Routing for Deadline-constrained Bulk Data Transfers Based on Transfer Failure Probability

Yaoquan Zhong, Wei Guo, Yaohui Jin, Weiqiang Sun, Weisheng Hu

Abstract—Bulk data transfers, which require reliable and efficient transfer of terabits or even petabits of data for data-intensive applications, have been extensively studied. In these applications, usually, a specified amount of data needs to be transferred within strict deadline. Previous researches mainly focus on routing deadline-constrained bulk data transfers to improve the utilization of network resources for guarantying deadline requirements, without considering network failures. This paper proposes novel routing algorithm for deadline-constrained bulk data transfers based on Transfer Failure Probability using continuous-time Markov model. The goal of the novel algorithm is to guarantee the deadline requirement by decreasing the failure probability of data transfer request under network failure settings. Simulation evaluates the performance and demonstrates the effectiveness of our algorithm in terms of deadline violation ratio and service blocking ratio under network failures.

Index Terms—bulk data transfer, data-intensive applications, deadline-constrained, transfer failure probability

I. INTRODUCTION

Bulk data transfers, which require reliable and efficient transfer of terabits or even petabits of data for data-intensive applications, have been extensively studied. In these applications, usually, a specified amount of data needs to be transferred within strict deadline. For example, e-VLBI applications [1], which combine many radio telescopes by the technology of interferometry, can achieve the effect of large telescope observation. Traditionally, each telescope transfers huge e-VLBI data to the same data computing center to process the accurate science results using supercomputer and large capacity storage facility. Usually, these data transfers have deadline requirement to realize the real-time data transfer for rapid turnaround. There are mainly two situations which may cause the transfer request not being completed before the deadline. First situation is that there is no enough network resource to support the transfer, stated as resource-constrained blocking in this paper, which can be shown by service blocking ratio. Another case is that, although the lightpath is scheduled to support the transfer request, because of the network failures (e.g., fiber or amplifier failure), the transfer request cannot be finished before deadline, stated as deadline violation, which can be shown by deadline violation ratio. Both of these situations will lead the deadline requirement to be violated.

Previous researches mainly focus on scheduling bulk data transfers to improve the utilization of network resources for guarantying deadline requirements, which mainly pay attention to reducing the probability of the first situation mentioned above. Lin et al formulate and investigate the bandwidth-oriented path computation problem for four types of data transfer in [2] and propose heuristic algorithms for two types of data transfer in [3] to improve the utilization of network resources and meet the transport requirements of application users. Chen et al present multi-interval scheduling for deadline-constrained bulk data transfers to minimize network congestion in [4]. Andrei et al address the Bulk Data Routing and Transfer problem and also present linear programming-based solution and heuristic routing strategies to minimizing overall network congestion in [5]. However, the routing scheme for deadline-constrained bulk data transfers to avoid the second mentioned case is rarely been studied in our knowledge. Because of the tremendous bandwidth carried by fiber, network failures (e.g., fiber or amplifier failure) are inevitable. And usually, in bulk data transfer, the transfer time is comparable with failure repair time. Hence the network failures usually cause the deadline requirement being violated.

Protection is one of the important strategies to recover bulk data transfer request when link failure occurs. By preserving extra resources, dedicated-path protection (DPP) or shared-path protection (SPP) can be applied [6] [7] [8] [9]. However, because of reserving extra resources, protection schemes will decrease the network resource efficiency. This will lead other bulk data transfer requests being abandoned because of insufficient network resources, as shown in the first situation. The problem how to minimize the deadline violation ratio without increasing the network congestion situation in deadline-constrained bulk data transfers routing is an important topic which is rarely studied.

In this work, we will propose Transfer Failure Probability to measure the probability that bulk data transfer cannot be finished before deadline for the reason of network failure on the scheduled lightpath and analyze Transfer Failure Probability using continuous-time Markov model. Results show that Transfer Failure Probability may vary by paths and is affected by other factors (e.g., failure rate, data transfer time, etc.). We then formulate the routing problem for deadline-constrained bulk data transfers based on Transfer Failure Probability and propose novel routing algorithm in which routing decisions are dictated by Transfer Failure Probability.

This study is organized as follows. The definition and analytical model of Transfer Failure Probability are proposed in Section II. Section III formulates the routing problem of for deadline-constrained bulk data transfers based on Transfer
Failure Probability and present novel routing algorithm. In Section IV, a performance evaluation of our novel routing algorithm is conducted in terms of deadline violation ratio and service blocking ratio under network failure settings. Section V concludes this study.

