Assessing the impact of geographically correlated failures on overlay-based data dissemination

Kyungbaek Kim and Nalini Venkatasubramanian
Dept. of Computer Science
University of California, Irvine, USA
Email: {kyungbak, nalini}@ics.uci.edu

Abstract—This paper addresses reliability of data dissemination applications when there are severe disruptions to the underlying physical infrastructure. Such massive simultaneous physical failures can happen during the geographical events such as natural disasters (earthquakes, floods, tornados) or sudden power outages - infrastructure failures in these cases are geographically correlated. In particular, we focus on overlay based data dissemination mechanisms and explore their ability to tolerate such large geographically correlated failures. Due to the tight correlation between multiple overlay links and a single physical link, a few physical failures may affect lots of overlay links. To enable reliable dissemination under such conditions, we propose overlay network construction methods that incorporate proximity-aware neighbor selection methods to improve the performance of the overlay data dissemination, to the extent possible, in terms of reliability and latency. In this approach, the overlay nodes select neighbors which are most likely distinct in presence of a geographical failure; we show how an overlay structure constructed using our proximity-aware neighbor selection techniques can disseminate data to over 80% of reachable end clients without any significant additional latency under various geographical failure conditions.

I. INTRODUCTION

The use of overlay networks for data dissemination has increased significantly over the past years [3], [4]. There exists today a multitude of protocols and systems that support various forms of information dissemination in overlay networks - large data dissemination [8], streaming data dissemination [16], fast dissemination of medium size data [19] etc. Most efforts emphasize system performance in terms of latency, bandwidth usage, network jitter, ability to deal with heterogeneity. Reliability of information dissemination is critical, especially in cases where the overlay networks are used for mission critical applications such as crisis alerting, enterprise networks, military command and control, and etc [11]. The typical approach to achieving reliable data dissemination in the presence of overlay node/link failures uses techniques such as data retransmission and overlay structure reconfiguration - these approaches are however, time-consuming and incur significant additional latency. For example, the overlay reconfiguration takes in the order of tens of seconds to initiate the reconfiguration process [9], [10]; the recovery time for data retransmission techniques minimally incur the round-trip delay between communicating nodes [6], [7]; and the physical infrastructure recovery processes can vary from hundreds of seconds to days [5]. The above methods incur additional latency in order to ensure reliability of information transfer and are unsuitable in mission critical applications where time is of the essence.

Typical structures used in constructing reliable overlay networks for data dissemination include forest [12], [13], [14], [15], [16], [17] and mesh [11], [18], [19] structures. In these overlay networks, each node has multiple neighbors and hence there exist multiple data paths via which information may reach their final destinations; unless all paths are unavailable simultaneously, data dissemination works without any retransmission or reconfiguration. Path diversity, as enabled in such protocols can support reliable data dissemination, especially when the overlay failures are random. However, in situations where failures are not random, but geographically correlated, as is the case when there are massive disruptions of underlying geography (earthquakes, floods, or tornados), the physical network within the region of the geographical events is damaged catastrophically. With massive events, the region of the simultaneously damaged physical network may be wide. Moreover all the overlay nodes/links which are corresponding to the damaged physical network are not reachable from rest of overlay nodes and the effect of the geographically correlated failures to the reliable overlay network might become significant.

In this paper, we explore the impact of the geographically correlated failures on data dissemination over the reliable overlay network structures such as forest/mesh. Specifically, we propose an overlay network construction mechanism to support reliable data dissemination (to the extent possible) under geographically correlated failures. Our overlay construction mechanism relies on choosing neighbor nodes in the distributed overlay network that share limited physical nodes within a region of a certain size - the expectation is that this will reduce the likelihood that overlay links share the same physical infrastructure under a geographically correlated failure.

To determine the degree of correlation and select more effective neighbors, the neighbor selection process is performed in a proximity-aware manner. The proximity factor is calculated using geographical information about the physical path (set of physical links) of an overlay link. Each overlay/physical node is mapped into a geographical position on the given Euclidean geographical space, and if the Euclidean distance between two nodes is below a certain threshold, the link is given a higher proximity factor value. The proximity factor between overlay neighbors can be calculated by combining the proximity factors of the physical nodes related to the overlay links of neighbors. The neighbor selection process then
chooses the neighbor which has the least proximity factor.

