Modeling inter-DC bulk data transfer in optical networks with edge storage

Shengnan Yue1 · Xiao Lin2 · Weiqiang Sun1 · Weisheng Hu1

Received: 24 March 2022 / Accepted: 13 June 2022 / Published online: 20 July 2022 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
In communication networks that span large geographical regions, transfer of bulk data is often costly and hard to manage. One prominent reason for that is bandwidth contention among long-lasting flows, and between long-lasting and short-lived ones on inter-data-center links. Delaying the transfer of bulk data to off-peak hours by temporarily storing transfer at edge of the network can help reduce bandwidth contention during peak hours, and improve overall bandwidth utilization and network utility. Although extensive research has been conducted on its use, to date no model has been developed to quantify the performance of storage assisted bulk transfers across multiple congested links. This research gap has limited our understanding of the operation of storage assisted transfer under different network conditions. In this paper, we model bulk data transfer in inter-datacenter optical networks with edge storage, and derive closed-form equations for bulk data transfer on inter-DC links that exhibit diurnal traffic patterns and varying peak hours. Our study reveals that when the data transfer requests arrive uniformly in a day, the probability of a successful transfer increases linearly with the time data is allowed to wait in edge storage. Our study also shows that this probability decreases linearly with $\beta + D$, where $\beta$ and $D$ are the durations of background traffic peak and size of data transfers, respectively, when the network is moderately loaded. And, the benefit of storage will barely increase when the allowed waiting time goes beyond $\beta + D$.

Keywords Bulk data transfers · Edge storage · Performance analysis · Network modeling · Optical networks

1 Introduction
Alongside the enormous amounts of traffic generated by mobile and interactive devices at the edge of the network (Xia et al. 2020), a significant portion of Internet traffic is bulky and delay insensitive (Luo et al. 2020). Examples of such traffic include the data downloads

* Weiqiang Sun
sunwq@sjtu.edu.cn

1 State Key Laboratory of Advanced Optical Communications Systems and Networks, Shanghai Jiao Tong University, Shanghai, China
2 Physics and Information Engineering, Fuzhou University, Fuzhou, China
from planetary-scale observatory instruments and new-generation high-throughput genomic sequencing equipments (Stephens et al. Jul. 2015). Such data are typically transferred to the core of network and propagated through the inter-datacenter optical network (Nicolaescu et al. 2021; Lu et al. 2015). During peak hours, these bulky data might have to compete for bandwidth with interactive applications and among themselves, resulting in congestion on inter-DC optical links and unpredictable completion times (Noormohammadmour and Raghavendra 2018; Wang et al. 2019).

One option for mitigating bandwidth contention is capacity expansion on congested links. However, increasing inter-DC link bandwidth in a wide area network is expensive (Greenberg et al. 2009) and capacity expansion technologies, such as dense wavelength division multiplexing (DWDM), on existing optical fiber is bounded by the nonlinear Shannon limit (Essiambre and Tkach 2012; Zhao et al. Sep. 2013). Real-time Internet backbone traffic statistics indicate that peak-time bandwidth requirements can exceed average values by a factor of two or three (“Center for Applied Internet Data Analysis.” 2020), and designing a network that can handle such levels of peak demand will inevitably lead to increased capital expenditures and inefficient bandwidth utilization. As an alternative to bandwidth expansion, putting off the delivery of delay-insensitive bulk transfer requests to off-peak hours by storing them at edge of the networks is a cost-effective way to reduce bandwidth contention during peak-hours and improve overall network utilization (Nicolaescu et al. 2021), particularly when bandwidth demand is unevenly distributed in space and time (Li et al. Jun. 2018; Lu and Zhu Dec. 2017). However, introducing storage into the transfer process adds an additional dimension of complexity to the problem (Lin et al. Jul. 2020; Lin et al. Mar. 2016), making performance evaluation more difficult. While a considerable amount of interesting research work has been conducted experimentally, or algorithmically, our understanding of the performance gain that storage may bring to data transfer across large geographic areas lags far behind.

In this paper, we aim to model the performance of data transfer across a large geographic area that traverses multiple links with slightly different diurnal traffic patterns (Laoutaris and Rodriguez 2018; Wu et al. 2017). We propose a visually intuitive tool named Network State Progression Diagram (NSPD) to capture the randomness of request arrivals and bandwidth availability. With the tool, we derive closed-form equations for data transfer with edge storage (ES) in a three-DC linear network. The equations are then used to demonstrate how data transfer demand, network load, and traffic peak variability may influence network performance.

The remainder of this paper is organized as follows. In Sect. 2, we review the existing work on storage assisted data transfer. In Sect. 3, we introduce the system model along with the assumptions. In Sect. 4, we mathematically model the network and scheduling with NSPD. In Sect. 5, we present numerical results. We conclude the paper in Sect. 6.

2 Related work

A review of the existing research on storage assisted transfer reveals that such studies have been primarily focused on scheduling algorithms, aiming to maximize flow, minimize cost, or to achieve other objectives.

Researches on flow maximization include the study by Patel et al. (Patel et al. 2008), in which the allocation of bandwidth and buffer space was considered. Iosifidis et al. (Iosifidis et al. 2011; Iosifidis et al. Jun. 2017) developed the joint storage control-routing policy for
max-flow problem defined by utilizing the minimum necessary storage capacity and proper storage placement. In (Tong et al. 2016), Tong et al. coped with randomly varying residual bandwidth and stabilized an inter-datacenter network without prior knowledge of the link residual bandwidth.

