Predictive Modeling of Effective Peak Throughput for Regional In-Network Caching

Sunmin (Annie) Eom
Princeton University
Princeton, New Jersey, USA
ae8786@princeton.edu

Alex Sim (advisor)
Lawrence Berkeley National Laboratory
Berkeley, California, USA
asim@lbl.gov

Kesheng Wu (advisor)
Lawrence Berkeley National Laboratory
Berkeley, California, USA
kwu@lbl.gov

1 Introduction

Large-scale scientific experiments, particularly in high-energy physics (HEP), generate massive volumes of data repeatedly accessed worldwide. To reduce redundant long-distance transfers, institutions employ regional data caching, also known as in-network caching, which shares data efficiently within a region, lowering access latency and improving application performance [3, 5].

This work characterizes three years of access records from the Southern California Petabyte-Scale (SoCal) Cache, which supports HEP experiments such as the Large Hadron Collider (LHC). With the High-Luminosity LHC expecting a tenfold increase in network use by 2029, effective network engineering is increasingly critical [6].

We analyze historical daily peak throughputs, or data transfer rates, for cache misses and hits, examining the bursty maximum and the more stable 99th and 95th percentiles. We then train Long Short-Term Memory (LSTM) networks [2], well-suited to sequential data and proven effective for throughput prediction [4], to forecast future peaks and inform bandwidth allocation.

2 Background and Data

The SoCal Cache consists of 28 XCache nodes [1] at Caltech, UCSD, and ESnet Sunnyvale. 15.6 M accesses were recorded over 1,188 days from July 2022 to September 2025. Each access is a cache miss if the file is fetched from the remote source, causing wide-area network (WAN) traffic, or a hit if served locally from the cache.

Preprocessing filters out 16.8% of records whose operation time is zero, negative, over two days, or too long for the volume of data transferred. Of the remaining 13.0 M records, transferring 9.7 PB, 8% are misses and 92% hits, both by count and volume (Tables 1, 2).

Table 1: Summary of data access counts from SoCal Cache

	Total Accesses	Cache Misses	Cache Hits
Total	12,953,244	1,016,241 (7.8%)	11,937,003 (92.2%)
Daily Avg	10,903	855 (7.8%)	10,048 (92.2%)

Table 2: Summary of data access sizes (TB) from SoCal Cache

	Total Accesses	Cache Misses	Cache Hits
Total	9,731.0	799.2 (8.2%)	8,931.8 (91.8%)
Daily Avg	8.2	0.7 (8.2%)	7.5 (91.8%)

3 Methods

We compute per-minute miss and hit throughput by proportionally allocating each transfer's data size across the minutes it spans, then summing and dividing by 60 seconds. We then compute each day's maximum, 99th percentile, and 95th percentile.

Forecasting these values is challenging: demand is user-driven, and predicting daily peaks rather than averages compounds the variance. For each metric we build LSTM models in PyTorch for horizons of 1, 7, 14, and 30 days and tune hyperparameters. Each model takes eight daily features over an input window—access counts and sizes for cache misses, cache hits, and combined, plus the peak throughputs themselves—and outputs daily peak miss and hit throughput across the horizon. The models are trained on the first 80% of the data and tested on the rest.

Models are evaluated by test-set root mean square error (RMSE) and relative error (RE), the ratio of RMSE to the full-period standard deviation, and the best model per metric and horizon minimizes average miss and hit RE. Though no single architecture is clearly best, longer horizons favor two layers, higher learning rates, and longer input windows. The best 1-day-ahead and 7-day-ahead models use input windows ranging from 7 to 30 days.

4 Results

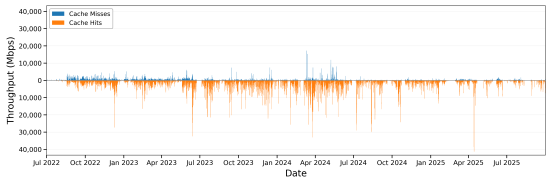


Figure 1: Per-minute throughput of cache misses (blue) and hits (orange) in Mbps

Per-minute throughput is bursty (Figure 1), with sharp spikes pushing daily maximum throughput to roughly 20 Gbps for misses and 40 Gbps for hits; the 99th and 95th percentiles reduce these to roughly 5 Gbps for misses and 30 Gbps for hits (Figure 2).