We evaluated the existing overlay construction techniques and the proposed proximity-aware mechanism using a synthetic infrastructure topology generated by Inet topology generator. We consistently observed that the overlay network constructed by using the proximity-aware neighbor selection disseminates the data more reliably and faster than the previous efforts under geographically correlated failures. The rest of the paper is organized as follows. In Section II, we motivate the correlation between the overlay and physical networks and discuss prior efforts in topology-aware techniques for overlay construction. Section III presents the correlation between overlay links from the view of a geographically correlated failure and introduces the proximity-aware overlay neighbor selection technique to construct reliable overlays using forest/mesh structures. In Section IV, we describe our experimental setup and results on the reliability/speed of neighbor selection approaches; we conclude in Section V with future research directions.

II. OVERLAY NETWORK VS PHYSICAL NETWORK:
OBSERVATIONS

An overlay network is constructed over a physical network which consists of a set of geographical distributed physical nodes. An overlay link (between two overlay nodes) essentially consists of multiple physical links, e.g. links between IP routers. Techniques to map overlay links to an underlying physical path use the underlying routing mechanisms available. Overlay nodes have little control over the physical links that connect them to other overlay nodes; due to commonality of the underlying routing mechanisms, it is possible that multiple overlay links share the same physical link. As a result, the failure of a single physical link may cause multiple overlay links to fail.

Despite the fact that a physical node/link failure affects multiple overlay links, physical failures are, typically, infrequent and randomized, and their impact to the overlay network is also randomized and limited. Overlay structures such as forests and meshes inherently support some degree of redundancy; this is often sufficient to support reliable data dissemination in the presence of limited random failures. In a reliable overlay network, where each node has multiple neighbors, dissemination latency is reduced since there are plenty of choices for using the fastest overlay link to disseminate the data.

Techniques that consider the relationship of overlay networks to physical links have been designed, primarily to support better dissemination performance (higher speed and fewer redundant transmissions). Some previous approaches [24], [25] generate a tree-like overlay network by minimizing the usage of the physical links while reaching all the overlay nodes. This topology-aware overlay network minimizes the bandwidth consumption of physical networks and reduces dissemination latency. Efforts such as [20], [21] aim to increase the communication reliability between two overlay nodes by supplying a backup overlay path whose physical path is less likely to overlap with that of the primary overlay path - i.e. use distinct links. These efforts model correlation between overlay and physical networks as the case where exact physical nodes/links overlap. However, when the massive and wide geography disruption occurs, multiple physical nodes/links located within the region of the geographical event are damaged simultaneously. One related effort [23], explores the impact of the geographically correlated failures in order to identify vulnerable network locations that would have the maximum effect on the physical network capacity - this effort is limited to the vulnerabilities of physical nodes and does not model or address impact to the overlay networks, built on the physical networks. In contrast, our effort focuses on constructing an overlay network that is capable of disseminating information in a timely and reliable manner to the rest of the infrastructure (that has not failed) despite unexpectedly geographically correlated failures in the physical infrastructure.

III. PROXIMITY-AWARE RELIABLE OVERLAY NETWORK CONSTRUCTION

A. Correlation between overlay links under a geographical event

As shown in Figure 1, each overlay node is mapped to its gateway router which is a physical node mapped into a position in the given Euclidean geographical space. A gateway router calculates all the routing paths to the other gateway routers, which are the primary routing paths determined by the given routing mechanism in the physical network. An overlay link between two overlay nodes is corresponding to the given primary routing path between their gateway routers. Surely, there are multiple physical routing paths such as backup routing paths. However, because the information of the physical network is not transparent to overlay nodes, overlay nodes can get the limited information of the physical network such as the primary routing paths used currently.

