



Leveraging user access patterns and advanced cyberinfrastructure to accelerate data delivery from shared-use scientific observatories

Yubo Qin^{a,*}, Ivan Rodero^a, Anthony Simonet^a, Charles Meertens^b, Daniel Reiner^b, James Riley^b, Manish Parashar^{a,c}

^a Rutgers Discovery Informatics Institute, Rutgers University, NJ, USA

^b UNAVCO, Boulder, CO, USA

^c Scientific Computing Imaging Institute, University of Utah, Salt Lake City, UT, USA

ARTICLE INFO

Article history:

Received 7 February 2020
Received in revised form 19 December 2020
Accepted 9 March 2021
Available online 25 March 2021

Keywords:

Cyberinfrastructure
Virtual Data Collaboratory
Distributed data sharing
Distributed facilities
Data pre-fetching
Observatories

ABSTRACT

With the growing number and increasing availability of shared-use instruments and observatories, observational data is becoming an essential part of application workflows and contributor to scientific discoveries in a range of disciplines. However, the corresponding growth in the number of users accessing these facilities coupled with the expansion in the scale and variety of the data, is making it challenging for these facilities to ensure their data can be accessed, integrated, and analyzed in a timely manner, and is resulting significant demands on their cyberinfrastructure (CI).

In this paper, we present the design of a push-based data delivery framework that leverages emerging in-network capabilities, along with data pre-fetching techniques based on a hybrid data management model. Specifically, we analyze data access traces for two large-scale observatories, Ocean Observatories Initiative (OOI) and Geodetic Facility for the Advancement of Geoscience (GAGE), to identify typical user access patterns and to develop a model that can be used for data pre-fetching. Furthermore, we evaluate our data pre-fetching model and the proposed framework using a simulation of the Virtual Data Collaboratory (VDC) platform that provides in-network data staging and processing capabilities. The results demonstrate that the ability of the framework to significantly improve data delivery performance and reduce network traffic at the observatories' facilities.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Almost all of science is transitioning from “data-poor” to “data-rich” in the 21st century, and large-scale observatories and shared-use instruments are playing essential roles in catalyzing scientific discoveries. With the increasing availability of shared-use instruments and observatories, observational data is becoming an important part of application workflows and contributor to scientific discoveries in a range of disciplines, and has enabled key scientific discoveries [1,2].

However, the corresponding growth in the number of users accessing these facilities coupled with the expansion in the scale and variety of the data, is making it challenging for these facilities to ensure their data can be accessed, integrated and analyzed in a timely manner, and is resulting in significant demands on their cyberinfrastructure (CI). At the same time, applications that use data from these observatories expect robust, highly available and

performant data services with low access latency [3–5]. Consequently, it is essential to address the associated data management and delivery challenges.

Recent studies have addressed these challenges from different perspectives. Several large-scale facility projects [3,6–10] co-design their facilities, cyberinfrastructure, and applications to support highly-customized and high-performance end-to-end scientific workflows. Although this supports the integration of processes from data generation to final discovery for its targeted workflows, developing such a specialized cyberinfrastructure is time-consuming and expensive, and the resulting solutions are typically inflexible.

Existing research has also explored reducing the volume of data transferred using methods such as data streaming [11], in-transit processing [12], edge processing [13], continuum computing [14], and proxy caching [15–18]. However, these solutions require appropriate capabilities at the edge and/or in the network and a co-design of user applications, limiting their broad adoption.

Recent CI projects [19–21] equipped with high-speed networks aim to address the data access challenge. For example, the Virtual Data Collaboratory (VDC) leverages the science DMZ

* Corresponding author.

E-mail address: yubo.qin@rutgers.edu (Y. Qin).

network model [22–24] to construct a dedicated regional network to enable high-throughput data transfers among users, data sources, and computing resources. However, considering that observational data is being generated at rates greater than increases in network bandwidth, the advances in networks alone cannot fundamentally resolve this challenge.

Furthermore, users access observational data in multiple ways with different performance expectations and access patterns. Users can manually download data from the observatory website, run scripts that automatically query data, or subscribe to real-time data streams. For example, automated script-based downloads are typically a part of application workflows [4,9], and while they often required stricter performance guarantees, their query and access patterns are more predictable.

We believe that an integration of knowledge of users' data access behaviors with the CI's data-delivery mechanisms can address data access challenges, improving data delivery performance as well as overall CI efficiency. In this research, we study user access and data usage patterns by analyzing the access traces of two NSF-funded large-scale observatories, the Ocean Observatories Initiative (OOI) operated by Rutgers and the Geodetic Facility for the Advancement of Geoscience (GAGE) operated by UNAVCO. At the same, we experiment with the VDC platform, a Science DMZ-based data CIs, and explore how its Data Transfer Nodes (DTNs) can be exploited to cache and pre-fetch data. Based on this study and its findings [25], we develop a set of optimization strategies that include a distributed cache layer, hybrid pre-fetching model (history-based and data mining-based), data placement strategy, and streaming mechanism. Furthermore, we combine these strategies and propose a push-based data delivery framework that can accelerate data delivery performance for large-scale shared-use observatories. The proposed framework is designed to run on DTNs within the VDC platform between the end-users and the data sources, allowing users to discover and access data through their local DTN, and the actual transfer of data is handled by the network of DTNs that are part of the framework using the optimization strategies developed in this work. From an end-users' perspective, data access is typically local as data is often pre-fetched and cached at the local DTN.

We have evaluated this model and framework using a simulated VDC platform and access traces from OOI and GAGE. Specifically, we measure latency and throughput to quantify the impact on data access performance of end-users under different network conditions, request traffic, and cache configurations. Moreover, we compare our hybrid pre-fetching model with state of the art pre-fetching models [26,27] for spatial-temporal data. The results show that our framework, combined with the proposed optimization strategies, can significantly improve data delivery performance and reduce the load on the observatory's CI. The main contributions of this work are as follows:

- Analysis of user access and data usage patterns for two representative large-scale shared-use observatories.
- Development of a hybrid data management model designed to support pre-fetching mechanisms.
- Use of DTN's within a Science DMZ to support data pre-fetching and caching based on user data access patterns.
- Design and prototyping of a push-based data delivery framework along with optimization strategies that leverage user data access patterns for data pre-fetching and caching.
- Experimentally evaluate the push-based data delivery framework using a simulated VDC platform and access traces from OOI and GAGE.

The remainder of this paper is organized as follows. Section 2 presents motivations for the presented work. Section 3 analyzes

the user data access patterns using OOI and GAGE access traces. Section 4 describes the design push-based data delivery framework along with optimization strategies for data pre-fetching and caching. Section 5 presents a performance evaluation of the system. Section 6 concludes the paper.

2. Motivation and background

2.1. Observatory data access challenge

Large-scale, shared-use scientific facilities, such as shared instruments, observatories and experimental platforms, have become key enablers for scientific discoveries in a range of disciplines. For example, the Event Horizon Telescope (EHT) generated the first images of a black hole event horizon [2], and the Laser Interferometer Gravitational-Wave Observatory (LIGO) enabled gravitational waves to be detected for the first time [1]. As a result, ensuring efficient, pervasive and democratic access to data from such facilities is essential and critical to amplifying their scientific impact.

However, rapidly growing data volumes and rates and the lack of sufficiently CI capabilities can limit effective access, and prevent scientists from using the data in a timely manner. For example, the LIGO project generates terabytes of data every day during its "observing" mode [28], and the in-construction Square Kilometer Array (SKA) project is estimated to generate an exabyte of raw data every day though it will then be compressed to around 10 petabytes [29]. As noted at the 2019 NSF Workshops on Connecting Large Facilities and Cyberinfrastructure [4, 30], many large facilities still struggle to provide broad data transfer and access services, especially in remote environments. Furthermore, using the wide-area network (WAN) is not the most effective approach for transferring large data volumes. For example, Dart et al. [31] note that it took three months to transfer 56 terabytes of climate model output data from the distributed Coupled Model Intercomparison Project (CMIP5) archive to the National Energy Research Supercomputing Center (NERSC). As most facilities do not provide co-located general-use computing and storage resources for end-users, users and application workflows typically have to retrieve the data over the network to local or national resources (such as those provided by Extreme Science and Engineering Discovery Environment (XSEDE)). Therefore, addressing data access challenges for large facilities is critical for both scientific facilities and their users.

2.2. Cyberinfrastructure and science DMZs

Cyberinfrastructure (CI) aims to knit together the end-users, data sources, software services, and computational resources using high-performance networks [32], with the overarching goal of enabling advanced science and engineering application. As the scale and complexity of applications increase, general-purpose networks quickly become insufficient to support their requirements, which can include managing and transporting of terabyte- or petabyte-scale dataset [33].

The Science Demilitarized Zone (Science DMZ) network model [34] is developed to optimize the transfer of large-scale scientific data. It achieves this goal by building a high-bandwidth network (e.g., 100 Gbps) among campuses, using dedicated Data Transfer Nodes (DTNs), by-passing traditional campus firewalls, and exploiting other techniques, as discussed in [22,23], to support high-throughput disk to disk data transfer [24].

Several recent-developed CI use the Science DMZ, such as the Pacific Research Platform (PRP) [19] and the Virtual Data Collaboration (VDC) [20]. In this work, we leverage VDC, which architecture is presented in Fig. 1. VDC constructs a high-bandwidth

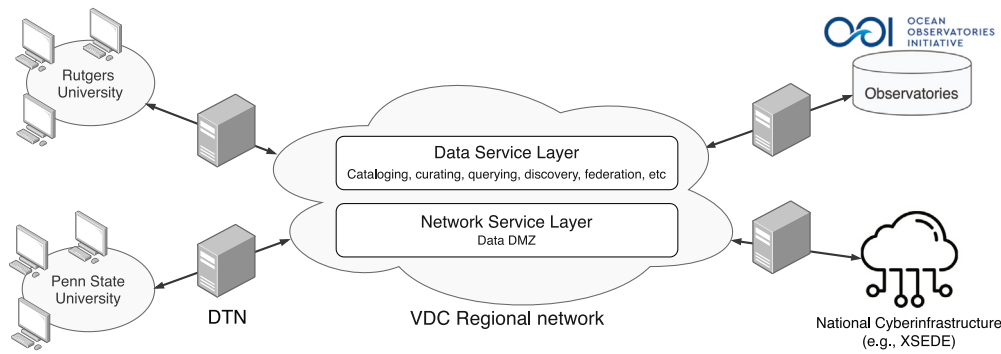


Fig. 1. Overall architecture of the Virtual Data Collaboratory (VDC). VDC is a data DMZ-based regional network. Users can connect to VDC using Data Transfer Nodes (DTN). The Data Services Layer support data discovery, access and processing across a federated environment that integrates multiple data sources and CI.

regional network connection. Its network services layer adopts the Science DMZ model to allow users on different campuses, data sources, and computational resources to connect to the VDC network using data transfer nodes (DTN). Its data service layer provides services, such as cataloging, curating, and querying, which enables users to discover data across observatories that are part for the VDC and integrate this data into application workflows running on remote computational resources.