Researches on cost minimization include the online optimizer Postcard (Yuan et al. 2012), which was proposed to minimize operational costs on inter-datacenter traffic with store-and-forward (SnF). Nandagopal et al. (Nandagopal and Puttaswamy 2012) minimized the overall peak bandwidth to reduce billable bandwidth usage, and the purchasing capacity is incremented once the residual capacity is insufficient. Chhabra et al. (Chhabra et al. 2010) proposed a scheme in which varying capacities and costs were assigned to both the network and individual links to minimize cost on a layered storage network.

In (Wang et al. Aug. 2014), Wang et al. sought to reduce the temporal peak traffic load on links by applying an SnF transfer mode and a lexicographical minimization approach to balance the traffic load. S. Su et al. scheduled dynamic bandwidth resources for multiple bulk data transfers and employed max–min fairness to periodically optimize the network concurrent allocation rate (Su et al. Dec 2014). Wu et al. proposed an underlying bulk data transfer service featuring optimal routing for distinct chunks over time to maximize the overall weight of all jobs (Wu et al. 2017). Long Luo et al. took into account the revenue earned from deadline-met transfers and the penalty paid for deadline-missed ones (Luo et al. 2018). Aiqin Hou et al. investigated bandwidth scheduling for multiple reservations, with the objective of maximizing the ratio of successful scheduling (Hou et al. 2018).

In an effort to study real-world application of storage assisted transfer, Laoutaris et al. showed that without additional transit cost, 24 terabytes of data may be transferred across different time-zones (Laoutaris et al. Dec 2013). The authors previously introduced NetStitcher, a system for stitching together unutilized bandwidth via unconstrained assistive storage (Laoutaris et al. Aug 2011). In another work (Li et al. 2012), the authors introduced an ISP-friendly bulk data transfer strategy to significantly reduce inter-domain traffic. In (Shi et al. 2011), iDTT, an overlay architecture providing delay-tolerant data transfer service for P2P applications based on the underutilized network capacity, was introduced. In (Zhang et al. Feb 2017), H. Zhang et al. performed on-line admission control and enforcement to implement deadline-based network abstraction (Zhong et al. 2019).

Although storage assisted transfer has been studied extensively in various contexts in the past literature, to the best of our knowledge, little analytic modeling efforts have been carried out.

3 System Model and Assumptions

The inter-DC network is simplified into a concatenation of $N$ links, as in Fig. 1. Background traffic $f(t)$ on the links exhibits diurnal patterns, with alternating busy and idle periods and a repeating period of $T$, as shown in Fig. 2(a). During busy periods, link bandwidth is heavily utilized by background traffic and bulk data is assumed to be off limits to avoid severe network congestion, and hence the state of the links can be binary, i.e., busy or idle. Transfer through a link is only possible during idle periods, and in the case of multiple links, only when there exists overlapping idle periods on all the links. The time of center and duration of busy periods, $\mu$ and $\beta$ respectively, are assumed to be randomly distributed on different links, and are assumed to have a joint probability distribution $\psi(\mu; \beta)$. The distribution of a network can be obtained by predicting based on historical traffic data or
studying the characterization of the traffic features (Alasmar et al. 2019; Krishnaswamy et al. 2020; Muelas et al. 2020; García-Dorado, et al. 2011). Requests arrive to the network with distribution \( p_r(t) \) in \( T \), and have a maximum allowed waiting time of \( W \) (i.e., deadline). Upon arrival, a request under consideration will see the network (i.e., the links) in some random state and its fate will be determined based on whether or not there exists an overlapping window wide enough to complete its transfer in the future. Benefiting from the sparsity of bulk requests (Jurkiewicz et al. 2021), we can handle only one such request in our model, by assuming that other requests “disappear” into the background traffic.

## 4 The mathematical model

The basic idea is that, when a request arrives into the network, it will see the network in some random state. By knowing the distribution of request arrivals and future network state evolutions, based on which we can decide whether a request can be transferred, we will be able to determine the probability of successful transfers. Below, we first introduce the Network State Progression Diagram (NSPD), then discuss how calculations can be made with NSPD, and present the modeling of a 2-link case as an example.

### 4.1 The network state progression diagram (NSPD)

Let \( D \) be the transfer time of a request. Given \( \mu \) and \( \beta \), we use \( g_{\beta,\mu}(t) \) to denote the window within which a successful transfer may start on a certain link. The relation of \( g_{\beta,\mu}(t) \) and \( f(t) \) is shown in Fig. 2a and b. Note that, since the “fate” of a request is solely determined by the network state evolution it sees upon arrival, what matters to us is the relative time difference between the arrival and current network state, instead of the absolute time of arrival. Thus, we can further normalize \( g_{\beta,\mu}(t) \) by \( \hat{g}_{\beta}(t) \) whose values...
in one period start with 0 and end with 1 in Fig. 2(c), and every point on it corresponds to a possible evolution requests could see. For example, the same network state evolution over time will be seen when a request with the arrival time \( t_0 \) arriving at \( g_{\beta,\mu}(t) \) (in Fig. 3a) and when it with the arrival time \( t_0 + \Delta t \) arriving at \( g_{\beta,\mu+\Delta t}(t) \) (in Fig. 3b), because both the relative time differences of the two cases are \( \mu - t_0 \). Also, the point \( x \) of \( \hat{g}_{\beta}(t) \) (in Fig. 3c) can represent that network state evolution (in Fig. 3d) as the same network state evolving from that time. Mathematically, for all requests arriving at an arbitrary time \( t \), they will see the same network state evolution \( x \) if:

\[
    x = t - \mu + \beta/2 + D, \quad 0 < x \leq T
\]  