For 1-day-ahead predictions, RE is 0.20–0.25 for misses and 0.62–0.71 for hits (Table 3), far smaller than the data's natural variation, and predictions follow the general trend despite missing extreme peaks (Figure 3). Longer horizons struggle more: RE for the best 7-day-ahead model rises 6–20% from its 1st-day to 7th-day predictions (Table 3), and 7th-day predictions track significantly less day-to-day variation (Figures 4, 5). At 14 and 30 days, RE rises to 0.28–0.31 for misses and 0.73–0.99 for hits, and neither captures demand

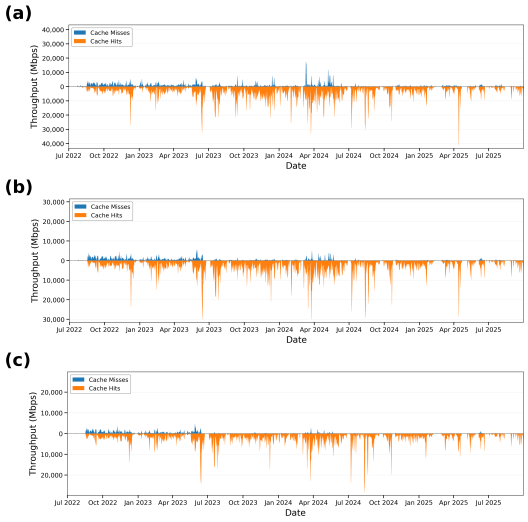


Figure 2: Daily (a) maximum, (b) 99th-percentile (p99), and (c) 95th-percentile (p95) per-minute throughput for cache misses (blue) and hits (orange)

Table 3: Test-set RMSE and RE for the best 1- and 7-day-ahead models, evaluated on day-1 and day-7 predictions

Metric	Target	1-Day Model, Day 1 Pred.	7-Day Model, Day 1 Pred.	7-Day Model, Day 7 Pred.	Std. Dev.
Max	Miss	297.1 (0.246)	312.3 (0.259)	368.1 (0.305)	1207.4
	Hit	3192.0 (0.705)	3383.8 (0.747)	4004.0 (0.884)	4527.3
p99	Miss	146.5 (0.227)	155.0 (0.241)	173.6 (0.270)	643.9
	Hit	2530.6 (0.672)	2554.7 (0.678)	3070.6 (0.815)	3768.0
p95	Miss	89.7 (0.200)	103.6 (0.231)	109.7 (0.245)	448.7
	Hit	1795.9 (0.615)	1733.5 (0.594)	2070.0 (0.709)	2917.8

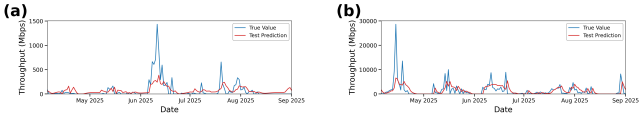


Figure 3: Daily p99 throughput day 1 test-set prediction of (a) cache misses, (b) cache hits, from the 1-day-ahead model

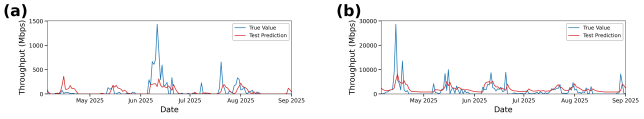


Figure 4: Daily p99 throughput day 1 test-set prediction of (a) cache misses, (b) cache hits, from the 7-day-ahead model

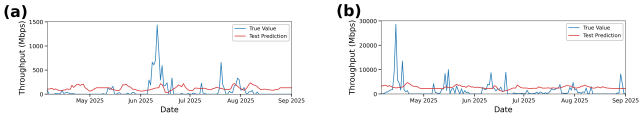


Figure 5: Daily p99 throughput day 7 test-set prediction of (a) cache misses, (b) cache hits, from the 7-day-ahead model

closely enough to be usable in practice. Across all horizons, the best model's RE decreases from maximum to 99th to 95th percentile (Table 3), confirming smoother metrics are more predictable.

