www.ietdl.org

Published in IET Communications
Received on 24th August 2013
Revised on 14th May 2014
Accepted on 28th May 2014
doi: 10.1049/iet-com.2013.0709

ISSN 1751-8628

Maintenance of cooperative overlays in multi-overlay networks

Wu-Chun Chung[1], Chin-Jung Hsu[2], Kuan-Chou Lai[3], Kuan-Ching Li[4], Yeh-Ching Chung[1]
[1]Department of Computer Science, National Tsing Hua University, Hsinchu 30013, Taiwan
[2]Department of Computer Science, North Carolina State University, Raleigh, NC 27695, USA
[3]Department of Computer Science, National Taichung University, Taichung 40306, Taiwan
[4]Department of Computer Science and Information Engineering, Providence University, Taichung 43301, Taiwan
E-mail: wcchung@sslab.cs.nthu.edu.tw

Abstract: In overlay-based applications, multiple overlay networks are deployed to fulfill different service requirements. A multi-overlay environment may exist in which a number of nodes simultaneously participate in the networks. When there are multiple overlay-based applications running over a set of nodes, some of the nodes take extra effort to maintain multi-overlay networks. Therefore, maintaining these co-existing overlays incurs redundant maintenance overhead. This research presents a cooperative strategy for exploiting a master–slave model to handle the common overlay-maintenance. The purpose is to eliminate the redundant maintenance overhead. To evaluate system performance, this study not only analyses various combinations of multiple overlays but also considers the effectiveness of the master selection approach. Experimental results demonstrated that the proposed cooperative strategy significantly decreases the redundant overlay-maintenance overhead. In some cases, the overall reduction ratio of maintaining multiple overlays is as high as 60%.

1 Introduction

The overlay network is a virtual network built upon the underlying computer network. Techniques of overlay networks are widely employed in distributed systems. With the explosion of overlay-based applications, nodes may simultaneously participate in multiple overlays to serve a variety of distributed applications. Therefore, multiple overlays co-exist in a distributed computing environment. In such a multi-overlay environment (MOE), maintaining the co-exiting overlays needs a considerable amount of effort. For instance, the failure-detection operation [1] is commonly adopted for every overlay to ensure the overlay resilience. The network-proximity estimation operation is also widely applied [2–5] to enable the locality-aware overlay. The overall maintenance cost is increased because of the incremental amount of redundant overhead when these co-existing overlays serve different applications. However, some redundant costs are able to be trimmed.

Several investigations related to MOEs have been studied on exploiting the synergy of multiple co-existing overlays. Maniymaran *et al.* [6] investigate the co-existence of Pastry and interest-based Gossip protocol. Another work of Lin *et al.* [7] estimate the potential benefits on different types of synergies among multiple co-existing overlays. However, their approaches are applied solely to specific overlays, mainly focused on Gossip-based protocols. In addition, those previous studies missed to consider a MOE in which different intersection ratios (IRs) of overlays may affect the system effectiveness. Our work not only takes the IR into consideration but also conducts the reduction of redundant overlay-maintenance operations in MOEs.

To tackle the problem of redundant overlay-maintenance in a MOE, a cooperative strategy is introduced to simplify the overall maintenance cost. The master–slave model and the selection criteria of a master overlay are also presented. Based on the proposed cooperative strategy, both cooperative failure-detection (CFD) and cooperative network-proximity estimation (CNPE) mechanisms are illustrated, respectively, to eliminate the redundant operations of failure-detection and network-proximity estimation. Expanding from our previous work [8], this paper further discusses the analysis of these proposed mechanisms, refined cost models of overlay maintenance and selection criteria of the master overlay. In order to evaluate the performance improvement, this work compares the maintenance cost of the original MOE with the one which adopts our proposed cooperative strategy. Experimental results show that the proposed strategy is efficient in diverse MOEs, where the reduction ratio is up to 60% in some cases.