When a request arrives to the network, the evolution is the combination of evolutions of all links starting from the time of arrival. We “attach” \( \hat{g}_{\beta,\mu}(t) \) of all links to an \( N \)-dimensional space, such that any point in the space can represent a certain combination of evolutions of all links, as shown in Fig. 4. With (1), the probability that an arriving request sees evolution \( x \) can be obtained by
where $x = (x_1, x_2, \ldots, x_N), \mu = (\mu_1, \mu_2, \ldots, \mu_N)$, and $\beta = (\beta_1, \beta_2, \ldots, \beta_N)$.

Considering the periodicity of $\hat{g}_{\beta_n}(t)$, it is necessary to include only one period of $\hat{g}_{\beta_n}(t)$ into the space. Network state evolutions will “wrap” around the boundary of the standard $\hat{g}_{\beta_n}(t)$. For example, Fig. 4 shows an exemplary evolution. With $(x_1, x_2)$, the evolution starts from a point where the transmittable window has not yet open on either links, followed by an open transmittable window on the vertical axis, and then by open transmittable windows on both axes. When it reaches $T$ on the vertical axis, the window on that axis closes but that on $x_1$ remains open until $T$ is also reached on the horizontal axis. After that, the process will repeat. It is worth noting that in the case of edge storage, evolutions will always progress along some diagonal path, because requests stored at the edge of the network will see evolutions on all links progressing simultaneously at the same pace.

For requests arriving at point $x$ (i.e., requests see that particular evolution), a successful transfer can be made, if $[D + \beta_n, T]$ (for $n = 1, 2, \ldots, N$), of all links can be reached simultaneously within the waiting time limit $W$ (Fig. 5). We define $S_W$ as the set of points under which a request can be successfully transferred within $W$. That is, $x \in S_W$, if there exists a waiting time $w \in [0, W]$ that satisfies

$$D + \beta_n + i_n T \leq x_n + w \leq T + i_n T \quad \text{for all } n = 1, 2, \ldots, N$$

where $0 \leq x_n \leq T$, $i_n = \left[ \frac{x_n + w}{T} \right]$, and $\prod_{n=1}^{N} i_n = 0$ which means it also can reach $[D + \beta_n, T]$ in the following periods for some of the links.

With the discussions above, we have the probability of a successful transfer, denoted as $P$, the integral of $p_c$ over $S_W$, that is:

$$P = \int_{S_W} p_c(x, \beta) dx d\beta$$

### 4.2 General picture for the two links case

Even though NSPD can be used to model networks with potentially arbitrary number of links, for simplicity reasons, we limit our following discussions to the two links case. The center of busy-hour ($\mu$) on the two links follows a quadrilateral pyramidal distribution.
(QPD) uniquely determined by the maximum offset interval $d$, as an alternative to the truncated normal distribution (Muelas et al. 2020) with the same interval $d$ and maximum probability value, as shown in Fig. 6. Durations of busy periods are assumed to be $\beta$ with probability 1 on both links. Requests are assumed to arrive uniformly in a day. We also assume that $d$ is smaller than three hours, as the center of busy-hour appears typically only in a window of a few hours (Muelas et al. 2020).

Using (1), (2), and the assumption of QPD on busy-hour distribution, we can get $p_e(x)$. The process of obtaining $p_e(x)$ is not important to our discussions and thus is omitted here. The readers are referred to Appendix A for details. We have for the two links case:

$$p_e(x_1, x_2) = \begin{cases} 
\frac{3(d - |x_1 - x_2|)^2}{2Td^3}, & |x_1 - x_2| \leq d \\
\frac{3(d - T + |x_1 - x_2|)^2}{2Td^3}, & T - d \leq |x_1 - x_2| \\
0, & d \leq |x_1 - x_2| \leq T - d 
\end{cases} \tag{5}$$

The areas within which $p_e(x)$ takes a non-zero value is shown in Fig. 7(a). Physically, these areas include all possible state evolutions on both links, under the constraint of the maximum offset interval $d$. The stripe along the center diagonal contains the evolutions where the transmittable windows on both links are fairly aligned, while the two triangles on the corners contain the cases where the transmittable windows are further apart, but still overlap. Note that $p_e(x)$ integrated over all shaded areas in Fig. 7(a) will result in a 1, i.e., probability of 1.

The calculation of $S_W$ can be separated into three different cases, when the network is loaded differently as measured by $\beta + D$. Under the three cases, different parts of the shaded areas in Fig. 7(a) are covered. This process can be intuitively understood by sweeping $S_0$ diagonally (in the 2-dimensional space) toward smaller values of $(x_1, x_2)$, as shown in Fig. 7(b-d). $S_0$ itself (with vertical line shades), together with the areas it sweeps through (with horizontal line shades), overlaps with the shaded areas in Fig. 7(a), and the resulting overlapping areas indeed form the collection of points where a successful transfer may take place. Like $p_e(x)$, calculation of $S_W$ is placed in Appendix B for brevity.

