An Overlay Network Construction Technique for Minimizing the Impact of Physical Network Disruption in Cloud Storage Systems

Citation:

Katsuya Suto, Hiroki Nishiyama, Nei Kato, Takayuki Nakachi, Tatsuya Fujii, and Atsushi Takahara, "An Overlay Network Construction Technique for Minimizing the Impact of Physical Network Disruption in Cloud Storage Systems," International Conference on Computing, Networking and Communications (ICNC 2014), Honolulu, Hawaii, USA, Feb. 2014.

URL:

http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6785307

An Overlay Network Construction Technique for Minimizing the Impact of Physical Network Disruption in Cloud Storage Systems

Katsuya Suto[§], Hiroki Nishiyama[§], Nei Kato[§], Takayuki Nakachi[†], Tatsuya Fujii[†], and Atsushi Takahara[†]
[§]Graduate School of Information Sciences, Tohoku University, Sendai, Japan
[†]NTT Network Innovation Laboratories, NTT Corporation, Yokosuka, Japan
E-mails: [§]{suto, bigtree, kato}@it.ecei.tohoku.ac.jp, [†]{nakachi.takayuki, fujii.tatsuya, takahara.atsushi}@lab.ntt.co.jp

***Abstract*—Cloud storage exploiting overlay networks is considered to be a scalable and autonomous architecture. While this technology can ensure the security of storage service, it requires addressing the "server breakdown" problem, which may arise due to malicious attacks on servers and mechanical troubles of servers. In existing literature, an overlay network based on bimodal degree distribution was proposed to achieve high connectivity to combat these two types of server breakdown. However, it cannot ensure the high connectivity against physical network disruption that removes numerous nodes from overlay network. To deal with this issue, in this paper, we propose a physical network aware overlay network, in which the neighboring nodes are connected with one another in the overlay. Moreover, the numerical analysis indicates that the proposed system considerably outperforms the conventional system in terms of service availability.**

I. Introduction

Recently, cloud storage, which provides storage service to users through the Internet, is evolving as a dominant choice for data storage for the common users [1]. However, since both non-malicious and malicious users access the service, a cloud storage with high security level is required [2]. As a consequence, cloud storage services commonly use a distributed architecture, in which a data file is distributed to distinct servers by using erasure coding in order to protect the stored data from the prying eyes of the malicious users. However, this architecture complicates data management, and it becomes difficult to search the data with the increase of data and servers. In addition, since a few central servers manage critical information such as the location of the stored data, malicious attackers may easily find the data. Due to these reasons, a scalable and distributed management architecture will be a significantly important requirement for future cloud storage services.

A cloud storage utilizing overlay networks is able to provide scalable and distributed management. In this architecture, a node in the overlay network has a route information of its neighboring nodes, and the desired data can be found by using a query flooding scheme (e.g., unstructured peer-to-peer (P2P)). However, when the nodes are removed due to server breakdown, the overlay network is unable to find data due to network disruption. In literature, there exists numerous works which address this issue [3]–[5]. The work in [3] proposed an overlay network tolerant to mechanical trouble of servers, i.e., a node is supposed to be randomly removed regardless of the degree of the node. According to the work, while the network following power-law degree distribution achieves the highest connectivity against mechanical troubles, there still remains the issue related to vulnerability of Distributed Denial of Service (DDoS) attacks on servers, i.e., a node with higher degree is supposed to be removed with a relatively higher probability. In order to overcome this shortcoming, the works in [4], [5] proposed an overlay network following the bimodal degree distribution, where there are only two types of nodes, namely, super nodes (SNs) with higher degree and leaf nodes (LNs) with lower degree. This network, dubbed THUP (churn/DoS Tolerant, Hierarchical, Unstructured, P2P network), optimizes connectivity against both mechanical troubles and DDoS attacks by controlling number and degree of each node.

While these works addressed only server breakdown issue, this paper takes into account the physical network disruption. The Physical network disruption, which is caused by router/communication line breakdown due to mechanical trouble or DDoS attacks, leads to the removal of numerous nodes from the overlay network. In order to improve the connectivity of THUP, in this paper, we propose a neighbor selection scheme by considering a physical network comprising neighboring nodes which completely connect with each other. Our extension to the original THUP with our proposed neighbor selection scheme is dubbed as "THUP Plus".