Major contributions of our work include: (i) exploiting the synergy of multi-overlays using the proposed master–slave model, (ii) taking the IR of co-existing overlays into consideration, (iii) analysing the cooperative mechanisms of failure-detection and network-proximity estimation, (iv) characterising the selection criteria of a master overlay and finally, (v) evaluating the effectiveness of eliminating redundant operations in a comprehensive MOE.

The rest of this paper is organised as follows. Section 2 introduces the most relevant works on multiple overlays.

The cooperative strategy and the master–slave model are proposed in Section 3. Section 4 further demonstrates two types of overlay maintenance: CFD and CNPE. The experimental results are presented in Section 5 to illustrate the performance improvement. Finally, concluding remarks and future works are given in Section 6.

2 Related work

The literature of multi-overlay studies conducts the multi-overlay framework [9, 10], the race condition problem [11–14] and the maintenance cost reduction [6, 7]. The most relevant issue to our work is the problem of reducing the maintenance cost, which focuses on the cooperation of multiple overlays in a MOE.

When these overlays work independently, the maintenance of co-existing overlays brings extra costs. The related study [6] proposes a novel approach to leverage the co-existence of the interest-based Gossip protocol and Pastry by constructing a joint overlay. The authors have proved that the routing table in Pastry could be replaced by the cluster view routing table in the interest-based Gossip protocol. Meanwhile, the random peer sampling view routing table in the interest-base Gossip protocol could also be replaced by the leaf set table in Pastry. However, their approach only focuses on two specific overlays. The cooperation between co-existing overlays is also discussed in the work [7]. The authors argue that the synergy of overlays benefits system performance instead of causing a negative impact introduced by the overlay competence. That paper analyses different types of synergies and compares their potential benefits, but mainly focusing on Gossip protocols.

From previous investigations, a few studies have reported on a general approach to making co-existing overlays cooperation. Moreover, previous studies have not considered the network environment with different IRs of multiple overlays, which is a key factor affecting the performance of cooperation. Our work not only focuses on eliminating those redundant operations of overlay-maintenance but also takes the complex MOEs into consideration.

3 Cooperative overlay maintenance

Overlay networks can be typically classified into unstructured and structured adhered to the topology [15]. This study focuses on the unstructured overlay, the ring overlay and the tree overlay. The parameter K defines the lower bound of the size of neighbour tables. Nodes in an overlay network joining or leaving arbitrarily will lead to a dynamic environment. To represent the dynamic degree of an overlay network environment, the churn rate is defined as $R = I/N$ where I is the amount of nodes that join/leave per unit time and N is the network size. The churn rate can be converted into the median session time to symbolise the average lifetime of nodes [4]. These two terms are interchangeably used in this paper.

In a dynamic environment, nodes have to maintain their routing tables to ensure that the network works well. For example, in the Chord protocol, nodes need to update the finger table when nodes in that table fail. The quantities of nodes that join and leave are assumed the same in the next discussion. These overlay-maintenance operations will incur maintenance costs. Accordingly, the overlay-maintenance cost M of an overlay is defined as the overhead occurred by overall communication messages that have been sent by all nodes within the overlay.

3.1 Multi-overlay environment

Definition 1 (MOE): A multi-overlay environment is defined as $\text{MOE} = \{O_i\}$ and $|\text{MOE}| \geq 2$, where O_i represents an overlay i that co-exists over the same underlying computer network.

As the scale of MOE expands, the total cost of overlay-maintenance increases. To tackle this problem, this work tries to reduce the amount of overlay-maintenance costs by the cooperation of co-existing overlays. However, in a MOE, nodes may not participate in all overlays. This paper also exposes the IR to represent how many nodes simultaneously appear in two separate overlays.

Definition 2 (IR): In a MOE, the IR of any two overlays is defined as $\text{IR}_{O_iO_j} = |\{n | n \text{ nodes co-exist in both } O_i \text{ and } O_j\}|/N$, where $i \neq j$.