DTNs within the VDC architecture are used to temporarily host data as moves across the network. Physically, the DTN is a purpose-built Linux server with significant storage and non-trivial computing power that is configured with the data transfer services such as Globus' gridFTP [35–37], and optimized for receiving WAN transfers at high speed [38,39]. For example, a DTN is capable for executing computation-intensive tasks such as machine-learning as proposed by the Chase-CI project [21]. In this work, we propose to deploy a cache/pre-fetching layer on the DTN to accelerate the data access performance for data from observatories. We also propose to leverage DTNs for data fusion and in-transit data processing.

2.3. Proxy caching and pre-fetching

Proxy caching is a well-known strategy for improving data access performance for users in geographically distributed scenarios. It achieves this goal by keeping content that is likely used in the near future at the locations close to its users [15,16]. Web data pre-fetching is a mechanism for predicting and fetching data prior to future requests [40–42]. Caching, combined with the pre-fetching, is an effective strategy for improving the user data access experience.

The proxy caching mechanism is typically implemented at three levels based on the cache location: client level, proxy level, and server level [15]. The proxy level caching is widely used in industry to keep the content close to users but not occupy users' local storage space. The Content Delivery Networks (CDN) [17] is one implementation of proxy level caching.

The role of the DTN within the Science DMZ network is equivalent to a proxy level cache within a web system. Therefore, by deploying a cache layer onto the DTNs in VDC essentially constructs a proxy server network that functions as the CDN to accelerate user data access performance [43] but without adding any additional hardware.

The cache replacement algorithm is crucial for web caching due to cache storage limitations [15,16,44]. Wong et al. categorize commonly used web cache replacement algorithms into: (1) Recency-based (e.g., Least-Recently-Used [45]); (2) Frequency-based (e.g., Least-Frequently-Used [46]); (3) Size-based, which evicts the largest object first [47]; and (4) Function-based, which

evicts objects according to their utility value [48]. The recency-based algorithm is the most commonly used for web caching because the web data has timeliness properties. Moreover, by leverages machine learning techniques, efforts such as Ali et al. [49, 50] further improve the Least-Recently-Used algorithm's performance.

Pre-fetching mechanisms enhance the caching performance by fetching data prior to the users' requests [40–42]. Pre-fetching algorithms are usually categorized as content-based and history-based. Content-based methods predict a user's future requests by analyzing data content. Xu et al. [51] propose to predict a user's future requests based on semantic preferences of previously retrieved Web documents. The history-based method conducts predictions based on a user's history records, and is the most commonly used approach for web pre-fetching. For example, existing studies [26,52] leverage data mining techniques to learn user access patterns using mining the correlations from a user's access records. Other work [27,53] has used Markov-based approaches to predict future requests by matching the user's current access sequence with history access sequences.

However, as existing literature [54] indicates, the factors determining which data objects will be re-accessed can vary significantly for different situations. Understanding user access patterns and particular environment characteristics are important for designing effective cache replacement and data pre-fetching mechanisms.

In order to exploit the locality of DTNs and the benefit of caching and pre-fetching, we propose to design a push-based data delivery framework along with a set of optimization strategies, including a hybrid pre-fetching model. To maximize the effectiveness of caching and pre-fetching we start by analyzing the access traces from real-world observatory and learn user data access patterns. We then develop a customized caching and pre-fetching mechanism that leverages this knowledge.

3. A study of observatory data access and usage patterns

In this section, we analyze traces from the NSF Ocean Observatories Initiative (OOI) and the Geodetic Facility for the Advancement of Geoscience (GAGE) with the goal of understanding how users access and utilize the data from these observatories.

OOI [3,55,56] is a networked ocean research observatory with arrays of instrumented water column moorings and buoys, profilers, gliders, and autonomous underwater vehicles (AUVs) distributed across different open ocean and coastal regions. OOI provides ocean scientists, educators, and the public the means to collect sustained, time-series datasets to enable the examination of complex, interlinked physical, chemical, biological, and geological processes operating throughout the coastal regions and open ocean.

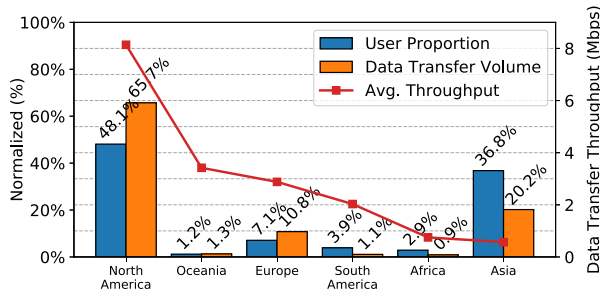


Fig. 2. Study of the impact of WAN performance on GAGE users' data access from different continents. The plot compares the proportion of users, data transfer volume, and the associated average data transfer throughput for each continent.

GAGE [57] (operated by UNAVCO) is a non-profit university-governed consortium that facilitates geoscience research and education using geodesy. GAGE services include the operation of the Network of the Americas (NOTA), which integrates a number of existing networks, and engineering, instrumentation, and data services for user of terrestrial and satellite geodetic technologies.

The analyses presented below is conducted on two traces: a one-month (November 2018) OOI user access log that contains 17.9 million user requests, and a one-year (2018) GAGE log that contains 77.8 million user requests. Each entry in the logs includes a publicly available IP address and the request metadata (e.g., filename, access timestamp, etc.).

3.1. Data transfer volume and network conditions

We use the GAGE trace to study how the impact of network performance exacerbates on data access. Since GAGE has a global user community, we analyze access performance across different continents. First, we estimate the user locations by reverse tracking their public IP address. Then, we compute the user distribution, the associated data transfer volume, and the average data transfer throughput as shown in Fig. 2. It is worth noting that while GAGE does have users from Antarctica who operate GPS receivers and work with data from the McMurdo station, their access records appear to come from other continents.

We observe that available network speed directly impacts the volume of user data transfers. North America, Oceania, and Europe have the highest average network throughput, and there is a positive correlation between their user percentages and the associated data transfer volumes. In contrast, there is a negative correlation for the rest of the continents. In particular, for Asian users, although they account for 37% users, they only account for 20% of the data transfer volumes and also have the lowest network performance (about 0.568 Mbps). This implies that low network performance limits user data access from the observatory.

3.2. Classification of users and requests

Data from the observatories is spatial-temporal and contains geographic and timing information in its metadata. To query this data, users need to provide the name of the data object and an observation time range, i.e., the starting and ending timestamp.

Our analysis of the OOI and GAGE traces reveals two types of requests: *human requests*, which are interactive requests from humans browsing the data catalog and manually downloading data, and *program requests*, which are from workflows or scripts that routinely download data using automated APIs.

We distinguish between these types of requests based on their access patterns, which differ in terms of their access frequencies

Table 1

Percentage of Human Users (HU) and the Program Users (PU), and their data transfer volumes for OOI and GAGE, respectively.

	Number of users (%)		Data transfer volume (%)	
	HU	PU	HU	PU
OOI	86.7%	13.3%	9.9%	90.1%
GAGE	94.1%	5.9%	9.4%	90.6%

and the queried data sets. Human requests query a set of data within a short period and then disappear. In contrast, program requests regularly query the latest data at a consistent frequency. We classify the users corresponding to these request types as *human user* and *program user* respectively.

To identify the request types, we maintain a running time window (e.g., one week) and analyze (i) each user's request frequency and (ii) patterns of repetition in the requests within that time window. If a user requests the same set of data objects more than once per day and this pattern repeats every day during the time window, we classify this user's future requests for this set of data objects as program requests. All other requests are classified as human requests.

Since access frequencies and the queried datasets differ between human users and program users, they require different optimization strategies. As a result, it is important to understand the relative percentages and associated data access volumes for each user type.

3.3. A quantitative study of users and their data accesses

The user percentages and associated data transfer volumes for human users and program users for OOI and GAGE are presented in Table 1. More than 86.7% of users are human users for both the observatories making them the primary user type. However, they generated less than 9.9% of the data transfers, which implies that program users are the primary data consumers. As a result, we mainly focus our optimizations on requests by program users and address human users to the extent possible.

3.4. Analysis of program requests

We analyze program requests based on the request metadata, i.e., access time, data object, and time range, and identify three access patterns: *regular*, *overlapping* and *real-time* request. These three request types are illustrated in Fig. 3 using a sample from the OOI trace on 23 November 2018. The horizontal axis is the time of the access request, and the vertical axis is the time range requested (i.e., the starting and ending timestamp of the requested data). Each vertical blue bar in the plot is a request, and the length of the bar represents the requested time range.

Regular requests query new data since the last request without any overlap. For example, in Fig. 3(a), a user makes a request every hour, and each request queries the most recent one-hour of data.

Regular request represent the most common request type. Observatories typically receive new data and update their databases at regular intervals which is determined by the characteristics of the instruments, the CI design, and other factors such as the data acquisition mechanism (e.g., satellite, cruise recovery). As a result, users develop programs that download the most recently updated data at these regular intervals.

Real-time requests are high-frequency (e.g., once per minute) regular requests. This type of request is shown in Fig. 3(b), where a user queries the past one-minute data every minute.