Baseline comparisons provide a complementary measure of performance. Forecasts beat persistence (previous day's value) and weekly (previous week's values) baselines at nearly every horizon; against the harder input-window average, models win at 1 and 7 days for all metrics, but at longer horizons only for some misses. At the same prediction distance, however, models trained for different horizons perform comparably (Table 3; Figures 3, 4; 14- and 30-day results consistent), so longer-horizon training does not itself degrade performance substantially at a given prediction distance.

5 Conclusion

In this study, we examine how cache miss and hit throughputs in the SoCal Cache behave over a three-year period and use LSTMs to predict future peak throughputs. Cache hits dominate traffic, confirming the SoCal Cache absorbs most of it locally rather than over the WAN. Per-minute and daily maximum throughput are bursty, while the 99th and 95th percentiles are smoother and better suited to prediction and planning.

1-day-ahead forecasts are robust, following the general shape of demand at low relative error, but longer horizons grow progressively less reliable, and from 14 days onward baselines outperform the models. Forecasts of each peak metric offer distinct guidance: the predicted maximum covers the day's peak but leaves most capacity unused, while 99th or 95th percentile covers nearly all minutes at far less bandwidth. Future work will involve developing stronger long-horizon models and hourly models for finer granularity.

Acknowledgments

This work was supported by the Office of Advanced Scientific Computing Research and Office of Workforce Development for Teachers and Scientists, Office of Science, of the U.S. Dept. of Energy under Contract No. DE-AC02-05CH11231.

References

[1] Edgar Fajardo, Derek Weitzel, Mats Rynge, Marian Zvada, John Hicks, Mat Selmeci, Brian Lin, Pascal Paschos, Brian Bockelman, Andrew Hanushevsky, Frank Würthwein, and Igor Sfiligoi. 2020. Creating a content delivery network for general science on the internet backbone using XCaches. *EPJ Web of Conferences* 245 (2020), 04041. doi:10.1051/epjconf/202024504041

[2] Alex Sherstinsky. 2020. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena* 404 (2020), 132306. doi:10.1016/j.physd.2019.132306

[3] Alex Sim, Erica Wang, Ronak Monga, Kesheng Wu, Justas Balcas, Brendan White, Chin Guok, Inder Monga, Diego Davila, Frank Würthwein, and Harvey Newman. 2025. Comparing cache utilization trends for regional data caches. *EPJ Web of Conferences* 337 (2025), 01341. doi:10.1051/epjconf/202533701341

[4] Caitlin Sim, Kesheng Wu, Alex Sim, Inder Monga, Chin Guok, Frank Würthwein, Diego Davila, Harvey Newman, and Justas Balcas. 2023. Effectiveness and predictability of in-network storage cache for scientific workflows. In *2023 International Conference on Computing, Networking and Communication (ICNC)*. IEEE, Honolulu, HI, USA, 226–230. doi:10.1109/ICNC57223.2023.10074058

[5] Veronica Yakubovich, Ivy Sim, Phillip Kim, Nick Monozon, Jinny Chung, Alex Sim, and Kesheng Wu. 2026. Time-series forecasting for network utilization in large-scale scientific workflows. In *2026 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, Maui, HI, USA, 1–5. doi:10.1109/ICNC68183.2026.11417004

[6] Jason Zurawski, Dale Carder, Eric Colby, Eli Dart, Carol Hawk, Kenneth Miller, Abid Patwa, Kate Robinson, and Andrew Wiedlea. 2023. *High energy physics network requirements review: two-year update.* Technical Report LBNL–2001605, OSTI ID:2405935. Energy Sciences Network. doi:10.2172/2405935