Analyzing scientific data-sharing patterns for in-network data caching

Elizabeth Copps
Middlebury College
14 Old Chapel Road
Middlebury, VT 05753
ecopps@middlebury.edu

ABSTRACT

The volume of data generated by new scientific projects is exponentially increasing, driving up network traffic and data delivery requirements. In-network regional data caching systems are being used to minimize the negative effects of the increased data volume by storing files locally. They have been shown to save network traffic volume, reduce data latency, and improve overall application performance. In this work, an in-network XCache system is analyzed to determine how much traffic volume is saved over time. The researchers also studied the system to predict how new caches can be distributed to further decrease traffic volume and increase data availability. The cache system was shown to reduce network traffic by an average factor of 2.91.

1 INTRODUCTION

A large portion of the popular datasets from new scientific projects and simulations has been found to be transferred multiple times to the same user or to users in the same network [3]. In-network regional caching systems allow networks to fetch files from the source and store them locally for use by all users in the region, significantly reducing the number of redundant transfers. This has the potential to greatly increase CPU efficiency and application performance [2].

We have studied a regional cache that is particularly attractive from a research standpoint because it spans from Sunnyvale, California to San Diego, roughly 500 miles. Regional caches of this size can span across entire countries in Europe and between major cities in the US, allowing data to be shared by many institutions. Previous experiments have shown that the average read time for jobs at Caltech was five times higher before installing an Xcache node [1]. Researchers at Lawrence Berkeley National Laboratory and UCSD determined that the overall network demand was reduced by a factor of three once an XCache node was installed at the Energy Sciences Network (ESnet) [3].

In this study, we have analyzed data from the regional cache in Southern California with the goal of determining how an in-network caching system can be distributed to decrease traffic volume and increase data availability. We have determined how user requests are spread across the regional cache and the effect of adding or removing a single XCache installation. This research will influence how regional cache installations will be implemented as network demands continue to grow from science projects and simulations.

2 DATASET

This study used data from June through August of 2020 on the National Energy Research Scientific Computing Center's Cori supercomputer. Each time a user made a request, it was sent to a redirector and transferred to one of the XCache nodes in the regional cache. The regional cache that we studied consists of 14 XCache nodes in three different locations with varying capacities (Table 1). The two nodes at Caltech are significant for having seven times the disk capacity of the nodes located at UCSD, allowing us to determine whether larger nodes are beneficial to the regional cache.

Table 1

	UCSD	Caltech	ESnet
Number of Nodes	11	2	1
Node names	Cache 00 - Cache 10	Xrd1, Xrd2	ESnet
Disk Capacity	24 TB Each	180 TB Each	40 TB

The researchers focused on several key metrics of the data, including request time, data file size, the cache where the request was fulfilled, and whether the file was already in the cache or needed to be transferred. To study node downtimes, each time that a node had zero accesses for a period of time over one hour was recorded. On each remaining node, the researchers compared the average hourly data for an appropriate period of time before and after the node had gone down. The individual downtimes were divided into groups based on how many nodes in the cache had initially been down.

3 RESULTS

Table 2

	Accesses	Data transfer size (GB)	Shared data size (GB)	Share Percentage
June 2020	1,804,697	532,037.7	818,956.9	60.62%
Jul. 2020	1,426,585	354,452.8	764,351.3	68.32%
Aug. 2020	995,324	249,583.5	586,188.8	70.14%
Total	4,226,606	1,136,074.0	2,169,497.0	65.63%

The data in Table 2 show that the percentage of the total data size that is shared increases over the summer. This indicates that more files are being stored in the regional cache as time goes on, leading to fewer files that must be transferred from the data source. Around 2.17 PB of data was shared over the study period, and 1.14 PB was transferred, leading to an average share percentage of 65.63%. The number of daily accesses varied from day to day, with a slight downward trend over the study period (Figure 1). Although two of the other caches have seven times the disk capacity of other nodes, they only take 1.5 times more accesses than the other nodes at their peaks in mid June.