Real-time or near real-time data monitoring is typically used to develop science gateways [11] that monitor event occurrences,

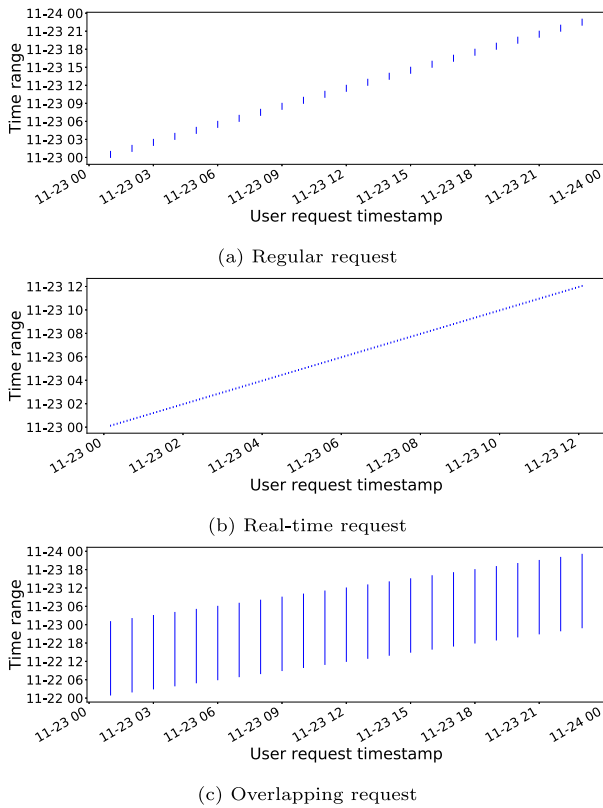


Fig. 3. Examples of the three types of user requests in the OOI trace on 23 November 2018. Fig. 3(a) illustrates *regular* requests where a user requests the past one-hour of data every hour. Fig. 3(b) illustrates *real-time* requests where a requests the past one-minute of data every minute (only a subset of the requests are plotted for visualization purposes). Fig. 3(c) illustrates *overlapping* requests where a user requests the past one-day of data every hour.

such as earthquake detection [5]. Since most observatories do not support subscribe-based data delivery, programs use high frequency *pull-based* requests to implement such monitoring. Real-time requests can result in large request traffic and significant load at the observatory. Furthermore, the data returned for each real-time request is typically small. As a result, these requests result in a large number of relatively small data transfers, which can potentially increase the system overheads and, in turn, reduce the quality of service provided by the observatory.

Overlapping requests are similar to the regular requests but with overlapping time-intervals across consecutive requests. These types of queries are illustrated in Fig. 3(c), where a user queries the past one-day data every hour, resulting in a 23-hour overlap between consecutive requests. While such queries are convenient from a user's perspective (e.g., the user is running an application every hour that needs the most recent 24 h of data), they can have a negative impact on the observatory as they result in redundant data transfers and unnecessary resource utilization.

Table 2 lists the percentage data transfer resulting from each of the three request types for OOI and GAGE. Note that for OOI overlapping requests dominate (i.e., 60.8%), while for the GAGE regular request dominate (i.e., 77.2%).

3.5. Analysis of overlapping requests

To quantitatively analyze the redundant data transfers resulting from overlapping requests, we divide the data transferred into *fresh data*, which is the percentage of data that is not part of previous request, and *duplicate data*, which is the percentage of redundant data.

Table 2

Data transfer volume for the three types of requests; Breakdown of the data transferred for overlapping request.

	Data transfer volume for the three request types			Data transfer volume for overlapping requests	
	Regular	Real-time	Overlapping	Fresh data	Duplicate data
OOI	13.8%	25.7%	60.8%	9.6%	90.4%
GAGE	77.2%	6.1%	17.2%	10.5%	89.6%

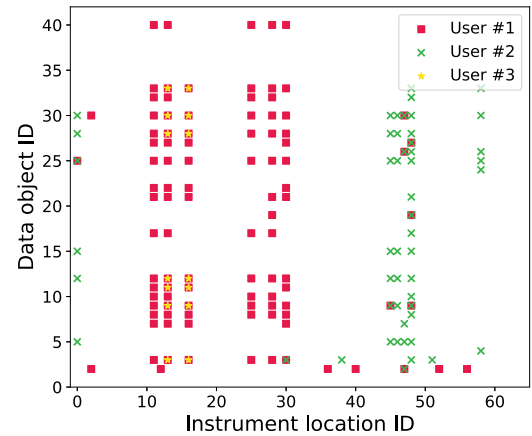


Fig. 4. Analysis of user request patterns. The x-axis represents instrument locations, and the y-axis represents requested data object IDs. The visible patterns suggest the existence of spatial correlation across the requests.

Table 2 shows that 90.4% of data transferred in response to overlapping requests is redundant data for OOI. Considering that 60.8% of OOI requests are overlapping, there is a significant amount of redundant data transfer observed in the OOI log.

Considering there is a large amount of redundant data transfer, adding a cache layer into the CI is an effective approach for reducing traffic and resource requirements at the observatory and improving performance. Furthermore, since a large percentage of the requests have predictable patterns, pre-fetching mechanisms can also be effective.

3.6. Analyzing correlations across requests

We observe a spatial-temporal correlation across human user requests in the OOI and GAGE traces indicating that human users tend to query data objects in a particular region as they navigate an observatory's data products. There are a couple of potential explanations for this. Scientist may be studying phenomena in a particular regions and as a result, query data from instruments and sensors in that region. Furthermore, observatory data portals typically use maps to help users navigate the observatory and explore desired instruments/sensors and associated data product and as a result, user request tend to spatially correlated. Consequently, it is possible to leverage this spatial-temporal correlation to predict human user requests.

Fig. 4 shows the queried data objects (using their instrument name and the instrument location) by three selected users from the one-day OOI trace used in Section 3.4. Since OOI deploys the same type of instrument at multiple locations [58], we use these two pieces of information to represent a specific data object. In the figure, we serialize the instrument name into an ID, which is represented in the y-axis. We then serialize the instrument locations, sort them by their proximity, and use their IDs to plot the x-axis. Thus, the dots in the same row represent the same type of instrument at different locations, while the dots in the same column represent different instruments in the same location.

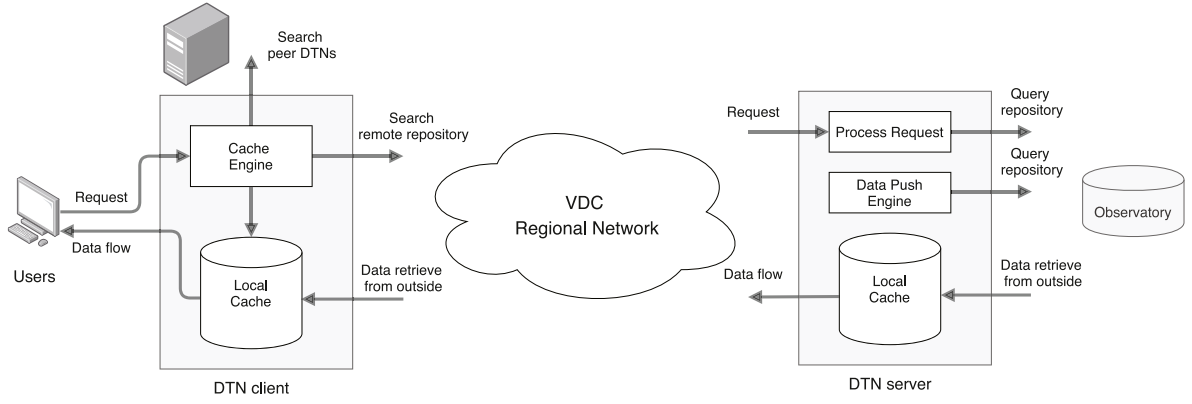


Fig. 5. Push-based data delivery framework deployed on top of the VDC Science DMZ-base architecture. The framework leverages the DTNs and their storage to implement a distributed and interconnected cache layer. The DTNs also host the data push engine. The framework client runs at the user's local DTNs and handles the user requests. The framework server runs at DTNs at the observatory and is responsible for managing the distributed data cache layer for data pre-fetching and streaming.

We observe a clear spatial correlation across requests in the plot, as users query multiple data objects of one region (vertical) and the same type of data object in nearby regions (horizontal). We also observe temporal correlation in our analysis. For example, in Figs. 3 where the consecutive blue bars, which represent the time interval of a request, have the same length. This implies that user (especially program users) tend to use a moving window while querying data. These observations confirm that we can potentially leverage this spatial-temporal correlation to predict user requests.

4. Design of a push-based data delivery framework

The proposed push-based data delivery framework exploits DTNs to cache queried data, manage the cached data, learn user access patterns, and proactively push data toward appropriate users.

The framework has a server-client architecture, as shown in Fig. 5. Clients run at user-side DTNs, pre-process user requests by searching for the requested data in caches at the local and peer DTNs, forwarding the request to the server if data is not found in these caches. The server runs at server-side DTNs that are local to the observatory. It handles requests arriving at the observatory, runs, executes the data pre-fetching and streaming mechanisms, and manages the placement of cached data.

Functionally, the architecture is composed of two main components: the cache layer and the data push engine. The cache layer spans the client and server side DTN to create a distributed interconnected cache network, and uses storage at the DTNs and caches data generated in response to user requests. The data push engine hosts the data pre-fetching and streaming mechanisms. It predicts user requests based on access patterns and can pre-fetch or stream data to users according based on these predictions.

4.1. Design of the hybrid data pre-fetching model

The goal of pre-fetching in this work is to predict user accesses based on observed access patterns presented in Section 3, allowing appropriate data objects to be pushed to the user. This is achieved as follows.

First, we exploit the spatial-temporal correlations to design an *association rule mining-based prediction model* for predicting the human requests as well as other requests that cannot be classified, based on the work from Xiong et al. [26]. Next, we design a *history-based prediction model* to handle program requests using the frequency of these requests. Finally, we develop a data streaming mechanism to process real-time requests.

We use the Autoregressive Integrated Moving Average (ARIMA) model, a widely used technique for time series prediction [59,60], to predict the timestamp of the user's next request. This combination of models can optimize the prediction accuracy and pre-fetching performance.

4.1.1. Representing user requests

We model user requests as a sequence $R_i = \langle r_1, r_2, \dots, r_n \rangle$. Each request tuple r_i includes the request timestamp ts , the name of the data object (or data stream) d , and the requested time range (i.e., the vertical bar in Fig. 3) tr . Furthermore, we represent the request of user j as the tuple R_j in Eq. (1), where $TS_n = \langle ts_1, ts_2, \dots, ts_n \rangle$, $D_n = \langle d_1, d_2, \dots, d_n \rangle$, $TR_n = \langle tr_1, tr_2, \dots, tr_n \rangle$, n is the number of data objects queried by the user j :

$$\begin{aligned} R_j &= \langle r_1, r_2, \dots, r_n \rangle \\ &= \langle (ts_1, d_1, tr_1), (ts_2, d_2, tr_2), \dots, (ts_n, d_n, tr_n) \rangle \\ &= TS_n, D_n, TR_n \end{aligned} \quad (1)$$

4.1.2. History-based prediction

The history-based data pre-fetching model is used when a request from a program user is identified. Specifically, during the learning period, we monitor a user's request sequence R . If its request pattern repeats more than a threshold number of times, then we identify this user as a program user and mark this series of requests as predictable, and the framework starts pre-fetching data for the user in anticipation of future requests. The learning period and threshold are empirically set to one week and 3 times, respectively in our experiments.