The remainder of this paper is organized as follows. The relevant research works on the overlay networks based on bimodal degree distribution are discussed in Section II. In Section III, we present our envisioned neighbor selection scheme. Section IV analyses the network connectivity of both THUP and THUP Plus to evaluate the effectiveness of THUP Plus. Finally, concluding remarks are provided in Section V.

II. Overlay Networks based on Bimodal Degree Distribution

In order to achieve high connectivity, a network following the bimodal degree distribution has been studied [6]–[8]. These works analyzed each degree of the high and low pole, k_{s} and k_{l}, and the node ratio r for optimizing tolerance to

server breakdowns. An optimal bimodal degree distribution is expressed by the following equation.

$$P(k) = \begin{cases} N_{\rm s} = rN, & \text{if } k = k_{\rm s} = \sqrt{\langle k \rangle N}, \\ N_{\rm l} = (1-r)N, & \text{if } k = k_{\rm l} = \langle k \rangle, \\ 0, & \text{otherwise}, \end{cases} \quad (1)$$

$$r = \left(\frac{A^2}{\langle k \rangle N} \right)^{\frac{3}{4}},$$

$$A = \left\{ \frac{2\langle k \rangle^2 (\langle k \rangle - 1)^2}{2\langle k \rangle - 1} \right\}^{\frac{1}{3}},$$

where, N, $N_{\rm s}$, $N_{\rm l}$, and $\langle k \rangle$ denote the number of nodes, SNs, LNs, and the average degree of the overlay network, respectively.

On the other hand, an optimal topology of overlay network based on the bimodal degree distribution, dubbed THUP, was proposed and studied in literature [4], [5]. Fig. 1 shows the network topology of THUP. Each SN having higher degree, $k_{\rm s}$, connects with all other SNs to construct a complete graph since SNs transmit a number of messages to other SNs. In addition, each SN connects with LNs having lower degree, $k_{\rm l}$, until the degree of each SN is satisfied. Due to the limitation of the degree of SNs, the LNs are classified into Normal Leaf Nodes (NLNs) which connect with the SN, and Extra Leaf Nodes (ELNs) which do not connect with the SN but connect with the diagonally-cornered ELN. LNs connect with other LNs to construct a ring topology, where the NLNs connecting with the same SN become neighboring nodes and the ELNs are evenly located. We define a group that consists of a SN, NLNs connecting with the SN, and ELNs connecting with the NLNs, where each group has the same number of nodes.

Although THUP achieves the maximum connectivity against the server breakdowns, it is easily disrupted due to physical network disruption because numerous nodes which are connected with victim routers are removed. A malicious attacker can easily deny the storage services by using DDoS attack on the victim route. Hence, a method to improve the connectivity against the physical network disruption is required for realizing higher available cloud storage service.

III. Neighbor Selection Scheme Tolerant to Physical Network Disruption

In this section, we propose a scheme to select the neighboring nodes in order to increase the connectivity against the physical network disruptions. First, we explain an overview of our proposed neighbor selection scheme. Then, we propose a node joining algorithm which realizes the proposed scheme.

A. Overview of physical network aware neighbor selection

Physical network disruption causes the following phenomena. First, all servers connected to the targeted (i.e., victim) router are unable to communicate with the other servers. Second, these affected nodes (via which the servers are connected with the victim routes in the physical network) are removed from the overlay network. In addition, if the overlay network is constructed through random neighbor selection without consideration of the physical network, the physical network disruption randomly removes these nodes from the overlay network. Since a lot of nodes lose their links, some nodes may be isolated from the overlay network (i.e., the overlay network is disrupted). Indeed, THUP, is also disrupted by such physical network disruptions because it is constructed through random neighbor selection.