II. DEFINITION AND ANALYSIS OF TRANSFER FAILURE PROBABILITY

A. Definition of Transfer Failure Probability

We assume a bulk data transfer request $T$ as: $T = \{s, d, dl, D\}$, where $s$ is the source node and $d$ is the destination node. $dl$ is the deadline of bulk data transfer request $T$. $D$ is the size of the bulk data. We denote one of the $s$-$d$ lightpath candidates satisfying the data transfer requirement as: $P_k$. And $BW_{P_k}$ is the bandwidth that lightpath $P_k$ can provide for the transfer request. We assume the start time of data transfer is $t_b$. Then we can get the end time of this transfer $t_e = t_b + D/BW_{P_k}$.

Transfer Failure Probability is the probability that optical link failure occurs in transfer interval time $[t_b, t_e]$ on the scheduled lightpath $P_k$ for bulk data transfer request $T$.

Based on the assumption above, we will analyze the Failure Probability of the transfer request. First we employ a continuous-time Markov chain to derive the Failure Probability of one single link $L$ in lightpath $P_k$. Then we extend the analysis to the whole lightpath $P_k$ and get the Transfer Failure Probability for the transfer request.

B. Analysis of Transfer Failure Probability

We employ a continuous-time Markov chain to derive the Failure Probability of the single link $L$. We assume that the duration between failures and the failure repair time for link $L$ are exponentially distributed, and mean values are $\lambda$ and $\mu$.

![State-transition diagram of the optical link.](image)

Fig.1 shows the corresponding state-transition diagram. We assume that there are only two states for one optical link: the initial state ‘U’ means up and state ‘D’ means down.

The initial state for optical link means the link state is up when $t = 0$. From continuous-time Markov model, the Failure Probability which represents the probability that optical link state is down at $t$ can be calculated as:

$$P^L(t) = \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda + \mu}$$

The probability that optical link state is down in bulk data transfer interval time $[t_b, t_e]$ equals to the sum of the probability that optical link state is down at $t_b$ and the probability that optical link state is up at $t_e$, but the state transfers to down in $[t_b, t_e]$ as shown in (2).

$$P^L\left\{[t_b, t_e] \right\} = P(t_b) + \left(1 - P(t_b)\right)P^L_{t_b \rightarrow D}$$

The failure probability of the single optical link $L$ with $\lambda$ and $\mu$ in bulk data transfer interval time $[t_b, t_e]$ can be calculated as:

$$P^L\left\{[t_b, t_e] \right\} = 1 - e^{-\lambda(t_e-t_b)} \left[ \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} e^{-(\lambda+\mu)t_b} \right]$$

Suppose a lightpath $P_k = \{L_1, ..., L_m\}$, where $L_j$ is a fiber link, $1 \leq j \leq m$, $m$ is an integer.

The Transfer Failure Probability of bulk data transfer request $T$ using lightpath $P_k$ can be calculated as:

$$P^D_k\left\{[t_b, t_e] \right\} = 1 - \prod_{j=1}^{m} \left\{1 - P^L_{\left\{[t_b, t_e] \right\}}\right\}$$

III. ROUTING PROBLEM FOR DEADLINE-CONSTRAINED BULK DATA TRANSFERS BASED ON TRANSFER FAILURE PROBABILITY AND NOVEL ROUTING ALGORITHM

A. Routing Problem for Deadline-constrained Bulk Data Transfers Based on Transfer Failure Probability

Given a network $G = (V, E)$, where $V$ is the set of nodes, and $E$ is the set of fiber links. Each fiber link contains several wavelength channels. We assume a bulk data transfer request $T = \{s, d, dl, D\}$, as mention in Section II.

The routing problem for deadline-constrained bulk data transfers based on Transfer Failure Probability can be formally stated as follows: route light paths and assign wavelength channels for bulk data transfer request with minimal Transfer Failure Probability to guarantee the deadline requirement by decreasing the Transfer Failure Probability.

B. Novel Routing Algorithm Based on Transfer Failure Probability

We denote the $s$-$d$ lightpath candidates satisfying the data transfer requirement as: $P_{sd} = \{P_1, ..., P_k, ..., P_n\}$, $1 \leq k \leq n$, $n$ is an integer.

The objective of novel routing algorithm is looking for a lightpath so that the Transfer Failure Probability is minimized, which can be denoted as:

$$\min\{P^D_k, 1 \leq k \leq n\}$$
Since the typical value of MTTR is in the order of hours, while MTTF is usually thousands of hours. That is, $1/\lambda > 1/\mu$, $\frac{\mu}{\lambda + \mu} \rightarrow 1$ and $\frac{\mu}{\lambda + \mu} \gg \frac{\lambda}{\lambda + \mu}$.

So (3) can be reduced to:

$$P^L\{(t_b, t_e)\} = \lambda (t_e - t_b)$$

(6)

1. $1-e^{-x} = x$ is a good approximation for $x \rightarrow 0$.