Since different co-existing overlays introduce different MOEs, our proposed strategy will be evaluated under various overlay combinations (OCs) to demonstrate its generality.

Definition 3 (OCs): Suppose that there are l multiple co-existing overlays in a $\text{MOE} = \{O_1, O_2, \ldots, O_l\}$, the OC produces $\binom{l}{2}$ MOEs. The set of OCs is defined as $\text{OC} = \{\{O_1\} \times \{O_2, O_3, \ldots, O_l\}, \{O_2\} \times \{O_1, O_3, \ldots, O_l\}, \ldots, \{O_l\} \times \{O_1, O_2, \ldots, O_{l-1}\}\}$.

The key to the reduction of overlay-maintenance operations is to determine which operations are redundant. This paper discusses two important operations of a MOE that are common and reducible. One is the failure detection and the other one is the network-proximity estimation.

The failure means a node may leave or disconnect the overlay without any notification. The overlay resilience relies on the periodical failure detection by monitoring the status of nodes. Zhuang *et al.* [1] have defined that the failure-detection operation can be active or passive. In the case of active failure-detection operations, each node periodically probes its neighbours by sending keep-alive messages for checking the state of neighbours. In a MOE, duplicate failure-detection operations of these overlays are redundant. Therefore the number of failure-detection operations could be reduced by common overlay-maintenance.

Many overlay networks consider network information to reduce the routing latency [2–5, 16]. According to the work [3], common approaches to exploit network-proximity include geographic layout, proximity routing and proximity neighbour selection. These approaches involve the network-proximity estimation operation. Therefore the estimating operation of network proximity is also reducible by common overlay-maintenance.

To deal with the common overlay-maintenance, the master–slave coordination model is proposed to eliminate those redundant operations in such a MOE. Next, the idea of the proposed model and corresponding cooperative mechanisms are illustrated.

3.2 Master–slave model

In the master–slave model, one of the co-existing overlays has to be the master overlay in charge of maintaining the common overlay-maintenance operations for other overlays, named slave overlays. With this model, the unnecessary maintenance operations can be eliminated in slave overlays. The reduction ratio of overlay-maintenance costs is adopted to assess the performance of the proposed approach.

Definition 4 (RR): The RR of maintenance costs is defined as

$$\mathrm{RR} = \left(\frac{M - M'}{M}\right) \times 100\% = \left(1 - \frac{M'}{M}\right) \times 100\% \qquad (1)$$

where M is the maintenance cost of the original MOE and M' is the maintenance cost of the same MOE by the master–slave model.

To support the synergy of the master and the slave overlays, the inter-overlay and intra-overlay protocols are introduced. The inter-overlay protocol defines how the slave overlays communicate with the master overlay, whereas the intra-overlay protocol is the mechanism that the master overlay follows to handle the common overlay-maintenance.

3.2.1 Inter-overlay protocol: The inter-overlay protocol defines the interaction mechanism between the master overlay and the slave overlay. Different overlay-maintenance operations are handled by different inter-overlay protocols. Fig. 1 depicts two inter-overlay protocols in a MOE, where each node participants at least two overlays. The subscription/notification protocol is used to handle event-driven overlay-maintenance, while the query/response protocol endows the master overlay with the ability to handle information-retrieval overlay-maintenance for slave overlays.

In a MOE, the node can eliminate the redundant overlay-maintenance overhead by these inter-overlay protocols. That is, the subscription/notification protocol is applied to the common failure detection and common operations of information exchange between nodes. On the other hand, the query/response protocol provides an interface for slave overlays to access the needed information from the master overlay. If the master and slave overlays maintain common information, for example, network-proximity or semantic-locality information, slave overlays communicates with the master overlay via the query/response protocol to retrieve the needed information. In this way, slave overlays do not maintain the needed information itself.