From (5) we know points along the same diagonal path have the same value of $p_e$. Thus, an increased $W$ that “sweeps” $S_0$ diagonally toward the upper left corner will at first result in a linear increase in the overall probability of a successful transfer. When the “sweeping” exceeds boundary in the upper left corner and wraps around, part of the swept-through area will overlap with $S_0$, thus slowing down the increase in the overall probability. As shown in Fig. 7, when $d$ is small, this slowed down increase happens only briefly. When the entire space has been swept through by a long enough $W$, increasing it further will
not bring additional benefits, i.e., no more increase in the overall probability. It is thus clear that when $W$ increases from zero, the probability of successful transfer will increase linearly first, and then quickly converge to some constant value. And this holds true under all three cases, i.e., when $0 < \beta + D < d$ (light load), $d \leq \beta + D < T-d$ (medium load), and $T-d \leq \beta + D < T$ (heavy load), respectively.

Based on these observations, the probability of a successful transfer can be depicted roughly in Fig. 8, in which $W_L$ and $W_S$ are defined as the boundaries of the three sections. In $[W_L, W_S]$, $P$ exhibits a non-linear relation with $W$. When $d$ is small (no more than 3 h in our study), as relative to $T$ of 24 h, this region is small, so discussions are omitted here and the corresponding equations are placed in Appendix C instead.

Thus the relation between $P$ and $W$ can be expressed in general as

$$P = \begin{cases} kW + P_0 & W \leq W_L \\ P_S & W > W_S \end{cases}$$

(6)
Intuitively, \( P_0 \) is the probability of a successful transfer without any storage. \( P_S \) is the maximally achievable probability of a successful transfer when \( W \) is large enough. Based on the NSPD presented above, \( P_0, P_S \), as well as \( k \), can be calculated. Thus, we can get the equations of \( P \) for the three different cases in Fig. 7(b–d). And based on the NSPD, \( W_L \) and \( W_S \) of the respective cases can also be determined. It is clear that for all three cases, \( W_L \) equals \( \beta + D \), and \( W_S \) will have to be determined separately in each case. We summarize the components under all three cases in Table 1.

### 4.3 Equations under Medium Load (\( d \leq \beta + D < T-d \))

As \( W \) increases from zero, \( S_0 \) “sweeps” toward the upper left corner, and the probability of successful transfer will increase linearly with \( W \) before it reaches \( \beta + D \) (\( W_L \) in all three cases). As requests are assumed to arrive uniformly in \( T \), even increase on \( S_W \) will result in even increase on the probability of successful transfer, and such, the slope is \( 1/T \). When \( W \) goes beyond \( \beta + D \), i.e., when \( S_0 \) sweeps beyond the corner, the points with negative coordinates on both dimensions will be covered, but such points will not contribute to the probability of successful transfer, because they are equivalent to the points around the lower-right corner, which already have been covered by \( S_0 \). This leads to smaller and slowly decreasing (thus non-linear) slope between \( W_L \) and \( W_S \).

From Fig. 7(b), it can be seen that when \( W \) is no less than \( \beta + D + d \), all shaded area in Fig. 7(a) will have been swept through. This leads to successful transfer probability of 1, and increasing \( W \) further will not result in any further increase in \( P \). Thus, \( W_S \) equals \( \beta + D + d \).

### 4.4 Equations under light load (\( 0 < \beta + D < d \))

From Fig. 7(c), it can be seen that when \( W < \beta + D \), not only the areas along the diagonal are covered, but also the areas in the upper right and lower left corners. That leads to a slightly steeper slope \( k \). At the same time, all shaded areas will already have been fully swept through, when \( W \) reaches \( 2\beta + 2D \), leading to a \( W_S \) of \( 2\beta + 2D \). And for the same reason, \( P_0 \) is higher than that in the medium load case, with one additional component in the equation brought about by the partly covered areas in the upper right and lower left corners.

### 4.5 Equations under heavy load (\( T-d \leq \beta + D < T \))

From Fig. 7(d), not all the shaded areas with non-zero probability can be swept through no matter how large \( W \) is, and when \( W \) is larger than \( T \), the swept area will not increase any more. This results in the maximum successful transfer probability of \( 1 - \left( 1 - \frac{T-\beta-D}{d} \right)^3 \). This decreases rapidly if the variability of peak time and load increase. The slope \( k \) is \( \frac{1}{T} \left( 1 - \left( 1 - \frac{T-\beta-D}{d} \right)^3 \right) \), the maximum transfer probability divided by \( T \).