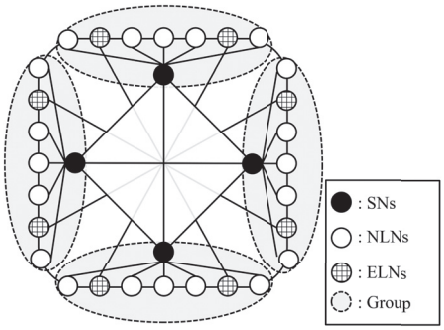


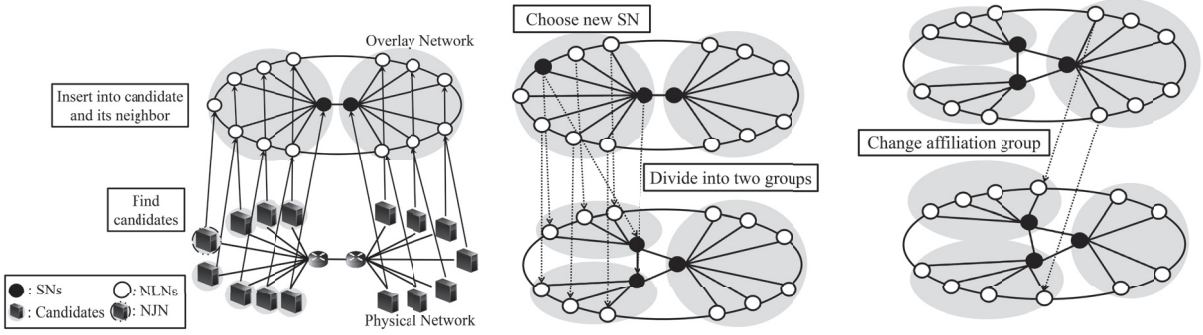
Fig. 1. Topology of THUP when $\langle k \rangle = 3$.

To deal with the above mentioned phenomena caused by physical network disruptions, our proposed scheme constructs the overlay network by considering the physical network, i.e., the servers connected to the same router in the physical network become neighboring nodes in overlay network. In the network constructed by the scheme, the physical network disruption causes the removal of a cluster which consists of many nodes. In other word, a few surviving nodes lose their links. Therefore, this scheme can minimize the decrease of connectivity against physical network disruption. It is expected that THUP can also improve the connectivity by using the proposed neighbor selection method. Our extension to the original THUP with our proposed neighbor selection scheme is dubbed as "THUP Plus".

B. Node joining algorithm in THUP Plus

We propose a node joining algorithm which performs the physical aware neighbor selection. The algorithm is autonomously executed in a newly joining node (NJN). First, an NJN receives an inter-group list from an SN. The inter-group list shared between SNs contains the following information: the average degree of the network, the SN's addresses, and the number of nodes in each group. The list is updated when node information in affiliation group is changed (i.e., node joining and node departure). Then, the NJN executes a neighbor selection process, which is described in the following.

Neighbor selection process – This process aims to connect with appropriate LNs, which connect with same router (or link) in physical network as shown in Fig. 2(a). First, the NJN searches appropriate group, in which the SN has the minimum hop count from the NJN on physical network, by using PING-PONG mechanism, where the NJN can know the IP address of SNs from the inter-group list. After finding the affiliation group, the NJN receives intra-group list shared within the group. The list contains the LN's addresses and degree in the group. After receiving the intra-group list, the NJN executes the PING-PONG mechanism against the LNs found in the



(a) Neighbor selection process (b) Group construction process (c) Regrouping process

Fig. 2. The main processes of node joining algorithm in proposed neighbor selection.

intra-group list in order to find adequate neighboring nodes which have the minimum hop count from the NJN on physical network. When the NJN finds some candidates, it randomly selects one of the candidates. Finally, the NJN connects with the selected candidate and its neighbor, and removes the link between the selected candidate and its neighbor.