As shown in Fig. 1*a*, the master overlay works well that node a in O_1 issues a subscription/notification request to the node a in O_2. Occasionally, the master overlay cannot deal with the request from the slave overlay because of the non-existence of the same node in the master overlay. For example, node b in O_3 cannot exploit the inter-overlay protocol to communicate with node b in O_2 since node b does not exist in O_2. The existence of non-overlapped nodes commonly happens in a MOE; however, it is ignored in most of previous works.

3.2.2 Intra-overlay protocol: The intra-overlay protocol defines an established method that the master overlay deals with the request from slave overlays. As shown in Fig. 1*b*, node a in O_3 adopts the query/response protocol to retrieve the requisite information. Once node a in O_2 receives the request, it responds to the required information itself or gather information via the pre-defined routing protocol if node a in O_2 cannot satisfy the request.

The design of the intra-overlay protocol varies under different sorts of common overlay-maintenance. A general protocol to meet different requirements is hard to define. Nevertheless, a common idea to eliminate the redundant overlay-maintenance is to reduce the duplicated maintenance operations of the same node. For the event-driven overlay-maintenance, the duplicated maintenance costs are incurred by duplicated links. For example, node a in O_1 links to node b in O_1, and meanwhile, node a in O_2 also links to node b in O_2. When the duplicated links exist, the duplicated overlay-maintenance operations could be saved. For the information-retrieval overlay-maintenance, the duplicated information-retrieval operations can be executed only once by the information sharing.

3.3 Master-overlay selection criteria

Selecting the master overlay in a MOE involves three steps: firstly, sampling the overlay-pattern to characterise an overlay; secondly, enabling multi-overlays to elect the master overlay; and finally, informing overlays of the selection result. The basic principle of selecting a suitable master overlay is to pick an overlay which can handle the most common overlay-maintenance. Some important criteria are given as follows:

- *Maximum information.* A candidate for a master overlay should share useful information as much as possible. In general, the routing table is the source of such information. Therefore, the routing table could be commonly maintained if two overlays have similar maintenance policy. Previous works [6, 7] adopt the same idea to reduce the total

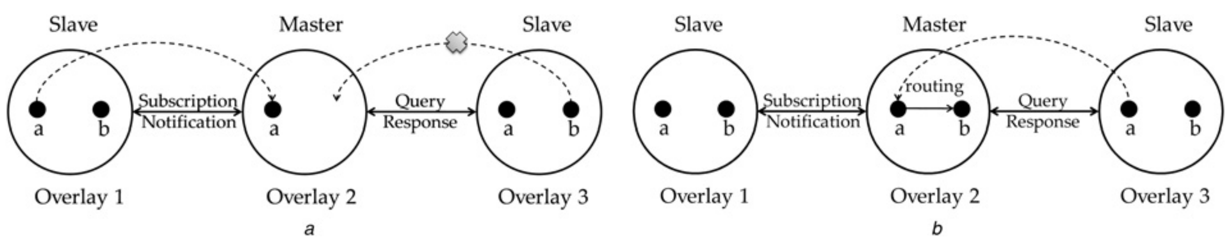

Fig. 1 *Two inter-overlay protocols in the master–slave model*

a Process of subscription/notification protocol
b Process of query/response protocol

overlay-maintenance costs. As a result, the overlay with a larger routing table can be selected as the master overlay.

- *Minimum requirement.* The master overlay handles the common overlay-maintenance operations for slave overlays. However, this cooperative strategy should not influence the original operations. Take the failure detection, for example, the master overlay ought to have lower detection duration. Suppose the detection durations of two overlays are 30 and 60s, respectively, the master overlay should be the one with the 30s detection duration. Otherwise, the master overlay would affect the original failure detection of another overlay.
- *Maximum IR.* The IR describes a number of the same nodes co-habiting in two overlays. In the case of MOE = $\{O_1, O_2, O_3\}$, it produces six combinations as well as six IRs by different master overlays. To gain a higher reduction ratio, the master overlay should have the maximum coverage of IRs. Given $\text{IR}(\{O_1\} \times \{O_2, O_3\}) = \max(\text{IR}(\{O_1\} \times \{O_2, O_3\}), \text{IR}(\{O_2\} \times \{O_1, O_3\}), \text{IR}(\{O_3\} \times \{O_1, O_2\}))$, O_1 is obliged to be the master overlay if it eliminates more redundant overlay-maintenance operations.
- *Minimum dynamics.* Since the dynamic status is not identical in each overlay, the selection of the master overlay should consider this factor. A relatively stable overlay network should be the master overlay to provide the qualitative service. Otherwise, it may need to pay large extra maintenance costs to handle such a condition.