It is worth noting that in this study, bulk transfer requests are assumed to arrive evenly during a day, a key assumption that leads to the linear relation between the probability of successful transfer and the waiting time, as what has been discussed above. If requests arrive unevenly, the relation between the probability of successful transfer and the waiting
| Case                     | $W_L$  | $W_S$  | $k$               | $P_0$                                      | $P_S$  |
|-------------------------|-------|-------|-------------------|--------------------------------------------|--------|
| Medium Load $d \leq \beta + D < T-d$ | $\beta + D$ | $\beta + D + d$ | $\frac{1}{4}$ | $\frac{1}{4} (T - \beta - D - \frac{d}{4})$ | 100%   |
| Light Load $0 < \beta + D < d$ | $2\beta + 2D$ | $\frac{1}{4} \left( 1 + (1 - \frac{\beta + D}{d})^3 \right)$ | $\frac{1}{4} (T - \beta - D - \frac{d}{4} + \frac{(d - \beta - D)^4}{4d^3})$ | 100%   |
| Heavy Load $T-d \leq \beta + D < T$ | $T$ | $\frac{1}{4} \left( 1 - (1 - \frac{T - \beta - D}{d})^3 \right)$ | $\frac{1}{4} (T - \beta - D - \frac{d}{4} + \frac{(d + D + T - \frac{d}{4})^4}{4d^3})$ | $1 - (1 - \frac{T - \beta - D}{d})^3$ |
time will no longer be linear. At the same time, the maximum peak time offset $d$ is assumed to be small (3 h in our numerical study). This results in a relatively small $[W_L, W_S]$ interval, during which the probability of successful transfer has a non-linear increase with regard to $W$, and gradually saturates to its maximum value. This interval will increase, when $d$ increases.

5 Numerical results

In this section, we first validate the rationality of QPD instead of truncated normal distribution (TND) in the model. Then the accuracy of the model is verified with simulations. We present the numerical results obtained with the equations above (under QPD), and compare them with simulated results applying the same conditions and two-link topology as in Sect. 4.2 except under TND to validate the rationality of QPD. In all situations, only marginal differences are observed between the analytical results under QPD and simulated results under TND. Further, the requests with the number $n$ uniformly distributed in a day and the duration following negative exponential distribution having an average duration of $D$ are considered in the simulations ($n$TND). Here, total network load cannot exceed network limit, i.e., $nD + \beta < T$. The number of flows above 500 MB of data is around $10^{-6}$ orders of magnitude in WAN, and decreases exponentially as the request size doubles (Jurkiewicz et al. 2021; "Flow-models." 2021; Bauer et al. 2021). When there are 100 million flows in a day, the number of bulk flow is around 100. So we set $n=100$. We average the results of 10 simulations, each with 15,000 periods (days) and a total of $1.5 \times 10^6$ bulky flows.

Figure 9 shows that longer waiting time results in linear increase in successful transfer probability. Under light or medium load, successful transfer can be guaranteed, regardless of the value of $d$. And to achieve high transfer probability, storage time must match the sum of request duration and network load. The gap between QPD and $n$TND appears only at $W=0$ h. When the load is high, the gap increases as waiting time

![Fig. 9 The successful transfer probability changes with storage time limit $W$ when (a) $d=1$ h and (b) $d=3$ h, under $n=100$ and $D=0.01$ h]
increasing, and could be around 15% when $W=24\text{ h, } d=3\text{ h, } \beta=23\text{ h, and } D=0.01\text{ h.}$ Both the maximal gain from storage and gap are highly dependent on $d$.

Figure 10 investigates how maximum offset interval $d$ affects the successful transfer probability. Little probability change is observed even under medium load in Fig. 10(a) with $\beta=12\text{ h.}$ Intuitively this is because with QPD (and similarly TND) peak hours occur most frequently around the center and decay quickly further apart from it. The small average offset has very little effect when there are enough network resources left. Under heavy load, relatively large decrease on the successful transfer probability can be observed. And the longer waiting time is, the bigger decrease will be (Fig. 10b). The gap between QPD and nTND will be larger when $d$ increases and a long waiting time allowed.

Figure 11 shows the impact of load on network performance, when $d=1\text{ h and } d=3\text{ h.}$ Without storage ($W=0\text{ h,}$), as the load increases, the network performance degrades almost linearly. And the difference between QPD and nTND also decreases. With storage, successful transfer can be guaranteed under low and medium load. And with longer allowed waiting time, the network will be able to serve at a higher load without any loss. When the allowed waiting time is half a day (12 h), the network will be able to serve without any loss under medium load, and will still be able to serve around half of the requests when the network is operating at very high load. Under heavy load ($\beta+D>T-d$), the transfer probability deteriorates rapidly as the load increases, and this explains why there is large difference between the model with one request and nTND with 100 requests, when $\beta=23\text{ h.}$ The model adapts well to all but extreme loads. For example, the gap is less than 1.5% even when $\beta=22.8\text{ h in Fig. 11(b).}$ The gap between model and nTND will be small when transfer probability decreases relatively slowly with increasing the overall network load. Figure 12 shows 3-dimensional view of the successful transfer probability against network load $\beta+D$ and storage time limit $W$, when $d=3\text{ h.}$

These simulation results suggest our model provides a mathematically tractable yet practically accurate approximation of the system except extreme load situations.

![Figure 10](image_url)

**Fig. 10** The successful transfer probability changes with maximum offset internal $d$ when (a) $\beta=12\text{ h}$ and (b) $\beta=23\text{ h, under } n=100\text{ and } D=0.01\text{ h.}$
Fig. 11 The successful transfer probability changes with busy duration $\beta$ when (a) $d = 1$ h and (b) $d = 3$ h, and (c) storage time limit $W$ when $d = 3$ h.