We use the ARIMA model to predict the timestamp of the user's next request, i.e., ts_{i+1} . Our implementation of the model is based on existing literature [59,60]. Training ARIMA requires the n most recent data points (i.e., a historical time interval). The higher the value of n is, the more accurate the prediction result is, but the training time is longer. In our experimental evaluation, we empirically set n to 60, which results in a training time of a few seconds and acceptable accuracy. Since the access intervals for *regular request's* are typically in the order of hours, this ARIMA prediction cost is acceptable.

Once the next timestamp of the user's next request, ts_{i+1} , is predicted, the framework can start pre-fetching relevant data. We use a pre-fetching offset to determine how far in advance this data should be pre-fetched. For instance, given offset to 0.8, the framework pre-fetches the data at a timestamp at $ts_i + 0.8 * (ts_{i+1} - ts_i)$, where ts_i is the timestamp of the user's last request.

The pre-fetching offset allows the user to control the pre-fetched timestamp and allows for time to transfer the data. Furthermore, ARIMA predictions are impacted by the considerable variance across consecutive access times and it can take several cycles for ARIMA to adapt, and an appropriate offset provides sufficient time to buffer the ARIMA prediction latency and achieve a higher pre-fetching success rate.

4.1.3. Association rule mining-based prediction

The association rule mining-based prediction model is used for request from human users and in cases where requests do not have identified patterns. This model uses the association rule mining Frequent Pattern Growth (FP-Growth) algorithm [61] to construct the prediction model similar to existing work [26,62–64]. FP-Growth will find an association among past requests to predict future requests for data objects, i.e., d_{i+1} . The time step for the next request is estimated based on the last two requests as follows: $ts_{i+1} = ts_i + (ts_i - ts_{i-1})$ and $tr_{i+1} = tr_i$.

The overall model is constructed as follows:

(a) FP-tree construction: The frequent-pattern tree (FP-tree) is a compact structure that stores quantitative information about frequent patterns in a database. The algorithm first scans the training dataset and counts the number of times a data point $d_i \in D_i$, $l_i \in L_i$ appears in the sequence of requests. This number is called *support*. Then, the algorithm finds the frequency 1-itemsets by comparing the itemset support with a predefined threshold. Finally, it re-scans the dataset and constructs the FP-tree.

(b) FP-Growth: Based on the FP-tree, which contains association rules between item sets and their corresponding *confidence*, the algorithm filters out association rules that have a confidence value lower than a predefined threshold, to form the complete set of frequent patterns.

Selecting appropriate threshold values for *support* and *confidence* is important to the model performance. In this work, we empirically set them to 30 and 0.5, respectively.

This method's prediction accuracy is lower than the history-based method due to the randomness in human user requests as observed in our experiments. Therefore, once the model generates a list of related data objects for a user sorted by their probability of repetition, the model only pre-fetches the first n data objects within the time range tr that are identical to the user's last request. The choice of n represents a trade-off between moving more data objects to increase the success rate and the resulting larger overheads. We empirically set n to 3 in our experiments.

4.2. Design of the data streaming mechanism

Real-time (or near real-time) data accesses to OOI and GAGE (we observed 25.70% and 6.10% real-time requests to OOI and GAGE respectively) are implemented as a sequence of high frequency requests, which results in significant traffic at the facilities. Like most such scientific observatories [3], these facilities do not support modern push-based models for delivering real-time streaming data. Furthermore, adding native support for data streaming can be a significant investment [30].

In this effort, we explore how these request for real-time (near-real-time) data can be supported more efficiently using in-network (using DTNs) resources, request prediction and data pre-fetching, and design a data streaming mechanism as part of the push-based data-delivery framework. This is done as follows: Once an incoming request is identified as a real-time request, the framework checks the observatory for data availability and pushes the most recent data back to the user for the duration of the request. Furthermore, the same request coming from multiple users can be combined into a single request to the observatory and redundant requests can be filtered out. Finally, this data can be cached at local DTNs to satisfy subsequent requests.

4.3. Design of the cache layer

The cache layer of the push-based data delivery framework designed to leverage storage at the DTNs to place data from the observatories as close to the user as possible. The design is consists of the choice of the cache eviction policy at each of the local DTNs, and an overall data placement strategy across the network of DTNs that is aware of their geographical distribution.

4.3.1. Choice of the cache eviction policy

The cache eviction policy at the local DTNs is based on the user access patterns. As observed in the analysis presented in Section 3, over 90.1% data transfer volume came from program requests that typically use a moving window to query data, and there is an overlap between consecutive requests.

Based on these observations, the least recently used (LRU) can be an effective cache eviction policy for the local DTN, which typically have a relatively large storage capacity (typical DTN configurations have tens to hundreds of GBs of memory and TBs of fast storage). Note that the LRU eviction policy can be replaced by other eviction policies as needed, and we compare the effectiveness of different policies in our experimental evaluation.

4.3.2. Data placement strategy

The goal of the data placement strategy is to place the data (and its replicas) at local DTNs that are close to potential users, and to keep the hot data, i.e., data with a high chance to be re-accessed, in the cache network as long as possible. The overall data placement strategy is composed of *virtual groups* and *local data hubs*.

A virtual group is a group of users that have common data interests and are geographically close to each other. We can place data objects of interest to a virtual group at a DTN that has the best network connective to the corresponding set of users.

In order to identify virtual groups with common data interests, we use K-Means to cluster past requests and then group users corresponding to the resulting clusters based on geographical proximity. We then map these user groups to appropriate DTNs which serve as the local data hub for the group. Note that this design assumes that users consistently access the observatory using institutional (e.g., university or research laboratory) resources and that the DTNs within the VDC align with these access locations.

The selection of a local data hub for a virtual group is based on three factors: network throughput, resource availability (e.g. storage) and user request frequency. As illustrated in Eq. (2), we select the DTN V_{dh} that maximizes the weighted sum of the three values as the local data hub within the virtual group. In this equation, P_{ij} represents the network throughput from DTN v_i to v_j and θ_p is the weight for the throughput; U_i represents the DTN device resource utilization (e.g., storage and CPU) with weight θ_u ; and F_i represents the request frequency of requests from the virtual group members that are connected to this DTN and has weight θ_f . We empirically set $\theta_p = 0.6$, $\theta_u = 0.2$, and $\theta_f = 0.2$.

$$V_{dh} = \max(\theta_p \sum_{j \neq i}^n P_{ij} + \theta_u U_i + \theta_f F_i), \quad 0 < i, j \leq n \quad (2)$$

Note that clustering virtual groups and the selection of local data hubs are performed periodically allowing the framework adapting to users changing their groups and interests. If the local data hub changes, the previous data hub keeps the data that was already cached, and new data is cached at the new data hub to minimize reconfiguration costs.

The virtual group concept is illustrated in Fig. 6 and is inspired by [18]. As shown in the figure, the virtual groups are overlaid on the physical infrastructure composed of DTNs that serve as the users' access points. Users belong to one or more virtual groups (e.g., *Virtual Group X* and *Virtual Group Y* in the figure) based on

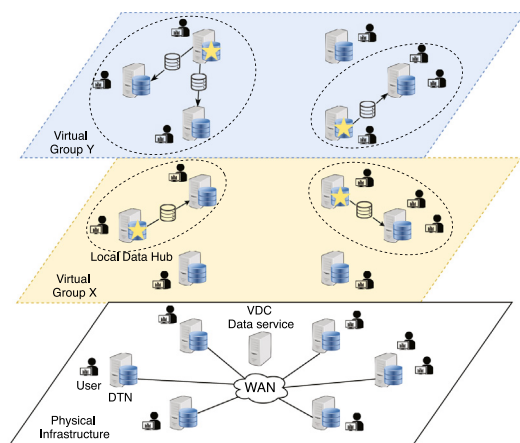


Fig. 6. Virtual groups and corresponding local data hub (i.e., the DTN with a star). The layered architecture is overlaid on top of the physical infrastructure. Users can belong to several virtual groups based on their data interests. These groups are split into sub-groups base on their geographic locations into sub-groups (represented by circles).

their data interests, and each virtual group is divided into sub groups based on their geographical location and mapped to DTNs in the physical infrastructure, which is the local data hub for the sub group serving as its access point and local cache.

4.4. Overall operation of the framework

The overall operation of the framework is illustrated in Fig. 5. When the client DTN receives a request, it first searches the cache layer starting with cache at the user's local DTN. Any portion of the requested data found at the local cache is returned, the peer DTNs are searched for any remaining requested data. If data is found cached at peer DTN, it compares data transfer costs from that DTN to transfer costs from the observatory and decided whether to transfer data from the DTN. Any remaining data is then requested from the server, and is also recorded by the pre-fetching engine allowing it to update its models. Finally, the data is pushed towards the user.

5. Experimental evaluation

5.1. Methodology and setup

We use a simulations along with OOI and GAGE request traces to evaluate the effectiveness and performance of the push-based data delivery framework presented in this paper under different conditions, and quantitatively compare it with two state-of-the-art pre-fetching models [26,27].

5.1.1. The simulator

We developed a simulator to emulate the VDC cyberinfrastructure as shown in Fig. 7. The simulated architecture is composed of seven geographically distributed DTNs interconnected via a wide-area network (WAN). We set DTN#1 as the VDC server that provides key data services including data cataloging, discovery, querying, etc.

The presented data delivery framework is deployed on the simulated VDC platform. The VDC server DTN is also server DTN for the framework and is the access point for the observatory. This DTN hosts the pre-fetching engine and manage data placement. The other DTNs (i.e., DTN#2 - #7) are the framework's client DTNs and collectively form the cache layer. The client DTNs can be configured to run different cache policies and use LRU by default.