While $P^L\{(t_b, t_e)\} \rightarrow 0$, (4) can be reduced to:

$$P^R\{ (t_b, t_e) \} = \sum_{j=1}^{m} P^L\{ (t_b, t_e) \} = \sum_{j=1}^{m} \lambda_j (t_e - t_b)$$

(7)

From (7), it can be concluded that Transfer Failure Probability of lightpath $P_k$ is related to the bulk data transfer interval time $h = (t_e - t_b)$ and $\lambda_j$ of all the optical links in lightpath $P_k$.

Equation (5) can be reduced to:

$$\text{min}\{\frac{D}{BW_{P_k}}, \sum_{j=1}^{m} \lambda_j, 1 \leq k \leq n\}$$

(8)

Note: $BW_{P_k}$ is the bandwidth that lightpath $P_k$ can provide to data transfer. From (8), we can conclude that the lightpath with minimal Transfer Failure Probability is the lightpath with minimal product of failure rate $\lambda_{P_k}$ and the reciprocal of the available bandwidth $BW_{P_k}$.

So, we can design the routing algorithm for Deadline-constrained Bulk Data Transfers based on Transfer Failure Probability (DBDT-TFP) according to (8). DBDT-TFP algorithm is described as follows:

---

Algorithm 1 DBDT-TFP algorithm

Input: Bulk data transfer request $T = \{s, d, dl, D\}$.

Network $G = (V, E)$, each link $L \in E$ is label with $\lambda_L$ and $\mu_L$.

Output: A lightpath between $s$ and $d$, or null.

1. Check the current network and gain the available sub-network $G_a = (V_a, E_a)$ for bulk data transfer request $T$.
2. K-Shortest Path algorithm [10] is used to pre-compute the candidate routes between $s$ and $d$ based on the link hops ($K$ lightpaths with minimal link hops). If no lightpath can be found, return null, else go to step 3.
3. Calculate Transfer Failure Probabilities of all the $K$ candidates found in step 2, return $P$ with minimal TFP.

---

To compare with DBDT-TFP algorithm, two typical routing algorithms for deadline-constrained bulk data transfer are described, Minimal Transfer Time routing algorithm for Deadline-constrained Bulk Data Transfer (MTT-DBDT) in Alg. 2 [2] and Shared-Path Protection routing algorithm for Deadline-constrained Bulk Data Transfer (SPP-DBDT) Alg. 3.

---

Algorithm 2 MTT-DBDT algorithm

Input: Bulk data transfer request $T = \{s, d, dl, D\}$.

Network $G = (V, E)$, each link $L \in E$ is label with $\lambda_L$ and $\mu_L$.

Output: A lightpath between $s$ and $d$, or null.

1. Check the current network and gain the available sub-network $G_a = (V_a, E_a)$ for bulk data transfer $T$.
2. K-Shortest Path algorithm is used to pre-compute the candidate routes between $s$ and $d$ based on the link hops. If no lightpath can be found, return null, else go to step 3.
3. Select the lightpath that can minimize the transfer time (maximal available bandwidth) from $K$ candidates found from Step 2.

---

Algorithm 3 SPP-DBDT algorithm

Input: Bulk data transfer request $T = \{s, d, dl, D\}$.

Network $G = (V, E)$, each link $L \in E$ is label with $\lambda_L$ and $\mu_L$.

Output: A lightpath between $s$ and $d$, or null.

1. Apply MTT-DBDT algorithm to find a primary path $P$ for transfer request $T$; if no path can be found, return null.
2. Apply SPP scheme to find a link-disjoint backup path $P_2$ with maximized sharing; if no path can be found, return null.
3. Return $P_1$ and $P_2$.

---

IV. SIMULATION AND ANALYSIS

In this section, we will evaluate the performance of DBDT-TFP algorithm by simulation, comparing with MTT-DBDT and SPP-DBDT algorithms stated in Section III.
A 24-nodes US WDM optical network topology (Fig. 2) is used in simulation. Each link is labeled with a number that shows the length of the link. We assume there are 16 wavelengths per link and the capacity of each wavelength is 1 Gbit/s. There is no wavelength conversion and the carried bandwidth is limited by the wavelength capacity.

![Network Topology](image)

Figure 2. The typical US network topology.