After the neighbor selection process, the type of the NJN is decided, i.e., NLN or ELN. If the degree of the SN in the affiliation group is not fully filled, the NJN connects with the SN and becomes NLN. Otherwise, the NJN establishes links to ELN as long as the degree of the node is lower than the average degree or the ELNs which do not satisfy the degree in the network. In order to connect with ELNs, the SN in the affiliation group finds the candidate instead of the NJN, and informs the address of the candidate to the NJN. The NJN attempts to establish link to the candidate. Then, the NJN updates the intra-group with addition of self-information, and transmits it to all nodes in the group. Following this, the SN updates the inter-group list (i.e., the number of nodes is incremented), and transmits it to other SNs. Then, the SN calculates the ideal number of SNs, $N_{\rm s}'$, by using Eq. (1) with the average degree and the number of nodes found in the inter-group list. When the current number of SNs, $N_{\rm s}$, is lower than the ideal number of SNs, the following process is executed.

Group construction process – The objective of this process is to construct new group to scale with the increase of nodes as shown in Fig. 2(b). First, the NJN selects a node from the biggest group as a new SN. The degree of the SN is calculated according to Eq. (1). Then, the SN attempts to construct a new group. That is, the SN divides the affiliation group into two groups which have same number of nodes.

The construction of group and participation of NJN change each group size. Therefore, the NJN calculates the range between the biggest group size and the smallest group size, δ. If each group size becomes disproportionate, namely, $\epsilon < \delta$, the NJN executes a regrouping process, where ϵ indicates the disproportion threshold. The network manager (or designer) configures the value of ϵ, arbitrarily.

Regrouping process – This process restructures the group to equalize each group size as shown in Fig. 2(c). First, the NJN calculates the average group size, $G_{\rm ave}$. The smaller size group gets the nodes from neighboring groups to approximate size of the group to $G_{\rm ave}$, where the nodes which change the group are chosen in order from a near node. They recreate the link to SN and receive the intra-group list in the new affiliation group. This procedure continues until all group sizes approximate $G_{\rm ave}$.

Algorithm 1 Node joining algorithm in THUP Plus.

```
Get an inter-group list
Neighbor selection process
if Degree of SN is not filled then
    Connect with SN
end if
while Degree of NJN is not filed & candidate exists do
    Connect with ELNs
end while
Update the intra-group list
if N_s < N_s' then
    Group construction process
end if
if ε < δ then
    Regrouping process
end if
```

IV. Analysis and Numerical Results

In this section, we evaluate the connectivity of THUP and THUP Plus by using complex network theory [6]–[8]. We analyze three values, CAR, CDO, and CAO, which quantify how many nodes can be removed from overlay network without network disruption under following three situations, respectively: (i) when the physical network disruption occurs, (ii) when the mechanical troubles of servers occur after physical network disruption, and (iii) when servers suffer DDoS attacks after physical network disruption.

A. Connectivity analysis in THUP

At first, we derive the value of $CAR_{\rm THUP}$. Since the physical network disruption randomly removes nodes in case of THUP, the value of $CAR_{\rm THUP}$ can be formulated as

follows by using percolation threshold in [6].

$$CAR_{\mathrm{THUP}} = \left(1 - \frac{1}{\overline{P(k^2)}/\overline{P(k)} - 1}\right)N, \tag{2}$$

where $\overline{P(k)}$ denotes the average degree of the degree distribution $P(k)$.

Then, we investigate the values of CDO_{THUP} and CAO_{THUP}. Since the DDoS attacks and mechanical troubles are caused after physical network disruption, the degree distribution after physical network disruption is required to conduct CDO_{THUP} and CAO_{THUP}. Supposed $\boldsymbol{S}$ and $\boldsymbol{R}$ are the set of surviving nodes and removed nodes, respectively, the interconnection links between the set $\boldsymbol{S}$ and $\boldsymbol{R}$ are removed by node removal model f_k. Since the physical network disruption randomly removes nodes, the node removal model, f_k, is given by following equation.

$$f_{\mathrm{k}} = \frac{N_{\boldsymbol{R}}}{N}, \tag{3}$$

where the $N_{\boldsymbol{R}}$ is number of nodes in set $\boldsymbol{R}$ (i.e., number of removed nodes). The degree distribution of THUP after physical network disruption, $P_\alpha(k)$, can be formulated as follows by using the node removal model f_k .