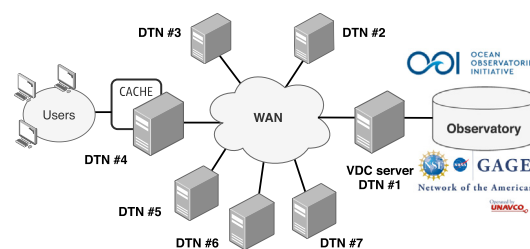


Fig. 7. Architecture of the simulator emulating the VDC cyberinfrastructure and used to implement push-based data delivery framework. DTN#1 as the VDC server and hosts the pre-fetching engine and manage data placement. The other DTNs (i.e., DTN#2 - #7) are client DTNs and collectively form the cache layer.

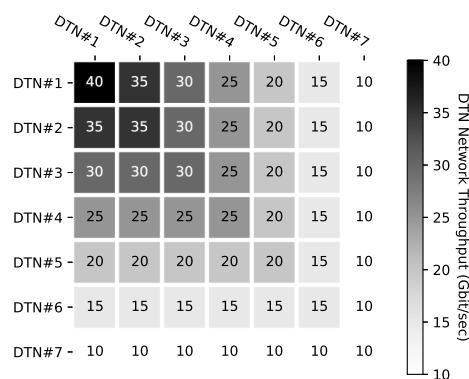


Fig. 8. Configuration of the interconnection bandwidths between DTNs used by the simulator.

In our experiment we also evaluate the Least-Frequently Used (LFU) policy.

To emulate GAGE’s average network throughput shown in Fig. 2, we limit the client DTNs’ bandwidth from 40Gbps to 10Gbps. The network bandwidth between DTNs is shown in Fig. 8. Since the bandwidth is not homogeneous across DTNs, we also evaluate the impact of the data placement strategy and the on the virtual group and local data hub schemes. The simulation assumes that users connect to their local DTNs at 100 Gbps.

The simulation uses a task queue at the VDC server to process the user requests, and ten service processes. User requests arriving at the server DTN are queued in the task queue and wait for the next available service process. Limiting the service processes to ten allows us to evaluate the impact of increasing request traffic on the observatory data service. When user requests arrive faster than the service processes can process, it results in longer queuing time and larger processing latency.

5.1.2. Evaluating the hybrid pre-fetching model

We evaluate the Hybrid Pre-fetching Model (referred to as **HPM**) used by the push-based data delivery framework by comparing it to two state-of-the-art spatial-temporal pre-fetching models [26,27]. The first reference model (**MD1**) is by Li et al. [27] and uses the Markov method to perform predictions. The authors connect geospatial data coordinates to convert the user access history into an “access path”. They observe that such paths follow Zipf’s law, and thus, they can predict user requests using a basic Markov model. In our evaluation, we add the geospatial coordinate information to the OOL and GAGE datasets and use this model.

The second reference model (**MD2**) is by Xiong et al. [26] and is a data mining-based method that uses a regional mesh and association rules to learn the spatial correlation, and use ARIMA to predict the temporal correlation. This model applies

the same prediction strategy to all user requests. In contrast, HPM distinguishes between request types (i.e., human or program requests) and uses the appropriate model for each request type.

5.1.3. Network conditions and request traffic variations

To evaluate data delivery performance under different network conditions, we consider three scenarios: *best*, *medium*, and *worst*. The *best* case maintains the original DTN bandwidth, as shown in Fig. 8. The *medium* and *worst* cases cap the bandwidth at each DTN at 50% and 1%, respectively.

Request traffic represents the number of requests that the observatory receives within a unit of time. Due to limited computational capacity and bandwidth, the observatory can only process a limited number of requests concurrently, and as a result, as the request traffic increases, the number of pending requests in the task queue also increases, and latency and data transfer times are correspondingly higher. This impacts the performance of the data delivery framework.

In our simulations, we consider three request traffic scenarios: *low*, *regular*, and *heavy*. We define the traffic corresponding to the one month traces from the OOI and GAGE as the regular request traffic, and emulate different request traffic scenarios relative to this traffic. To emulate the heavy request traffic, we compress one month of requests into a one-week time interval. It means that the observatory will receive four times the number requests within a unit time as compared to regular request traffic. Similarly, we emulate low request traffic by expanding one month of request to span a two-month time interval.

5.1.4. Simulator configurations

As noted previously, the simulator is configured to use the LRU cache eviction policy by default, and our evaluations compares LRU with LFU. We also compare the performance for different cache sizes. Specifically, due to the difference in data sizes, we evaluate the OOI trace with cache sizes of 128 GB, 256 GB, 512 GB, 1 TB, and 10 TB, and the GAGE trace with cache sizes of 32 GB, 64 GB, 128 GB, 256 GB, and 10 TB. Note that the 10 TB cache is large enough to cache the entire data and represents best-case performance.

Given that OOI and GAGE requested are distributed across the globe, in our simulations, we cluster request based on the continent they originate from. We use the client DTNs (DTN#2 - #7) to represent the six continents (excluding Antarctica) and distribute users across these DTNs according to their locations. By default, the simulator uses the *best* network configuration and *regular* request traffic configuration.

5.1.5. Evaluation metrics

In our experiments, we measure *latency* and *throughput* to quantify the data delivery performance to end-users. *Latency* is defined as the time between when the user submits a request to when the observatory starts processing it, and includes the time it spends in the observatory task queue. We compute *throughput*, by dividing the data size by the total data transfer time.

Furthermore, we use the *recall* metric to evaluate the performance of the pre-fetching mechanism, which is the percentage of the pre-fetched data that is accessed by the user. A higher *recall* value indicates that a smaller amount of pre-fetched data is wasted. Pre-fetched data may be wasteful if the prediction was incorrect, if the data is not delivered on time (i.e., before the user requests it), or if the cache configuration caused it to be evicted before it was accessed.

5.2. Experimental results

5.2.1. Data delivery performance

This experiment compares the framework's data delivery performance considering the cache configurations described in Section 5.1.4. Specifically, it compares the *throughput* and *latency* under the following conditions:

- *W/O Cache*: Represents the current observatory data delivery method that processes all requests by directly transferring data from the observatory to the user.
- *Cache Only*: Represents the baseline for our framework. It adds a cache layer using the DTNs, but does not use any optimization strategy.
- *HPM, MD1, MD2*: Represent our proposed framework that includes a cache layer and data placement optimization strategies, and is configured with three different pre-fetching models.

Our framework with pre-fetching achieves significant improvement in accelerating the data delivery performance, where the experiments' results are plotted in Figs. 9–12. For example, Fig. 9 shows that, in case of OOI, using the smallest cache size (128GB), HPM increases data transfer *throughput* by 2,689.8x, and reduces request *latency* by 34.8% as compared to the *No Cache* case. Only adding a cache layer (i.e., the *Cache Only* case), we observe a 739.6x improvement in *throughput*. This confirms our analysis presented in Section 3 that indicated significant overlap across request. It also validates the benefits of a cache layer as a simple yet effective method for improving the data delivery performance.

Using pre-fetching results in a 3.6x increase in *throughput* relative to the baseline. Moreover, HPM performs better than the reference models (MD1, MD2) in all cases. As seen in Figs. 9(c)–12(c), the pre-fetching *recall* values indicates that HPM achieves the best prediction accuracy. This is because there are over 90.1% program requests in the traces (Table 1) and the HPM's history-based prediction model can predict these requests with high accuracy. In contrast, the reference models (MD1 and MD2) treat all requests equally, which results in at least 66.7% wasted pre-fetching and redundant data movement. This validates the benefits of using user data access patterns to improve the data delivery performance.

Furthermore, MD2 has a higher *recall* value than MD1 for all cases. This indicates that the association rule-based prediction model performs better than the Markov-based method in predicting user data request for observatories. It also suggests that HPM's associated rule-based sub-prediction model performs well for non-program requests.

Figs. 9(b)–12(b) show that our framework decreases *latency* (e.g., 34.8% OOI, LRU, 128GB) as compared to the *No Cache* case. This is because the cache layer, and the pre-fetching and data streaming mechanisms reduce the number of data requests that are sent to the server DTN. Hence, it reduces the size of the task queue and allows each request to be processed faster. Specifically, Table 3 lists the normalized count of requests arriving at the server DTN for each configuration. HPM reduces the number of requests served by the observatory in all cases, especially as compared to the *Cache Only* case. This validates the benefits of using the data push, data pre-fetching and data streaming mechanisms for *real-time requests*, allowing users to mostly retrieve data from their local DTN.

The results show that the LRU cache eviction policy works better than the LFU eviction policy for small cache size configurations. LRU achieves 79.6% (OOI, 128GB) and 91.2% (GAGE, 32GB) higher *throughput* than LFU using 1TB and 256GB cache

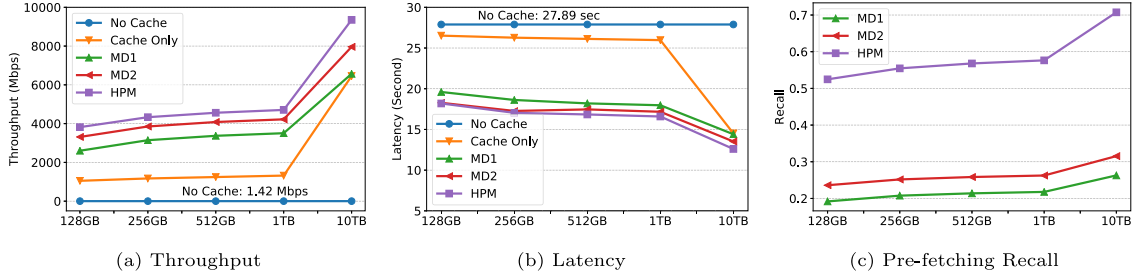


Fig. 9. OOI LRU cache performance.

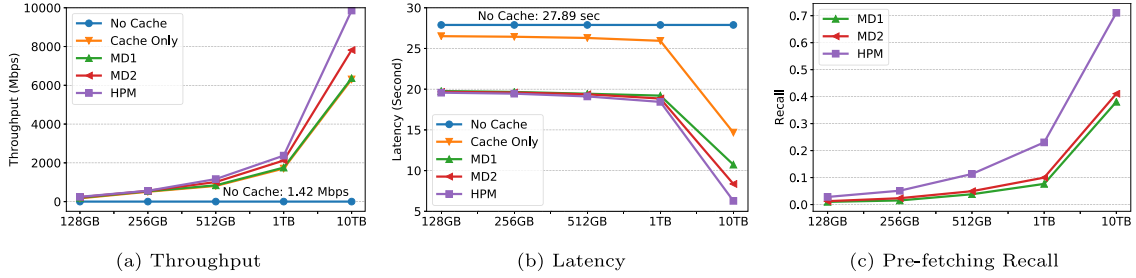


Fig. 10. OOI LFU cache performance.