Table I shows the typical data on failure rates and failure-repair times of fiber link, where FIT (failure-in-time) denotes the average number of failures in 10^6 hours.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Reference Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiber Link MTTR</td>
<td>24-96 hrs</td>
</tr>
<tr>
<td>( \lambda_{\text{link}} )</td>
<td>100 FIT/km</td>
</tr>
<tr>
<td>( C_{OA} )</td>
<td>0.02/km</td>
</tr>
<tr>
<td>( \lambda_{OA} )</td>
<td>2000 FIT</td>
</tr>
</tbody>
</table>

In this simulation, we assume that the failure rate \( \lambda_j \) for fiber link \( L_j \) is calculated using (9), where \( D_{L_j} \) is the length of link \( L_j \), \( \lambda_{\text{link}} \) is the link failure rate per kilometer, \( \lambda_{OA} \) is the optical amplifier failure rate per kilometer and \( C_{OA} \) is a constant for optical amplifier spacing per kilometer [12]:

\[
\lambda_j = \frac{D_{L_j}}{L_j} \left( \lambda_{\text{link}} + \frac{C_{OA}}{\lambda_{OA}} \right)
\]

(9)

Failures occur independently on each fiber link following a Poisson process. The link failure profile is generated according to Table I. Failure repair time follows a negative exponential distribution, and the mean value of failure repair-time is randomly distributed for each fiber link in three levels (24, 48 and 96 hours).

We assume dynamically-arriving bulk data transfer requests with uniform distribution among all nodes. The bulk data transfer requests will be provisioned with one lightpath (in DBDT-TFP and MTT-DBDT algorithms) or two lightpaths (in SPP-DBDT algorithm), or dropped if none can be found. One year period is investigated, with tens of thousands bulk data transfer requests. We average the results of several runs with different seeds, to improve statistical confidence.

A. Effects of bulk data size

![Deadline Violation Ratio](image)

Figure 3. Deadline violation ratio under different bulk data sizes.

We first compare the deadline violation ratios (number of transfers that violate the deadline requirement because of link failure over the number of total admitted bulk data transfers) of DBDT-TFP algorithm with MTT-DBDT and SPP-DBDT algorithms under different bulk data sizes. We use average bulk data sizes from 2 to 32 petabits, and data transfer arrival rate is fixed at 10 per day. The results are shown in Fig. 3. We observe that DBDT-TFP and SPP-DBDT algorithms have similar performance and always outperform MTT-DBDT algorithm, by reducing overall deadline violation ratio. As the increase of the bulk data size, the deadline violation ratio also increases. This is because larger bulk data size means the longer transfer holding time. This will cause the data transfer request easier in the impact of link failures. SPP-DBDT algorithm has better performance in deadline violation ratio by introducing shared-path protection. However, this performance improvement is at the cost of reserving extra network resource. Because of reserving extra resources, SPP-DBDT algorithm will decrease the network resource efficiency. That will lead other bulk data transfer requests being abandoned because of insufficient network resources, which can be concluded from the high service block ratio of SPP-DBDT algorithm, as shown in Fig. 4.

![Service Block Ratio](image)

Figure 4. Service block ratio under different bulk data sizes.

Service blocking ratio is defined as the rejected data transfers over the total requested transfers. We can observe that DBDT-TFP and MTT-DBDT algorithms have reasonable service block ratio, comparing with the SPP-DBDT algorithm. This is because that DBDT-TFP and MTT-DBDT algorithms...
provide single lightpath for bulk data transfer request. While SPP-DBDT algorithm provision two lightpaths for each data transfer request, which will decrease the network resource efficiency.

B. Effects of network load

We simulate these three algorithms when network load varies from 20 Erlangs to 120 Erlangs, with average bulk data size is set at 4 petabits. The deadline violation ratios are plotted in Fig. 5, which are also consistent with the result in Fig. 3 when average bulk data size is 4 petabits. We can observe that, on average, DBDT-TFP and SPP-DBDT algorithms can achieve roughly 30% lower deadline violation ratio than MTT-DBDT algorithm under this network failure profile. In addition, deadline violation ratios of all these schemes have little change for the investigated range of network loads, which means these schemes show little sensitivity to network load. That is mainly because deadline violation ratio depends on transfer request and failure profiles. We also find that the service blocking ratio of all these schemes will increase with the network loads. While the service block ratio of SPP-DBDT algorithm is highest in these three schemes, this is also because the extra shared-path protection consumes more network resources.

V. CONCLUSIONS

In this work, we propose the mathematical model of Transfer Failure Probability, to describe the failure probability of the scheduled bulk data transfers under network failure settings. Then we present DBDT-TFP algorithm to route lightpaths for dynamic deadline-constrained data transfer requests based on Transfer Failure Probability. Simulation results show that DBDT-TFP algorithm has better trade-off performance in deadline violation ratio and service blocking ratio than MTT-DBDT and SPP-DBDT algorithms.

ACKNOWLEDGMENT

This research is sponsored by the china NSFC under Contract No. 61071080, 60825103, 60632010, the China 973 program under grant 2010CB328204, 2010CB328205, and the China 863 program.

REFERENCES