$$P_\alpha(k) = \begin{cases} \displaystyle\sum_{i=0} f_i^i P_{\boldsymbol{S}}(i), & \text{if } \ k = 0, \\ \displaystyle\sum_{i=k} \binom{i}{k} f_i^{i-k}(1-f_i)^k P_{\boldsymbol{S}}(i), & \text{otherwise}, \end{cases} \tag{4}$$

where $P_{\boldsymbol{S}}(i)$ is degree distribution in set $\boldsymbol{S}$ before the links between the set $\boldsymbol{S}$ and $\boldsymbol{R}$ are removed. Thus it is decided as follows.

$$P_{\boldsymbol{S}}(i) = \frac{(1-f_i)P(i)}{1 - \sum_j P(j) f_j}. \tag{5}$$

Since the mechanical troubles randomly remove the nodes, the connectivity against mechanical troubles after physical network disruption, CDO_{THUP}, can be formulated as follows.

$$CDO_{\mathrm{THUP}} = \left(1 - \frac{1}{\overline{P_\alpha(k^2)}/\overline{P_\alpha(k)} - 1}\right)(N - N_{\boldsymbol{R}}). \tag{6}$$

Moreover, by solving a simultaneous equation in [6], [8], we can derive the connectivity against DDoS attacks on servers after physical network disruption, CAO_{THUP}.

$$CAO_{\mathrm{THUP}} = \left(p\overline{P_\alpha(k)} - \sum_{k=\widetilde{K}}^{K}(k-1)P_\alpha(k)\right)(N - N_{\boldsymbol{R}}), \tag{7}$$

where K and $\widetilde{K}$ denotes the maximum degree before the node removal and the maximum degree after the node removal, respectively. The value of p is given as follows.

$$p = 1 - \frac{1}{\overline{P_\alpha(\widetilde{k}^2)}/\overline{P_\alpha(\widetilde{k})} - 1},$$

where $P_\alpha(\widetilde{k})$ is the average degree of degree distribution $P_\alpha(\widetilde{k})$ from 0 to $\widetilde{k}$ before node removal.

B. Connectivity analysis in THUP Plus

In the case of THUP Plus, since servers connecting with the same router become neighboring nodes, a cluster is removed from the overlay network due to physical network disruption. In other word, THUP Plus does not lead to network disruption. Hence, the value of CAR_{PRO} can be formulated as follows.

$$CAR_{\mathrm{PRO}} = N. \tag{8}$$

Let $N^{\boldsymbol{S}}$ and $N^{\boldsymbol{R}}$ be the number of nodes in the sets the $\boldsymbol{S}$ and $\boldsymbol{R}$, respectively. Considering the cluster as a virtual nodes with degree k_{v}, the node removal model of physical network disruption, f'_k, is given by following equation.

$$f'_k = \begin{cases} 1, & \text{if } \ k = k_{\mathrm{v}}, \\ 0, & \text{otherwise}. \end{cases} \tag{9}$$

The virtual node, which is composed removed nodes in the set $\boldsymbol{R}$, has external links to the nodes in the set $\boldsymbol{S}$. There are four kinds of external links as follows: the links between removed SNs and surviving SNs, that between removed SNs and surviving NLNs, that between removed ELNs and surviving ELNs, and that between removed NLNs and surviving NLNs that constructs ring topology. Therefore, the degree of virtual node, k_{v}, is decided by

$$k_{\mathrm{v}} = N_{\mathrm{s}}^{\boldsymbol{S}} N_{\mathrm{s}}^{\boldsymbol{R}} + \left[N_{\mathrm{s}}^{\boldsymbol{R}}\{k_{\mathrm{s}} - (N_{\mathrm{s}} - 1)\} - N_{\mathrm{nl}}^{\boldsymbol{R}}\right] + N_{\mathrm{el}}^{\boldsymbol{R}} + 2,$$

where $N_{\mathrm{s}}^{\boldsymbol{S}}$, $N_{\mathrm{s}}^{\boldsymbol{R}}$, $N_{\mathrm{nl}}^{\boldsymbol{R}}$, and $N_{\mathrm{el}}^{\boldsymbol{R}}$ indicate the number of surviving SNs, removed SNs, removed NLNs, and removed ELNs, respectively. The degree distribution of the virtualized overlay network, $\tilde{p_k}$, is given by following equations.