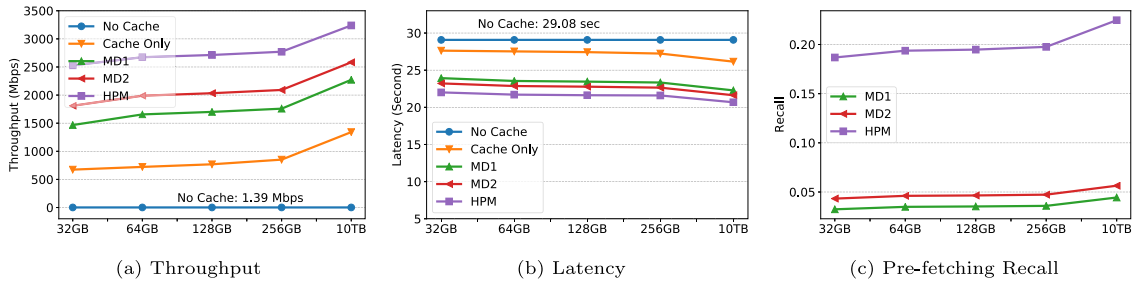


Fig. 11. GAGE LRU cache performance.

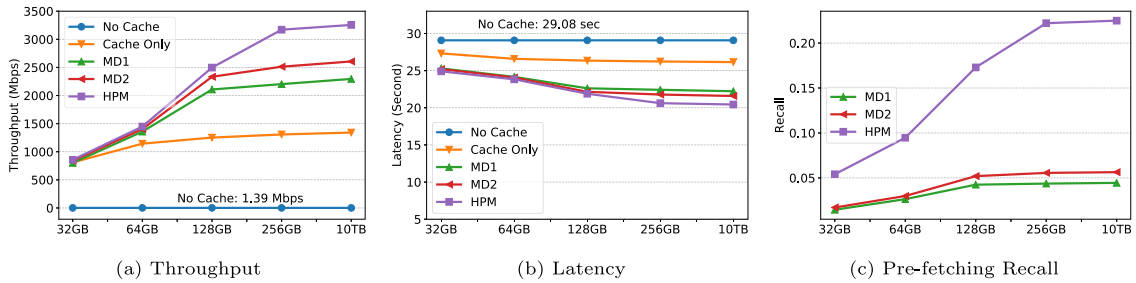


Fig. 12. GAGE LFU cache performance.

Table 3

Normalized number of user requests for the OOI and GAGE traces that have to be served by the observatory.

		No Cache	Cache Only	MD1	MD2	HPM
OOI	LRU	1.0000	0.5722	0.4653	0.4228	0.3928
	LFU	1.0000	0.9881	0.9874	0.9871	0.9849
GAGE	LRU	1.0000	0.9437	0.8707	0.8447	0.7912
	LFU	1.0000	0.9107	0.8941	0.8905	0.8893

sizes respectively. In contrast, the *throughput* curve rises steeply when using LFU and achieves higher throughput than LRU with large cache sizes (i.e., 10TB, OOI). This implies that the recency of

requests is more relevant than the frequency of request for observatories. Since DTN storage resources are limited and considering the typical scale of observatory data, performing well with a small cache size is important, and as a result, we suggest that using the LRU eviction policy is a better choice. We do not further evaluate advanced recency-based eviction models [49,50] in this paper and consider it as future work.

5.2.2. Analysis of the pre-fetching mechanism

The goal of these experiments is to understand how pre-fetching improves the local data reuse. Fig. 13 plots the average percentage of requests that are served from the local cache for the four strategies using the last experiment for the LRU cache

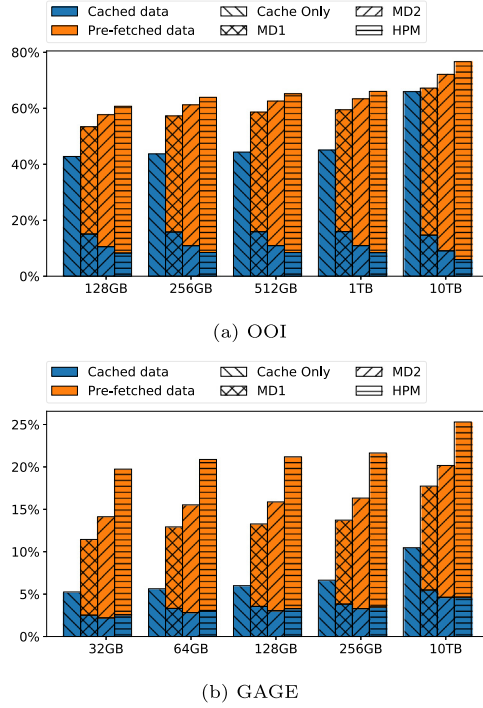


Fig. 13. Data movement from the local cache. These plots show the percentage of requests that are served by cached and pre-fetched data.

configurations. The color in the plots marks the data sources, where blue is the percentage of requests served using cached data, and the yellow is percentage of requests serviced using pre-fetched data. These plots illustrate that pre-fetching enables users to obtain more data from their local cache. For instance, using the smallest cache sizes (OOI 128GB, GAGE 32GB) and the *HPM* pre-fetching model, the percentage of local data access is 41.9% and 278.8% higher than the *Cache Only* case for OOI and GAGE, respectively.

As opposed to passively searching cached data, the pre-fetching mechanism proactively pushes data toward to user. It ensures that users can access more of their data locally regardless of whether the data is reused from the previous requests. For example, in Fig. 13b, GAGE has a smaller number of *overlapping requests*; however, the pre-fetching mechanism still enables a much higher percentage of local data accesses as compared to the baseline.

Furthermore, the pre-fetching mechanism can achieve near-optimal performance with a small cache size. The best performance is obtained for a 10TB cache size as the entire datasets for both traces fit into the cache. *HPM* with the smallest cache size can achieve 79.2% and 78.0% of this best performance, for OOI and GAGE, respectively.

Based on these experiments, we conclude that our push-based data delivery framework can enable users to access more data from their local DTN cache by proactively pushing data toward users.

5.2.3. Evaluation of the data placement strategy

In this experiment, we use *HPM* with LRU and the GAGE trace to evaluate the impact of the data placement strategy on the data delivery performance. Table 4 summarizes the results. The first row presents the percentage of cached data that has been optimized by the data placement strategy using different cache configurations. We observe that the data placement strategy is

Table 4

Impact of the data placement strategy (DP): Percentage of cached data optimized by the DP, throughput (Mbps) for retrieving the data from a peer DTN's cache, and the impact of the DP on total data transfer performance.

		32GB	64GB	128GB	256GB	Avg.
% Data opt. by DP		21.11%	17.13%	12.95%	10.91%	15.52%
Throughput (Mbps)	W/O DP	16968.15	16899.72	16763.17	16644.47	16818.88
	W/ DP	20291.43	20230.38	20095.91	19985.99	20150.93
Improv. %		19.59%	19.71%	19.88%	20.08%	19.81%
Tot. perf. improv. (Mbps)	W/O DP	2446.10	2603.19	2658.92	2723.77	2608.00
	W/ DP	2527.70	2673.68	2713.45	2771.17	2671.50
Tot. perf. improv. %		3.34%	2.71%	2.05%	1.74%	2.46%

more effective for a small cache size because it replicates hot data to the local data hub.

We also measure the average throughput for retrieving data from the peer DTN cache. The data placement strategy improves throughput by 19.81% on average. The performance increases slightly as the cache size gets larger because the local data hub can keep the data replica longer. Overall, the data placement strategy improves data delivery performance by 2.46% on average. We anticipate that our data placement strategy is likely to impact overall data delivery performance even more significantly as the scale of the DTN network grows, the network gets more heterogeneous and complex, and more the amount data retrieved from peer DTN caches increases.

5.2.4. Framework performance for different network condition and request traffic

The goal of this experiment is to evaluate the ability of the push-based data delivery framework presented in this paper to tolerate different network conditions and request traffic levels. As presented in Section 5.1.3, we create three network conditions, i.e., *best*, *medium* and *worst*, where *medium* and *worst* are 50% and 1% of the *best* case respective, and three levels of request traffic. Moreover, we use a LRU cache of size 1TB and 256GB for the OOI and GAGE traces respectively. We compare the achieved *throughput*, which is presented in Table 5.

The columns of the table compares performance of each strategy for different network conditions. The performance remains constant for *HPM*, *MD1* and *MD2* for the *best* and *medium* network conditions. However, for the *worst* network condition, the throughput drops down by 34.9% and 31.1% relative to the *best* case for OOI and GAGE, respectively. These results illustrate that the pre-fetching mechanism tolerates network bandwidth variations as it can exploit the potentially idle network resource to transfer data in advance. We can also see that the network conditions significantly impact the *No Cache* case as in this case data needs to be downloaded directly from the observatory. This indicates that the data delivery methods currently implemented by observatories are not very resilient to unfavorable network conditions. Conversely, the *Cache Only* strategy is not significantly affected by the network conditions because in this case, data is mostly retrieved from the local DTN's cache.

The rows of the table present the average data transfer throughput for each cache strategy and for the *low* to *high* request traffic levels. The request traffic impacts all the cases except *Cache Only* as there is limited concurrent processing in this case. Since our simulations use a fixed number (*ten*) of service processes, heavier request traffic implies longer queuing time. Moreover, given sufficient processing capability, multiple concurrent data transfers decrease the shared bandwidth available to each task. As a result, the data transfer time increases, which can result in the pre-fetched data arriving too late and being wasted. We conclude that controlling the request traffic at the observatory is important. The push-based data delivery framework presented in this paper addresses this by using caching and pre-fetching mechanisms to reduce the request traffic.

Table 5

Pre-fetching model performance comparison for different network conditions and request traffic levels. HPM is our hybrid pre-fetching model; DM1 and DM2 are the reference pre-fetching models.