Fig. 12 The successful transfer probability changes with load $\beta + D$ and storage time limit $W$ when $d = 3$ h.
6 Conclusions

In this paper, we model bulky data transfer across multiple congested links with a visually intuitive tool we call Network State Progression Diagram (NSPD). Closed form equations are obtained for the successful transfer probability of a single bulk transfer request, in a network with two links. We show how the successful transfer probability may be affected by network load and peak time distribution on the links. Our study provides useful insights in the operation of networks capable of edge storage. Even though the equations are obtained under the assumption of quadrilateral pyramidal distributed (QPD) busy hour, we prove through simulations that replacing it with TND will only bring marginal differences. And the model works well for multiple bulk requests except in extreme load cases.

Besides edge storage, intermediate storage in network such as storage in datacenters provides more flexibility and complexity in data transfer. In our next step, we wish to apply the concept of NSPD to model the performance of networks with intermediate storage, and study in closed forms how the two different storage strategy will behave under different network setups.

Appendix A

The calculation process of $p_e(x)$ for the 2 links case is as follows.

Under QPD, assuming the distribution center of $\mu_1$ and $\mu_2$ are both $\bar{u}$, the $\psi(\mu_1,\mu_2)$ shown as Fig. 5 is

$$\psi(\mu_1, \mu_2) = \frac{3}{d^2}(1 - \frac{2}{d} z), \text{ where } z = \begin{cases} \bar{u} - u_2, & \bar{u} - \frac{d}{2} < u_1 < \bar{u} + \frac{d}{2} < u_2 < u_1 \\ \bar{u} - u_2, & \bar{u} < u_1 < \bar{u} + \frac{d}{2} < u_2 < 2\bar{u} - u_1 \\ u_2 - \bar{u}, & u_1 < \bar{u} < u_2 < u_1 + \frac{d}{2} \\ u_2 - \bar{u}, & u_1 < \bar{u} < u_2 < 2\bar{u} - u_1 \\ u_1 - \bar{u}, & u_1 < \bar{u} < u_1 + \frac{d}{2} \end{cases}$$ (7)

For $p_e(x)$ in (2), transforming the integral object from $t$ to $\mu_1$ according to (1), the values of $\mu_1$ and $\mu_2$ outside $[0,T]$ should be mapped to values within $[0,T]$ considering the periodicity. And with $p_e(t) = 1/T$, and we have

$$p_e(x) = \frac{1}{T} \int_0^T \psi(\mu_1, \mu_1 + x_1 - x_2 + kT) d\mu_1, \text{ where } k = \begin{cases} -1, & \mu_1 + x_1 - x_2 > T \\ 0, & 0 \leq \mu_1 + x_1 - x_2 \leq T \\ 1, & \mu_1 + x_1 - x_2 < 0 \end{cases}$$ (8)

After piecewise integral, we can get (5).
Appendix B

The process of deriving $S_W$ is as follows.

With a maximum waiting time $W$ no greater than $T$, it is sufficient to consider the cases where $i_1$ and $i_2$ take values no greater than 1, in (3).

With (3), for all $(x_1, x_2)$, $0 \leq x_1, x_2 \leq T$, that belongs to $S_W$, we should have.

1) when $i_1 = 0$ and $i_2 = 0$,

$$\begin{align*}
\{ & \beta + D \leq x_1 + w \leq T \\
& \beta + D \leq x_2 + w \leq T 
\} 
\end{align*}$$

(9)

2) when $i_1 = 0$ and $i_2 = 1$,

$$\begin{align*}
\{ & \beta + D \leq x_1 + w \leq T \\
& \beta + D + T \leq x_2 + w \leq 2T 
\} 
\end{align*}$$

(10)

3) when $i_1 = 1$ and $i_2 = 0$,

$$\begin{align*}
\{ & \beta + D + T \leq x_1 + w \leq 2T \\
& \beta + D \leq x_2 + w \leq T 
\} 
\end{align*}$$

(11)

where $0 \leq w \leq W$ under all three cases above. $S_W$ is thus the union of (9) – (11), joint by $0 \leq w \leq W$.

When $0 \leq \beta + D \leq T/2$,

1) when $W < \beta + D$,

$$S_W = \left\{ \begin{array}{l}
\beta + D - W \leq x_1 \leq \beta + D \\
\beta + D - W \leq x_2 \leq T - \beta - D + x_1
\end{array} \right\} \cup \left\{ \begin{array}{l}
\beta + D \leq x_1 \leq T - W \\
\beta + D \leq x_2 \leq T
\end{array} \right\} \cup \left\{ \begin{array}{l}
T - W \leq x_1 \leq T \\
T - W \leq x_2 \leq T
\end{array} \right\}$$

(12)

2) when $\beta + D \leq W < T$,

$$S_W = \left\{ \begin{array}{l}
0 \leq x_1 \leq T \\
0 \leq x_2 \leq T
\end{array} \right\} - \left\{ \begin{array}{l}
0 \leq x_1 \leq 2\beta + 2D - W \\
T - \beta - D + x_1 \leq x_2 \leq \beta + D + T - W
\end{array} \right\} - \left\{ \begin{array}{l}
T - \beta - D \leq x_1 \leq \beta + D + T - W \\
0 \leq x_2 \leq \beta + D - T + x_1
\end{array} \right\}$$

