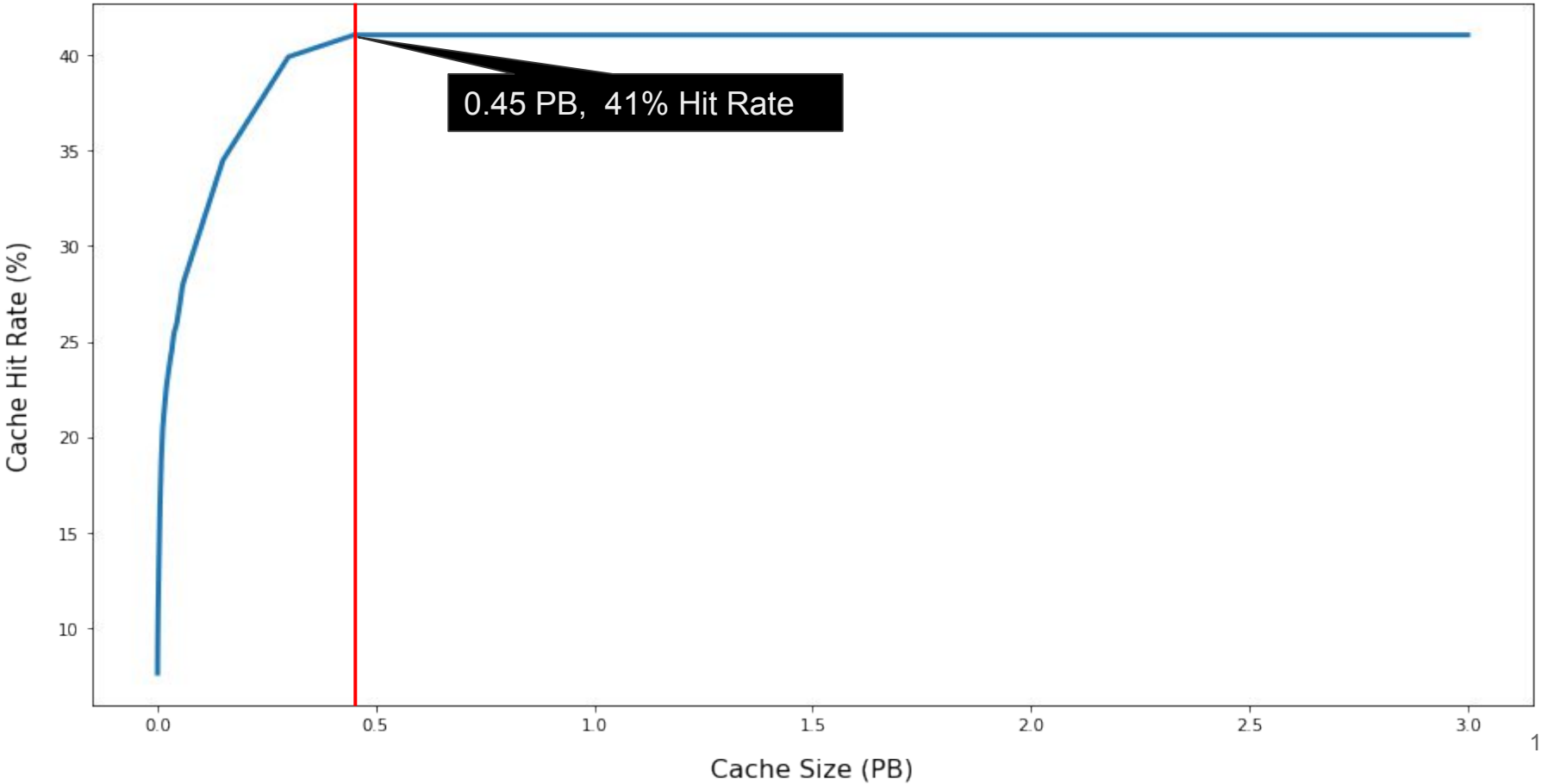
Counts hits and misses.

Hit rate = hits / (hits + misses).

Hit Rate vs Cache Size for PRAGUELCG2 (1 / 2)



Hit Rate vs Cache Size for PRAGUELCG2 (2 / 2)



CONCLUSIONS

A small amount of storage (7% of PRAGUELCG2s capacity) covers 41% of accesses.

Steep rise on the first 450 TB.

Increasing the size above 450 TB doesn't improve the hit rate.

FUTURE WORK

Use a larger data sample (longer time interval).

Compare different sites to identify patterns.

Find a model for the caching behavior.

THANK YOU