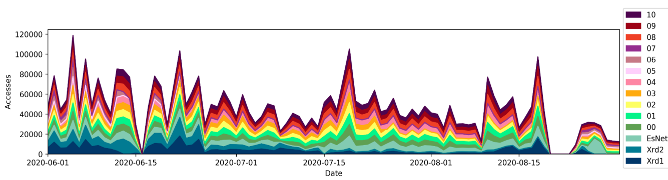

Figure 1: Daily total access count

The daily traffic reduction rate is a measure of how much network traffic the regional cache is saving by not having to transfer data from the source. It is calculated using the following equation:

$$\text{Network traffic reduction rate} = \frac{(total\ share\ size + total\ transfer\ size)}{total\ transfer\ size}$$

The traffic reduction rate was 2.91 over the full study period. The traffic reduction rates on individual nodes followed similar patterns to the total rate, shown in Figure 2.

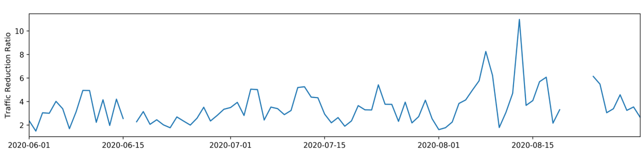

Figure 2: Daily traffic reduction rate

On average, the number of transfers increases after a downtime, while the number of shares decreases. The average hourly share count decreased by 264.59 shares per hour over all of the downtimes studied, and the average hourly transfer count increased by 67.81 transfers per hour. The average hourly share size decreased by 149 GB per hour, while the average hourly transfer size increased by 6.76 GB per hour. When a node went down, the other nodes evenly split the load that had been on that node.

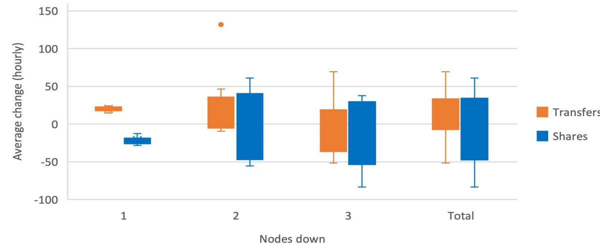

Figure 3: Box plots of hourly changes after a node downtime; 'Nodes down' refers to the number of nodes that were down initially

4 CONCLUSIONS

The proportion of data volume that is shared rather than transferred increases as more files are cached over time, increasing the traffic reduction rate. This can be expected to stop when the cached data volume reaches the regional cache's total capacity. With a higher number of nodes, more files can be cached, so there is a higher traffic reduction rate.

As data volume and network bandwidth demand continue to grow, adding additional nodes to regional caches can be expected to maintain network efficiency and data availability for new science projects and simulations. By adding a single node to a regional cache, the access load on other nodes will decrease proportionally according to how many nodes are already in the network. Adding a node with a larger disk capacity does not necessarily take a proportionally larger load of accesses from the network, but it does allow more files to be cached. Over time, this results in a higher traffic reduction rate as a larger number of requested files are already in the regional cache.

ACKNOWLEDGMENTS

Thank you to my mentors, Alexander Sim and John Wu, for their guidance and feedback throughout this project. This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internship (SULI) program.

REFERENCES

[1] Fajardo, E., Matevz, T., Balcas, J., Tadel, A., Würthwein, F., Davila, D., Jonathan, G., and Sfiligoi, I, "Moving the California distributed CMS XCache from bare metal into containers using Kubernetes", EPJ Web Conf. 245 04042 (2020), DOI: 10.1051/epjconf/202024504042

[2] Fajardo, E., Weitzel, D., Rynge, M., Zvada, M., Hicks, J., Selmeci, M., Lin, B., Paschos, P., Bockelman, B., Hanushevsky, A., Würthwein, F., and Sfiligoi, I., "Creating a content delivery network for general science on the internet backbone using XCaches", EPJ Web Conf. 245 04041 (2020), DOI: 10.1051/epjconf/202024504041.

[3] Sim, A., Zhang, K., Copps, E., Wu, J., Guok, C., Monga, I., Würthwein, F., Hernandez, E., and Davila, D., "Exploring in-network data caching, ESnet-US CMS collaboration study", WLCG Grid Deployment Board meeting, 2021, https://indico.cern.ch/event/876773/