(13)

3) when $W \geq T$,

$$S_W = \left\{ \begin{array}{l}
0 \leq x_1 \leq T \\
0 \leq x_2 \leq T
\end{array} \right\}$$

(14)

When $T/2 < \beta + D \leq T$,

1) when $W < \beta + D$,

$$S_W = \left\{ \begin{array}{l}
\beta + D - W \leq x_1 \leq T - W \\
\beta + D - W \leq x_2 \leq T - \beta - D + x_1
\end{array} \right\} \cup \left\{ \begin{array}{l}
T - W \leq x_1 \leq \beta + D \\
\beta + D - T + x_1 \leq x_2 \leq T - \beta - D + x_1
\end{array} \right\} \cup \left\{ \begin{array}{l}
\beta + D \leq x_1 \leq T \\
\beta + D - T + x_1 \leq x_2 \leq T
\end{array} \right\}$$

(15)
2) when $\beta + D \leq W < T$,
\[
S_w = \begin{cases} 
0 \leq x_1 \leq T - \beta - D \\
0 \leq x_2 \leq T - \beta - D + x_1 \\
T - \beta - D \leq x_1 \leq \beta + D \\
\beta + D - T + x_1 \leq x_2 \leq T - \beta - D + x_1 \\
\beta + D \leq x_1 \leq T \\
\beta + D - T + x_1 \leq x_2 \leq T \\
0 \leq x_1 \leq W - \beta - D + x_1 \\
\beta + D + T - W \leq x_2 \leq T \\
W - \beta - D + x_1 \leq x_1 \leq T - \beta - D \\
\beta + D + x_1 \leq x_2 \leq T \\
\beta + D + T - W \leq x_1 \leq T \\
0 \leq x_2 \leq x_1 - \beta - D
\end{cases}
\] (16)

3) when $W \geq T$,
\[
S_w = \begin{cases} 
0 \leq x_1 \leq T - \beta - D \\
0 \leq x_2 \leq T - \beta - D + x_1 \\
0 \leq x_1 \leq T - \beta - D \\
\beta + D + x_1 \leq x_2 \leq T \\
T - \beta - D \leq x_1 \leq \beta + D \\
\beta + D - T + x_1 \leq x_2 \leq T - \beta - D + x_1 \\
\beta + D \leq x_1 \leq T \\
0 \leq x_2 \leq x_1 - \beta - D \\
\beta + D \leq x_1 \leq T \\
\beta + D - T + x_1 \leq x_2 \leq T
\end{cases}
\] (17)

**Appendix C**

The successful transfer probability when $W_L \leq W \leq W_S$.

a. Medium load ($d \leq \beta + D < T - d$)
\[
P(\beta, D, d, W) = 1 - \frac{(\beta + D + d - W)^4}{4Td^3}
\] (18)

b. Light load ($0 < \beta + D < d$)
\[
P(\beta, D, d, W) = 1 - \frac{(\beta + D - d)^3(7\beta + 7D - 4W + d)}{4Td^3} - \frac{(\beta + D - W + d)^4}{4Td^3}
\] (19)

c. Heavy load ($T - d \leq \beta + D$)
\[
P(\beta, D, d, W) = 1 - \frac{1}{4T d^3} \left( (\beta + D + d - W)^4 - W(\beta + D + d - T)^3 \right) \quad (20)
\]

**Funding** Funding was provided by National Key Research and Development Project of China (Grant No. 2019YFC1709800), National Natural Science Foundation of China (Grant Nos. 61433009, 61671286).

**Declaration**

**Conflict of interest** The authors confirmed no potential conflict of interest.

**References**

"Center for Applied Internet Data Analysis." [Online]. Available: http://www.caida.org/data/realtime/passive/. Accessed on: Dec.1, 2020.

"Flow-models." [Online]. Available: https://github.com/piotrjurkiewicz/flow-models/tree/master/data/agh_2015. Accessed on: Oct 15, 2021.

Alasmar, M., G. Parisi, R. Clegg and N. Zakhleliu, "On the Distribution of Traffic Volumes in the Internet and its Implications." *IEEE INFOCOM*, pp. 955–963, 2019.

Chhabra, P., V. Erramilli, N. Laoutaris, R. Sundaram, and P. Rodriguez. Algorithms for constrained bulk-transfer of delay-tolerant data. (2010, May). Presented at *IEEE ICC*. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5502217&isnumber=5501741.

Essiambre, R., Tkach, R.W.: Capacity Trends and Limits of Optical Communication Networks. Proc. IEEE *100*(5), 1035–1055 (2012)

García-Dorado, José Luis, et al. "Characterization of the busy-hour traffic of IP networks based on their intrinsic features." *Computer Networks*, vol. 55, no. 9, pp. 2111–2125, 2011.

Greenberg, A., J. Hamilton, D. A. Maltz, and P. Patel, "The Cost of a Cloud: Research Problems in Data Center Networks." *Acm Sigcomm Comp Com*, vol. 39, no. 1, pp. 68–73, Jan. 2009.

Hou, A. Q. *et al*. (2018, Nov). Bandwidth Scheduling for Big Data Transfer with Deadline Constraint between Data Centers. Presented at *IEEE /ACM Innovating the Network for Data-Intensive Science (INDIS)*. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8648791.