$$\tilde{P}(k) = \begin{cases} 1/\tilde{N}, & \text{if } \ k = k_{\mathrm{v}}, \\ N_{\mathrm{s}}^{\boldsymbol{S}}/\tilde{N}, & \text{if } \ k = k_{\mathrm{s}}, \\ N_{\mathrm{l}}^{\boldsymbol{S}}/\tilde{N}, & \text{if } \ k = k_{\mathrm{l}}, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

where $\tilde{N} = N - N^{\boldsymbol{R}} + 1$ indicates the number of nodes in the virtualized overlay network. Thus, the degree distribution of THUP Plus after physical network disruption, $P_\beta(k)$, can be formulated as follows by using the node removal model f'_k.

$$P_\beta(k) = \frac{(1-f'_k)\tilde{P}(k)}{1 - \sum_j \tilde{P}(j) f'_j}. \tag{11}$$

As in the case in THUP, in THUP Plus, the connectivity against mechanical troubles after physical network disruption, CDO_{PRO}, can be formulated as follows.

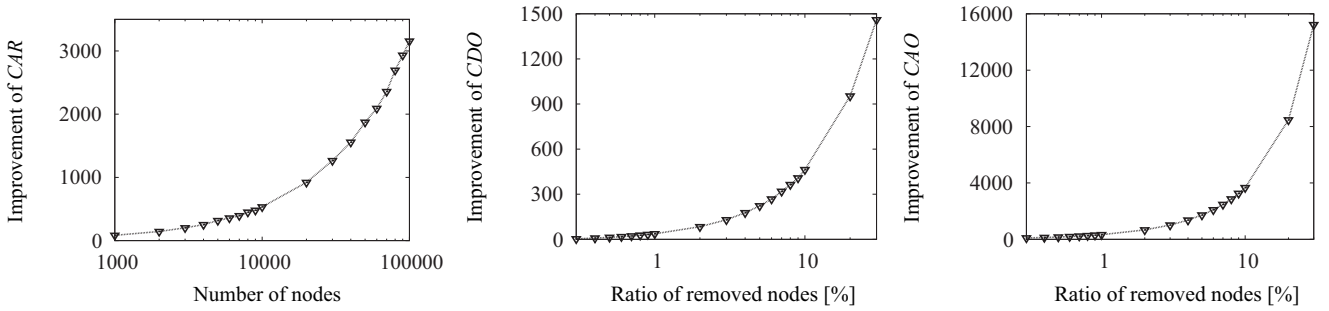
$$CDO_{\mathrm{PRO}} = \left(1 - \frac{1}{\overline{P_\beta(k^2)}/\overline{P_\beta(k)} - 1}\right)(N - N_{\boldsymbol{R}}). \tag{12}$$

In addition, the connectivity against DDoS attacks on servers after physical network disruption, CAO_{PRO}, can be formulated as follows.

$$CAO_{\mathrm{PRO}} = \left(p\overline{P_\beta(k)} - \sum_{k=\widetilde{K}}^{k_1}(k-1)P_\beta(k)\right)(N - N_{\boldsymbol{R}}), \tag{13}$$

where,

$$p = 1 - \frac{1}{\overline{P_\beta(\widetilde{k}^2)}/\overline{P_\beta(\widetilde{k})} - 1}.$$



(a) Physical network disruption

(b) Mechanical troubles of servers after physical network disruption

(c) DDoS attacks on servers after physical network disruption

Fig. 3. Connectivity improvement in THUP Plus.

C. Results and discussions

In the remainder of this section, we verify the performance of THUP Plus through numerical calculation. The improvement of the connectivity is defined as the delta between the connectivity of THUP Plus and THUP. At first, we evaluate the connectivity against the simple failure model whereby physical network disruption occurs. Then, we evaluate the improvement of the connectivity against the integrated network failure model, which considers physical network disruption, mechanical troubles of servers, and DDoS attacks to servers.