(a) OOI throughput for different network conditions and request traffic levels using LRU.															
Network	Low Request Traffic					Regular Request Traffic					Heavy Request Traffic				
	No Cache	Cache Only	MD1	MD2	HPM	No Cache	Cache Only	MD1	MD2	HPM	No Cache	Cache Only	MD1	MD2	HPM
Best	2.85	1326.97	3849.01	4881.04	5609.05	1.42	1322.24	3509.73	4221.31	4705.25	1.21	1311.97	2680.96	3201.53	3582.40
Medium	2.01	1327.27834	3858.24	4899.32	5615.12	1.11	1319.97	3506.01	4221.02	4707.16	0.92	1312.27	2677.13	3200.63	3582.07
Worst	1.01	1326.95858	2507.73	3185.79	3648.89	0.71	1318.08	2278.10	2743.75	3057.84	0.58	1311.96	1739.93	2080.23	2327.72

(b) GAGE throughput for different network conditions and request traffic levels using LRU.															
Network	Low Request Traffic					Regular Request Traffic					Heavy Request Traffic				
	No Cache	Cache Only	MD1	MD2	HPM	No Cache	Cache Only	MD1	MD2	HPM	No Cache	Cache Only	MD1	MD2	HPM
Best	1.52	851.84	1758.36	2091.16	2771.17	1.39	850.00	1501.11	1903.03	2488.08	1.31	849.02	1147.16	1348.49	1729.56
Medium	1.26	853.12	1760.41	2092.90	2771.26	1.19	851.99	1498.31	1904.62	2487.06	1.09	848.52	1146.77	1351.36	1730.24
Worst	0.69	853.45	1214.21	1443.12	1908.88	0.69	849.97	1035.21	1313.22	1715.93	0.60	849.50	789.01	927.91	1193.97

6. Conclusion

The paper addressed data access challenges for large-scale, share-use facilities such as instruments, observatories and experimental platforms with the goal of improving data access performance and overall user experience for these facilities. Specifically, in this work we analyze the data access and usage patterns for two large-scale facilities, OOI and GAGE and study how these patterns can be leveraged to predict future requests.

In this paper, we have presented a push-based data delivery framework that leverages user access patterns to design a hybrid data pre-fetching model along with optimized cache mechanisms aimed at accelerating data delivery performance for distributed observatories. Furthermore, we presented an implementation of this framework that leverages the VDC Science DMZ and uses in-network DTNs to host the data pre-fetching and caching services.

We evaluated our framework using a simulated VDC environment and OOI and GAGE user access logs. The experimental evaluation indicated a significant reduction in the amount of data transferred over the network for both, the OOI and GAGE observatories as compared to the current approach (60.7% and 19.7% improvements for OOI and GAGE respectively).

Potential future optimizations include replacing the ARIMA time-series prediction model with the portable Recurrent Neural Network (RNN) based predictor [65], which could further improve performance. Our current efforts include building a prototype implementation to be deployed in the Virtual Data Collaboratory.

CRedit authorship contribution statement

Yubo Qin: Writing - original draft, Investigation, Methodology, Software, Visualization, Validation, Conceptualization. **Ivan Rodero:** Resources, Investigation, Methodology, Conceptualization, Writing - review & editing, Project administration, Funding acquisition. **Anthony Simonet:** Investigation, Writing - review & editing. **Charles Meertens:** Resources. **Daniel Reiner:** Resources. **James Riley:** Resources. **Manish Parashar:** Resources, Investigation, Methodology, Conceptualization, Writing - review & editing, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research is supported in part by National Science Foundation (NSF), USA via grants numbers OCE 1745246, OAC 1835692, OAC 1826997, and OAC 1640834, and was conducted as part of

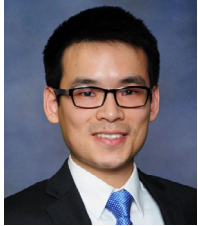
the Rutgers Discovery Informatics Institute (RDI²). The materials used are based in part on services provided by the GAGE Facility, operated by UNAVCO, Inc., with support from NSF and the National Aeronautics and Space Administration, USA under NSF Cooperative Agreement EAR-1724794 and NSF grant OAC 1835791. We thank the reviewers for their careful review, which helped improve the manuscript.

References

- [1] B.P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., Observation of gravitational waves from a binary black hole merger, *Phys. Rev. Lett.* 116 (6) (2016) 061102.
- [2] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett, D. Bintley, et al., First m87 event horizon telescope results. iv. imaging the central supermassive black hole, *Astrophys. J. Lett.* 875 (1) (2019) L4.
- [3] I. Rodero, M. Parashar, Data cyberinfrastructure for end-to-end science, *Comput. Sci. Eng.* 22 (05) (2020) 60–71.
- [4] E. Deelman, A. Mandal, V. Pascucci, S. Sons, J. Wyngaard, C. Vardeman, S. Petruzza, I. Baldin, L. Christopherson, R. Mitchell, et al., Cyberinfrastructure center of excellence pilot: Connecting large facilities cyberinfrastructure, in: 2019 15th International Conference on eScience (eScience), 2019, IEEE, 2019, pp. 449–457.
- [5] K. Fauvel, D. Balouek-Thomert, D. Melgar, P. Silva, A. Simonet, G. Antoniu, A. Costan, V. Masson, M. Parashar, I. Rodero, et al., A distributed multi-sensor machine learning approach to earthquake early warning, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, 2020, pp. 403–411.
- [6] P. Dewdney, P. Hall, R. Schillizzi, J. Lazio, The square kilometre array, *Proc. Inst. Electr. Electron. Eng.* IEEE 97 (8) (2009) 1482–1496.
- [7] A. Abramovici, W.E. Althouse, R.W. Drever, Y. Gürsel, S. Kawamura, F.J. Raab, D. Shoemaker, L. Sievers, R.E. Spero, K.S. Thorne, et al., Ligo: The laser interferometer gravitational-wave observatory, *Science* 256 (5055) (1992) 325–333.
- [8] T.U. Kampe, B.R. Johnson, M.A. Kuester, M. Keller, Neon: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure, *J. Appl. Remote. Sens.* 4 (1) (2010) 043510.
- [9] E. Deelman, K. Vahi, G. Juve, M. Rynge, S. Callaghan, P.J. Maechling, R. Mayani, W. Chen, R.F. Da Silva, M. Livny, et al., Pegasus, a workflow management system for science automation, *Future Gener. Comput. Syst.* 46 (2015) 17–35.
- [10] M. Albrecht, P. Donnelly, P. Bui, D. Thain, Makeflow: A portable abstraction for data intensive computing on clusters, clouds, and grids, in: Proceedings of the 1st ACM SIGMOD Workshop on Scalable Workflow Execution Engines and Technologies, 2012, pp. 1–13.
- [11] I. Rodero, Y. Qin, J. Valls, A. Simonet, J. Villalobos, M. Parashar, C. Youn, C. Wang, K. Thareja, P. Ruth, et al., Enabling data streaming-based science gateways through federated cyberinfrastructure, *Gateways* 2019, 2019.
- [12] A.R. Zamani, M. Zou, J. Diaz-Montes, I. Petri, O. Rana, A. Anjum, M. Parashar, Deadline constrained video analysis via in-transit computational environments, *IEEE Trans. Serv. Comput.* (2017).