During constructing a reliable overlay network such as forest or mesh, an overlay node selects its multiple overlay neighbors at the application layer. By using the multiple overlay links to the multiple neighbors, the overlay network can disseminate the data more reliably and faster. However,
Algorithm 1 Neighbor selection algorithm for an overlay node

Require: \( T_{\text{neighbor}} := \text{Target number of neighbors} \)
Require: \( M := \text{Set of random overlay nodes} \)
Ensure: \( \text{M.size()} > T_{\text{neighbor}} \)

1: \( N \leftarrow \text{NULL} \) (\( N := \text{set of selected neighbors} \))
2: for all \( m \in M \) do
3: \( \text{Calculate } L_m \) (\( L_m := \text{Overlay Latency to } m \))
4: \( e = m', \) where \( L_m \) is smallest among \( M \)
5: \( \text{N.add}(e), \text{M.remove}(e) \)
6: while \( \text{N.size()} < T_{\text{neighbor}} \) do
7: for all \( m \in M \) do
8: \( \text{CORR}_m \leftarrow 0 \)
9: for all \( n \in N \) do
10: \( \text{CORR}_m = \text{CORR}_m + \text{CORR}(m, n) \)
11: \( e = m', \) where \( \text{CORR}_{m'} \) is smallest among \( M \)
12: \( \text{N.add}(e), \text{M.remove}(e) \)


Algorithm 2 \( \text{CORR}(m, n) \) algorithm for calculating the correlation between two overlay nodes \((m, n)\). \( D(R_{mi}, R_{nj}) \) means Euclidean distance between \( R_{mi} \) and \( R_{nj} \).

Require: \( T_{\text{distance}} := \text{Distance threshold} \)
Require: \( P_m, P_n := \text{Physical path to } m, n \)
\( , \text{where } P_m := \{r_{m1}, r_{m2}, ..., r_{mr}\} \)
\( , \text{where } r_{mr} = \text{rth} \text{Physical node on } P_m \)

1: \( \text{ProximityFactor} \leftarrow 0 \)
2: for all \( r_{mi} \in P_m \) do
3: for all \( r_{nj} \in P_n \) do
4: if \( D(r_{mi}, r_{nj}) < T_{\text{distance}} \) then
5: \( \text{ProximityFactor} = \text{ProximityFactor} + 1 \)
6: Return \( \text{ProximityFactor} \)

the inherent property between an overlay link and a physical routing path may nullify an advantage of the multiple neighbors in the presence of a geographically correlated failure. That is, the overlay links, whose corresponding physical paths are more correlated to each other under the geographically correlated failure, are not helpful for constructing a reliable overlay network.

Let us consider that the overlay node A selects two neighbors in Figure 1. Two overlay links (A, D) and (A, E) share the same physical node e, and D and E should be selected alternatively to exploit the advantage of multiple neighbors. Overlay links (A, C) and (A, D) seems to be distinct to each other under a random node failure. However, when we consider the geographical event like Figure 1, the physical node d used in (A, C) and the physical node e used in (A, D) can fail simultaneously and the two overlay links are also correlated to each other for the geographical event. Because the correlation between physical nodes depends on the size of the geographical event, as the geographical distance between two physical nodes is shorter, the correlation between them is bigger. According to this, the best two neighbors for node A are node B and node E.

B. Proximity-aware overlay neighbor selection

As we described, intelligent neighbor selection is essential to increase the performance of the overlay network in terms of reliability and latency. Algorithm 1 illustrates how an overlay node selects a set of neighbors \((N)\) from a known set of randomly picked overlay nodes \((M)\). We assume that an overlay node can gather some information about the limited number of other random overlay nodes. Initially, the overlay node selects its first neighbor as the one with the least end-to-end latency to realize the objective of fast communication with neighbors. After selecting the first neighbor, we compute the correlation factor \((\text{CORR}_m)\) of unslected random overlay nodes with the least latency neighbor. While calculating \(\text{CORR}_m\) for an unselected overlay node \(m\), note that the correlation factors for the already selected neighbors \((\text{CORR}(m, n))\) are accounted for. For example, if there are 2 already selected neighbors \((A\) and \(B)\), the \(\text{CORR}_m\) is computed by adding \(\text{CORR}(m, A)\) and \(\text{CORR}(m, B)\). Given all the \(\text{CORR}_{m}\), the unselected overlay node \(m\) whose \(\text{CORR}_m\) is the smallest is selected as a new neighbor. If many nodes have same minimum \(\text{CORR}_m\), latency is used as a tie-breaker and the node having the least latency is selected as a new neighbor.