Fig. 3(a) demonstrates the change of the connectivity improvement against the physical network disruption when the number of nodes varies from 10^3 to 10^5 and the average degree is 3. From Fig. 3(a), it is clearly shown that THUP Plus achieves a higher value of CAR than THUP regardless of the number of nodes. This is because of the fact that our proposed neighbor selection method can decrease the number of nodes that lose their links.

Fig. 3(b) demonstrates the impact of the ratio of removed nodes due to a physical network disruption on the connectivity against mechanical troubles of servers after physical network disruption, when the number of nodes and the average degrees are 10^5 and 3, respectively. From Fig. 3(b), it is understood that our proposed THUP Plus improves the connectivity regardless of the ratio of the removed nodes. Fig. 3(c) demonstrates the connectivity against DDoS attacks on servers after physical network disruption with the same parameter as in Fig. 3(b). From Fig. 3(c), we can confirm the improvement of connectivity of THUP Plus regardless of the ratio of the removed nodes. Compared with Fig. 3(b), we can notice that THUP Plus is more effective against DDoS attacks on servers since the improvement of CAO is higher than that of CDO. Thus, we conclude that the proposed THUP Plus features as the most suitable network in terms of service availability.

V. Conclusion

In this paper, we proposed THUP Plus, which achieves a significantly high connectivity against physical network disruption, for future cloud storage utilizing overlay networks. In case of utilizing such overlay network, the network connectivity is one of the most important factors to provide high availability of cloud storage service. However, the existing schemes including THUP cannot ensure the connectivity against the physical network disruption even though they achieve high connectivity against server breakdowns. This is because of the fact that the scheme constructs the overlay network without considering the physical network information. Thus, in our proposed THUP Plus, each node selects the neighboring nodes based on the physical network information. From the numerical results, we confirmed that THUP Plus achieves a substantially higher performance in contrast with THUP.

Acknowledgement

Part of this work was conducted under the national project, “Research and development of “Movable ICT-Units” for emergency transportation into disaster-affected areas and multi-unit connection”, supported by the Ministry of Internal Affairs and Communications (MIC), Japan.

References

[1] C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou, “Toward Secure and Dependable Storage Services in Cloud Computing,” *IEEE Transactions on Services Computing,* Vol. 5, No. 2, pp. 220-232, Apr.-Jun. 2012.

[2] J Harauz, L. M. Kaufman, and B. Potter, “Data Security in the World of Cloud Computing,” *IEEE Security & Privacy*, Vol. 7, No. 4, pp. 61-64, July-Aug. 2009.

[3] D. Stutzbach, R. Rejaie, and S. Sen, “Characterizing unstructured overlay topologies in modern p2p file-sharing systems,” *IEEE/ACM Transactions on Networking*, Vol. 16, No. 2, pp. 267-280, Apr. 2008.

[4] K. Suto, H. Nishiyama, S. Shen, and Nei Kato, “Designing P2P Networks Tolerant to Attacks and Faults Based on Bimodal Degree Distribution,” *Journal of Communication*, vol. 7, no. 8, pp.587-595, Aug. 2012.

[5] K. Suto, H. Nishiyama, N. Kato, T. Nakachi, T. Fujii, and Atsushi Takahara, “THUP: A P2P Network Robust to Churn and DoS Attack based on Bimodal Degree Distribution,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 247-256, Sep. 2013.

[6] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H. E. Stanley, “Optimization of network robustness to waves of targeted and random attacks,” *Physical Review E*, Vol. 71, No. 4, 4 pages, Apr. 2005.

[7] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H. E. Stanley, “Optimization of the robustness of multimodal networks,” *Physical Review E*, Vol. 74, 8 pages, July 2006.

[8] B. Mitra, F. Peruani, S. Ghose, and N. Ganguly, “Analyzing the vulnerability of superpeer networks against attack,” in *Proc. of the 14th ACM conference on Computer and communications security*, Alexandria, Virginia, USA, Oct.-Nov. 2007, pp. 225?234.