- [13] E.G. Renart, D. Balouek-Thomert, M. Parashar, An edge-based framework for enabling data-driven pipelines for iot systems, in: 2013 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 2019, IEEE, 2019, pp. 885–894.
- [14] D. Balouek-Thomert, E.G. Renart, A.R. Zamani, A. Simonet, M. Parashar, Towards a computing continuum: Enabling edge-to-cloud integration for data-driven workflows, *Int. J. High. Perform. Comput. Appl.* 33 (6) (2019) 1159–1174.
- [15] W. Ali, S.M. Shamsuddin, A.S. Ismail, et al., A survey of web caching and prefetching, *Int. J. Advance. Soft Comput. Appl.* 3 (1) (2011) 18–44.
- [16] G. Zhang, Y. Li, T. Lin, Caching in information centric networking: A survey, *Comput. Netw.* 57 (16) (2013) 3128–3141.
- [17] J.D. Gagliardi, T.S. Munger, Content delivery network, *US Patent* 7, 962, 580 (Jun. 14 2011).
- [18] F. Jiang, C. Castillo, S. Ahalt, Cachalot: A network-aware, cooperative cache network for geo-distributed, data-intensive applications, in: NOMS 2018–2018 IEEE/IFIP Network Operations and Management Symposium, 2018, IEEE, 2018, pp. 1–9.
- [19] L. Smarr, C. Crittenden, T. DeFanti, J. Graham, D. Mishin, R. Moore, P. Papadopoulos, F. Würthwein, The pacific research platform: Making high-speed networking a reality for the scientist, in: Proceedings of the Practice and Experience on Advanced Research Computing, PEARC '18, 2018, pp. 29:1–29:8.
- [20] M. Parashar, V. Honavar, A. Simonet, I. Rodero, F. Ghahramani, G. Agnew, R. Jantz, The virtual data collaboratory, *Comput. Sci. Eng.* (2019).
- [21] I. Altintas, K. Marcus, I. Nealey, S.L. Sellars, J. Graham, D. Mishin, J. Polizzi, D. Crawl, T. DeFanti, L. Smarr, Workflow-driven distributed machine learning in chase-ci: A cognitive hardware and software ecosystem community infrastructure, 2019, arXiv preprint [arXiv:1903.06802](https://arxiv.org/abs/1903.06802).
- [22] D.R.C. Magri, T.C.M. de Brito Carvalho, F.F. Redigolo, M.A.T. Rojas, M.A.S. Junior, L.N. Ciuffo, G.N. Dias, A.S. de Moura, F. Vetter, Science dmz: Support for e-science in brazil, in: 2014 IEEE 10th International Conference on e-Science, Vol. 2, IEEE, 2014, pp. 75–78.
- [23] P. Calyam, A. Berryman, E. Saule, H. Subramoni, P. Schopis, G. Springer, U. Catalyurek, D.K. Panda, Wide-area overlay networking to manage science dmz accelerated flows, in: 2014 International Conference on Computing, Networking and Communications (ICNC), IEEE, 2014, pp. 269–275.
- [24] L. Farrell, Science DMZ: The fast path for science data, *Sci. Node.* (May 2016).
- [25] Y. Qin, A. Simonet, P.E. Davis, A. Nouri, Z. Wang, P. Manish, I. Rodero, Towards a smart internet-scale cache service for data intensive scientific applications, in: Proceedings of the 10th Workshop on Scientific Cloud Computing, ACM, 2019, pp. 11–18.
- [26] L. Xiong, Z. Xu, H. Wang, S. Jia, L. Zhu, Prefetching scheme for massive spatiotemporal data in a smart city, *Int. J. Distrib. Sens. Netw.* 12 (1) (2016) 4127358.
- [27] R. Li, R. Guo, Z. Xu, W. Feng, A prefetching model based on access popularity for geospatial data in a cluster-based caching system, *Int. J. Geogr. Inf. Sci.* 26 (10) (2012) 1831–1844.
- [28] LIGO: Laser Interferometer Gravitational-Wave Observatory, <https://www.ligo.caltech.edu/mit/>.
- [29] SKA: Square Kilometer Array <https://www.skatelescope.org/the-ska-project/>.
- [30] 2019 NSF Workshop on Connecting Large Facilities and Cyberinfrastructure, <https://facilitiesci.github.io/2019/>.
- [31] E. Dart, M.F. Wehner, et al., An assessment of data transfer performance for large-scale climate data analysis and recommendations for the data infrastructure for cmip6, 2017, arXiv preprint [arXiv:1709.09575](https://arxiv.org/abs/1709.09575).
- [32] Transforming Science Through Cyberinfrastructure, <https://www.nsf.gov/cise/oac/vision/blueprint-2019/Overview-Computational.pdf> (2019).
- [33] J. Crichigno, E. Bou-Harb, N. Ghani, A comprehensive tutorial on science dmz, *IEEE Commun. Surv. Tutor.* 21 (2) (2018) 2041–2078.
- [34] E. Dart, L. Rotman, B. Tierney, M. Hester, J. Zurawski, The science dmz: A network design pattern for data-intensive science, *Sci. Program.* 22 (2) (2014) 173–185.
- [35] K. Chard, S. Tuecke, I. Foster, Globus: Recent enhancements and future plans, in: Proceedings of the XSEDE16 Conference on Diversity, Big Data, and Science At Scale, 2016, pp. 1–8.
- [36] W. Allcock, J. Bresnahan, R. Kettimuthu, M. Link, The globus striped gridftp framework and server, in: SC'05: Proceedings of the 2005 ACM/IEEE Conference on Supercomputing, 2005, IEEE, 2005, pp. 54–54.
- [37] B. Radić, V. Kajić, E. Imamagić, Optimization of data transfer for grid using gridftp, *J. Comput. Inf. Technol.* 15 (4) (2007) 347–353.
- [38] Z. Liu, R. Kettimuthu, I. Foster, P.H. Beckman, Toward a smart data transfer node, *Future Gener. Comput. Syst.* 89 (2018) 10–18.
- [39] R. Kettimuthu, Z. Liu, D. Wheeler, I. Foster, K. Heitmann, F. Cappello, Transferring a petabyte in a day, *Future Gener. Comput. Syst.* 88 (2018) 191–198.
- [40] T.M. Kroeger, D.D. Long, J.C. Mogul, et al., Exploring the bounds of web latency reduction from caching and prefetching, in: USENIX Symposium on Internet Technologies and Systems, 1997, pp. 13–22.
- [41] G. Pallis, A. Vakali, J. Pokorny, A clustering-based prefetching scheme on a web cache environment, *Comput. Electr. Eng.* 34 (4) (2008) 309–323.
- [42] Y.-F. Huang, J.-M. Hsu, Mining web logs to improve hit ratios of prefetching and caching, *Knowl.-Based Syst.* 21 (1) (2008) 62–69.
- [43] K. Mokhtarian, H.-A. Jacobsen, Caching in video cdns: Building strong lines of defense, in: Proceedings of the ninth European conference on computer systems, 2014, pp. 1–13.
- [44] S. Podlipnig, L. Böszörményi, A survey of web cache replacement strategies, *ACM Comput. Surv.* 35 (4) (2003) 374–398.
- [45] A. Vakali, Lru-based algorithms for web cache replacement, in: International Conference on Electronic Commerce and Web Technologies, Springer, 2000, pp. 409–418.
- [46] L. Cherkasova, G. Ciardo, Role of aging frequency and size in web cache replacement policies, in: International Conference on High-Performance Computing and Networking, Springer, 2001, pp. 114–123.
- [47] S. Jin, A. Bestavros, Greedydual web caching algorithm: exploiting the two sources of temporal locality in web request streams, *Comput. Commun.* 24 (2) (2001) 174–183.
- [48] P. Cao, S. Irani, Cost-aware www proxy caching algorithms, in: Usenix symposium on internet technologies and systems, vol. 12, 1997, pp. 193–206.
- [49] W. Ali, S.M. Shamsuddin, A.S. Ismail, Intelligent web proxy caching approaches based on machine learning techniques, *Decis. Support Syst.* 53 (3) (2012) 565–579.
- [50] W. Ali, S. Sulaiman, N. Ahmad, Performance improvement of least-recently-used policy in web proxy cache replacement using supervised machine learning, *Int. J. Adv. Soft Comput. Appl.* 6(1) (2014).
- [51] C.-Z. Xu, T.I. Ibrahim, A keyword-based semantic prefetching approach in internet news services, *IEEE Trans. Knowl. Data Eng.* 16 (5) (2004) 601–611.
- [52] Q. Yang, H.H. Zhang, T. Li, Mining web logs for prediction models in www caching and prefetching, in: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, 2001, pp. 473–478.
- [53] A. Nanopoulos, D. Katsaros, Y. Manolopoulos, A data mining algorithm for generalized web prefetching, *IEEE Trans. knowl. Data Eng.* 15 (5) (2003) 1155–1169.
- [54] K.-Y. Wong, Web cache replacement policies: a pragmatic approach, *IEEE Netw.* 20 (1) (2006) 28–34.
- [55] L.M. Smith, J.A. Barth, D.S. Kelley, A. Plueddemann, I. Rodero, G.A. Ulses, M.F. Vardaro, R. Weller, The ocean observatories initiative, *Oceanography* 31 (1) (2018) 16–35.
- [56] I. Rodero, M. Parashar, Architecting the cyberinfrastructure for National Science Foundation Ocean Observatories Initiative (OOI), in: 7th International Workshop on Marine Technology: MARTECH, vol. 2016, 2016, pp. 99–101.
- [57] NSF Awards the Geodetic Facility for the Advancement of Geoscience (GAGE) to UNAVCO, <https://www.unavco.org/highlights/2018/award.html>.
- [58] OOI: Ocean Observatories Initiative, <https://oceanobservatories.org/>.
- [59] J. Contreras, R. Espinola, F.J. Nogales, A.J. Conejo, Arima models to predict next-day electricity prices, *IEEE Trans. Power Syst.* 18 (3) (2003) 1014–1020.
- [60] D.Ö. Faruk, A hybrid neural network and arima model for water quality time series prediction, *Eng. Appl. Artif. Intell.* 23 (4) (2010) 586–594.
- [61] J. Han, J. Pei, Y. Yin, Mining frequent patterns without candidate generation, in: ACM Sigmod Record, Vol. 29, ACM, 2000, pp. 1–12.
- [62] S. Pan, Y. Chong, Z. Xu, X. Tan, An enhanced active caching strategy for data-intensive computations in distributed gis, *J. Supercomput.* 73 (10) (2017) 4324–4346.

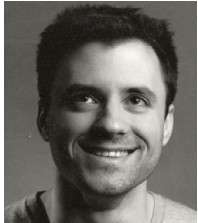
- [63] R. Li, W. Feng, H. Wu, Q. Huang, A replication strategy for a distributed high-speed caching system based on spatiotemporal access patterns of geospatial data, *Comput. Environ. Urban Syst.* 61 (2017) 163–171.
- [64] L. Xiong, L. Yang, Y. Tao, J. Xu, L. Zhao, Replication strategy for spatiotemporal data based on distributed caching system, *Sensors* 18 (1) (2018) 222.
- [65] S.-M. Tseng, B. Nicolae, G. Bosilca, E. Jeannot, A. Chandramowlishwaran, F. Cappello, Towards portable online prediction of network utilization using mpi-level monitoring, in: *European Conference on Parallel Processing*, Springer, 2019, pp. 47–60.



Yubo Qin received his M.S. degree, Ph.D. degree from the Rutgers University, and B.S. degree from the University of Electronic Science and Technology of China. His current research focuses on addressing data discovery and geo-distributed data sharing challenges on earth science. He is a student member of IEEE and ACM.



Ivan Rodero is Associate Research Professor at the Rutgers Discovery Informatics Institute at Rutgers University. His research focus on data-driven science and engineering, parallel and distributed computing and advanced cyberinfrastructure. He received his M.S. and Ph.D. degrees from the Technical University of Catalonia – Barcelona Tech. He has received various awards for his research and publications, including the IEEE TCSC Young Achievers in Scalable Computing Award. He is senior member of IEEE and ACM, and member of AAAS.



Anthony Simonet received a master's degree from the University of Bordeaux, and the Ph.D. degree from the École Normale Supérieure de Lyon, France, in 2015. He is a post-doctoral associate with the Rutgers Discovery Informatics Institute (RDI2) at Rutgers University in New Jersey. His research spans over multiple domains including distributed data management, models and middleware for hybrid distributed computing infrastructures and energy.



Charles Meertens is the director of Geodetic Data Services at UNAVCO. His research interests are in the broad area of geodesy, geophysics, hydrology, informatics, and volcanology. He holds a Ph.D. and M.Sc. degree from the University of Colorado Boulder and a B.A. from the University of California, Santa Barbara.



Daniel Reiner is a software engineer technical lead at UNAVCO. He received his B.S. from DeVry University.



James Riley is a software engineer, technical lead and senior web administrator at UNAVCO. He specializes in linked data, user-centered design, information systems and information architecture. He received his B.S. from Regis University.



Manish Parashar is the Distinguished Professor of Computer Science at Rutgers University and the founding director of the Rutgers Discovery Informatics Institute. He is currently on an IPA appointment at the National Science Foundation (NSF), where he is serving as office director for NSF Office of Advanced Cyberinfrastructure. His research interests are in the broad areas of parallel and distributed computing and computational and data-enabled science and engineering. He has received a number of awards for his research and leadership and is a Fellow of the AAAS, IEEE, and IEEE Computer Society and an Association for Computing Machinery Distinguished Scientist.