The most important part of the neighbor selection is \(\text{CORR}(m, n)\) presented in Algorithm 2. \(\text{CORR}(m, n)\) calculates the proximity factor between the overlay link to \(m\) and another overlay link to \(n\). First of all, the physical path corresponding to an overlay link is obtained as a set of physical nodes on the path. As alluded to earlier, a physical node is mapped into a point in the given Euclidean geographical space such as Euclidean \(n\)-space and we can calculate the Euclidean distance between any two physical nodes. That is, \(D((p_1, q_1), (p_2, q_2)) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}\), where \((p_1, p_2, p_n, q_1, q_2, q_0)\). When the Euclidean distance between two physical nodes is below a threshold, \(T_{\text{distance}}\), the proximity factor between two physical paths increases by 1. By comparing all pairs of routers in the two paths, the proximity factor between two overlay nodes is computed. \(T_{\text{distance}}\) is a knob that can be used to tune proximity factor values. As \(T_{\text{distance}}\) increases, the proximity factor is able to capture correlations for larger size geographical events.

IV. Evaluation

We evaluate the impact of the geographically correlated failure to the reliable overlay network such as forest and mesh and the effectiveness of proximity-aware overlay neighbor selection.

To figure out that the new selection method can be applicable to various reliable network structures, we apply the method not only to forest but also mesh. In both reliable overlay networks, we assume that there is a single root which generates the data. For forest, there are multiple trees starting from the root, and each tree has a given degree of children which is set to 3 in this evaluation. In mesh, each node has multiple neighbors including the root. For both of forest and mesh, an overlay node selects its children or neighbors among the given random overlay nodes, whose size is set to 30 in this evaluation, and this neighbor selection performs periodically.
In the evaluation, we focus on the data dissemination for the mission critical applications where the reliability and the time are essential. When the root starts disseminating data, it forwards the data to all the trees or all the neighbors in forest and mesh, respectively. According to this, a node can get multiple data from trees or neighbors, and it only forwards the first data, but discards the duplicated data. When failures occur during the dissemination, nodes do not hesitate to fix the overlay/physical network, but forward the current data by using only the currently available network immediately.

The physical network is generated by Inet topology generator [1]. Inet generates topologies on a XY plane (10000,10000) and each node locates on a specific point in the manner of the uniform distribution. We use this XY plane as a Euclidean 2-dimension space and the specific location of a router as the Euclidean coordinate. The inter-router latency is calculated by the same estimation method which is used by Modelnet network emulator [2]. The generated physical network is consisted of 10000 routers and 41152 inter-router links. 1000 stub routers are selected and 1000 end clients attach to each of the stub routers. Each stub router calculates the routing path to all the other stub routers based on a shortest path algorithm and each end-to-end overlay link is mapped to this pre-computed routing path which can be considered as a primary physical path.

The compared parameters are reliability and maximum latency. When geographically correlated failures occur, the overlay/physical nodes inside the failure region are most likely unreachable from outside nodes. Moreover, some outside nodes are also unreachable from the root because the physical paths to them are highly correlated to the failure region. According to this, the reliability is defined as $\frac{\#\text{ of nodes currently reachable from the root}}{\#\text{ of nodes receiving the message}}$. The maximum latency is the time when the last node gets the data. We run evaluations for 20 times and got the average value for each of the parameters.

A. Impact of geographically correlated failures

In Figure 2, we present how the reliable overlay network works on different failures: random overlay failures and geographically correlated failures. To generate a portion of random overlay failures, the random overlay nodes selected uniformly up to the portion of the failures. On the other hand, the geographically correlated failures are generated by choosing a certain size of a rectangular region inside the Euclidean 2-dimension space, and the overlay/physical nodes inside the region fail with the very high probability such as 0.9.

Because we assume that overlay/physical nodes are uniformly distributed in the Euclidean space, the portion of the region can be used to represent the portion of the geographically correlated failures.