Iosifidis, G., Koutsopoulos, I., Smaragdakis, G.: Distributed Storage Control Algorithms for Dynamic Networks. *IEEE/ACM Trans. Networking* 25(3), 1359–1372 (2017)

Iosifidis, G., I. Koutsopoulos and G. Smaragdakis. (2011, April). The Impact of Storage Capacity on End-to-End Delay in Time Varying Networks. Presented at *IEEE INFOCOM*. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5934938&isnumber=5934870.

Jurkiewicz, P., Rzyym, G., Borylo, P.: Flow Length and Size Distributions in Campus Internet Traffic. *Comput. Commun.* 167, 15–30 (2021)

Krishnaswamy N, Kiran M, Singh K, et al, “Data-driven learning to predict wan network traffic,” Proceedings of the 3rd International Workshop on Systems and Network Telemetry and Analytics. 2020.

Laoutaris, N., Sirivianos, M., Yang, X.Y., Rodriguez, P.: Inter-Datacenter Bulk Transfers with NetStitcher. *Acm Sigcomm Comp Com* 41(4), 74–85 (2011)

Laoutaris, N., Smaragdakis, G., Stanoevic, R., Rodriguez, P., Sundaram, R.: Delay-Tolerant Bulk Data Transfers on the Internet. *IEEE Acm T Network* 21(6), 1852–1865 (2013)

Laoutaris N. and P. Rodriguez. (2018). Good Things Come to Those Who (Can) Wait or How to Handle Delay-tolerant Traffic and Make Peace on the Internet. Presented at ACM *HotNets*. [Online]. Available: http://conferences.sigcomm.org/hotnets/2008/papers/20.pdf.

Li, W.X., Zhou, X.B., Li, K.Q., Qi, H., Guo, D.K.: TrafficShaper: Shaping Inter-Datacenter Traffic to Reduce the Transmission Cost. *IEEE ACN T Network* 26(3), 1193–1206 (2018)

Li, Y., H. Wang, P. Zhang, J. Dong, and S. Cheng. (2012, Sept). D4D: inter-datacenter bulk transfers with ISP friendliness. Presented at *CLUSTER*. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6337831&isnumber=6337766.
Lin, X., Sun, W.Q., Veeraraghavan, M., Hu, W.S.: Time-Shifted Multilayer Graph: A Routing Framework for Bulk Data Transfer in Optical Circuit-Switched Networks With Assistive Storage. J Opt Commun Netw 8(3), 162–174 (2016)

Lin, X., Yue, S.N., Tan, Y.L., Sun, W.Q., Veeraraghavan, M., Hu, W.S.: A Partial Store-and-Forward Scheduling Method for Inter-Datacenter Bulk Data Transfers. IEEE Access 8, 128167–128181 (2020)

Lu, P., Zhu, Z.Q.: Data-Oriented Task Scheduling in Fixed- and Flexible-Grid Multilayer Inter-DC Optical Networks: A Comparison Study. J Lightwave Technol 35(24), 5335–5346 (2017)

Lu, P., L. Zhang, X. Liu, J. Yao and Z. Zhu, "Highly efficient data migration and backup for big data applications in elastic optical inter-data-center networks," IEEE Network, vol. 29, no. 5, pp. 36–42, Sep.-Oct. 2015.

Luo, L., H. F. Yu, K. T. Foerster, M. Noormohammadpour, and S. Schmid, "Inter-Datacenter Bulk Transfers: Trends and Challenges," IEEE Network, vol. 34, no. 5, pp. 240–246, Sep. 2020.

Luo, L., H. F. Yu, Z. L. Ye, and X. J. Du. (2018, Apr). Online Deadline-aware Bulk Transfer over Inter-Datacenter WANS. Presented at IEEE INFOCOM. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8485828.

Muelas, D., García-Dorado, J.L., Albandea, S., López de Vergara, J.E., Aracil, J.: On the dynamics of valley times and its application to bulk-transfer scheduling. Comput. Commun. 164, 124–137 (2020)

Nandagopal T. and K. P. N. Puttaswamy. (2012, May). Lowering inter-datacenter bandwidth costs via bulk data scheduling. Presented at: CCGrid [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6217479.

Nicolaescu, A. -C., Maskorakis, S. and I. Psaras, "Store Edge Networked Data (SEND): A Data and Performance Driven Edge Storage Framework," IEEE INFOCOM , pp. 1–10, 2021.

Patel, A., Tacca, M. and J.P. Jue. (2008, Feb). Time-shift circuit switching. Presented at OFC/NFOEC. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=3784657.

Shi, C., M.H. Ammar and E.W. Zegura. (2011, Dec). Delay tolerant bulk data transfer for P2P file sharing systems. Presented at GLOBECOM. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5784872&isnumber=57848793.

Wang, Y.W., Su, S., Liu, A.X., Zhang, Z.B.: Multiple bulk data transfers scheduling among datacenters. Comput Netw 68, 123–137 (2014)

Wang, J. H., J. L. Wang, C. Q. An, and Q. L. Zhang, "A survey on resource scheduling for data transfers in inter-datacenter WANs," Comput Netw; vol. 161, pp. 115–137, Oct. 2019.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.