A single tree is very vulnerable to the random overlay failures. But the reliable overlay network, where each node selects its neighbors randomly, is very tolerant to substantial random overlay failures like Figure 2(a). Under the geographically correlated failures, a single tree looks almost useless, and the reliability of the reliable overlay networks decreases along with the portion of the failures like Figure 2(b).

B. The effectiveness of proximity-aware neighbor selection

From now on, we more focus on how the neighbor selection methods affect the performance of the data dissemination over the reliable overlay network. We compare the proximity-aware neighbor selection with other 3 neighbor selection methods: random, latency-aware, and topology-aware. The random method chooses the neighbors just randomly. The latency-aware method picks the neighbors with the increasing order of the latency [22]. The topology-aware method selects the neighbors whose primary routing paths are less likely sharing the same physical node [20], [21]. We apply all the methods to both of the forest and the mesh, and we get the very similar results for both of them. In this paper, we present the results for the forest.

In Figure 3(a), the proximity-aware neighbor selection keeps the most reliability under various geographically correlated failures. Under 25% failures, its reliability is 15% and 25% more than latency-aware/topology-aware methods and random method, respectively. The latency-aware method and topology-aware method can achieve more reliability than random method, but it is limited because they do not consider the impact of the geographically correlated failure unlike the proximity-aware method.

The proximity-aware neighbor selection also has a low maximum latency as shown in Figure 3(b). Obviously, the latency-aware method achieves the minimum value of the maximum latency, and the proximity-aware method has a very similar maximum latency to the latency-aware method. In the proximity-aware method, the overlay neighbor which has the least latency is firstly selected, and the other neighbors are selected by comparing the proximity factor-based correlation.
factor between them. Moreover, when multiple overlay nodes have the same correlation factor, their latency is used for the tie-breaker. According to this consideration about the latency, the proximity-aware neighbor selection can achieve high reliability as well as low latency.

In Figure 4(a), we show the performance of the proximity-aware neighbor selection with various distance thresholds. The distance threshold represents the estimated size of the geographical region which the calculated correlation factor supposed to cover. That is, the performance of proximity-aware neighbor selection with distance threshold 0 is similar to that of the topology-aware method. As the distance threshold increases, the proximity-aware neighbor selection prepares for wider geographically correlated failure. At a certain failure, the reliability increases along with the distance threshold, but it almost saturates after a certain point. At 16% and 25% failures, the reliability almost saturates after distance threshold 5000 and 7000, respectively. That is, in order to maximize the performance of the proximity-aware neighbor selection for a certain size of a geographically correlated failure, the distance threshold should be larger than the size of the failure.

Even the number of trees in forest structure increases, the effectiveness of the proximity-aware neighbor selection is still valid like Figure 4(b). As the number of trees increases, the natural reliability of the forest also increases. However, the other neighbor selection methods can not exploit the benefits of multiple trees effectively under the geographical failures, and its reliability is limited. On the other hand, the proximity-aware neighbor selection can achieve about 80% reliability under 25% geographical failures.

V. CONCLUSION

Existing overlay networks based on forest and mesh structures disseminate data with high reliability and speed under random overlay failures. However, these techniques do not work very well under geographically correlated failures; the reliability and speed of data dissemination are decreased significantly in such conditions. We explore a proximity-aware approach neighbor selection which helps construct overlay networks that are resilient to geographically correlated failures. Our technique captures the correlation between physical paths by checking the geographical proximity between every physical node on the alternative paths, and utilizes this information to generate geographically diverse neighbors for each overlay node. We observe that the data dissemination on the proximity-aware reliable overlay network can achieve high reliability (>0.8) with low latency (<3000ms) under a wide range of substantial geographically correlated failures. Our future work will exploit this approach to design more efficient techniques for data dissemination; we will explore the ability to predict failure propagation patterns and use this prediction to design warning mechanisms that are cognizant of the propagating failures.

REFERENCES

[1] Inet : heep/topology.eecs.berkeley.edu/inet/
[7] Stefan Rintner and Fabian Bustamante, "Resilient Peer-to-Peer Multicast without the Cost," in Proceedings of MCM’05 2005