Given these principles of selecting the master overlay, the next section will go through different criteria when the master–slave model is applied into different overlay-maintenance operations.

4 Cooperative mechanisms

To support the failure-detection and network-proximity estimation, two cooperative mechanisms are proposed based on the master–slave model. Detail mechanisms could be referred to our previous work [8]. This paper focuses on the key concept, refined cost model and selection criteria of the master overlay. Some notations to be used next are listed in Table 1.

Table 1 Notations

Symbol	Explanation
m, s	superscript indicating the master overlay or the slave overlay
n^{m}, n^{s}	node n in the master overlay or the node n in the slave overlay
$L^{m}_{n_i n_j}$, $L^{s}_{n_i n_j}$	link from n_i to n_j in the master overlay or the slave overlay
n_r	requester node in the slave overlay
n_h	handler node in the master overlay handling the request from n_r
n_t	target node in the master overlay that n_r is interested in
n_c	cooperator node in the master overlay that helps n_h handle the request from n_r
msg_i	message produced by n_h for informing n_t about the request from n_r
msg_f	message produced by n_t for forwarding msg_i to n_c
msg_n	message produced by n_c for notifying n_h about the failed n_c
msg_e	message produced by n_h to explore more answers for the request from n_r

4.1 Cooperative failure detection (CFD)

The cost reduction in failure detection comes from two main schemes: the elimination and the cooperation. The elimination approach is inspired by exploiting the duplicated link between two nodes which exist in the master and the slave overlays. Once a duplicated link exists, the detection operation does not repeatedly probe the status of the same neighbour. The detection operation is only executed once in the master overlay. Nevertheless, the elimination approach helps with reducing only a part of redundant failure-detection operations. That is because the link in a slave overlay might not exist in the master overlay.

To cope with such a condition, the cooperation approach is proposed. According to each node is monitored by some other nodes, the proposed approach is to find a cooperator node undertaking the redundant failure-detection operation. Hence, the master overlay can assist the slave overlay in handling failure-detection operations. However, this approach may incur extra overhead in the master overlay. Detecting failed or removed $L^{m}_{n_t n_c}$ may also generate many msg_i if a master overlay is responsible for handing a large amount of subscription requests from multiple slave overlays. An aggregator method is further adopted to cushion the impact on CFD performance. The proposed method aggregates multiple msg_f into a composite 'forward message'. Experimental results show that this method improves the reduction ratio of more than 20% in some of cases.

4.1.1 Maintenance cost model: The maintenance cost model of CFD captures the number of communication messages, that is, msg_i, msg_f and msg_n, generated by the subscription/notification protocol. The reduction ratio of CFD is derived by calculating the number of saved messages for the communication.

The first factor, msg_i, is sent from n_h to n_t. The message is generated in two cases. In one case, n_h receives a subscription request from n_r, but n_h does not monitor the status of n_t. In the other case, n_h does not monitor n_t anymore, and then, n_h sends msg_i to n_t in order to transfer the monitoring task. The total amount of incurred msg_i is represented as

$$C(\text{msg}_\text{i}) = L_\text{c} \times P_\text{s} + L_\text{d} \tag{2}$$

where L_c is the quantity of created links in slave overlays, P_s is the probability of non-existence of $L^{m}_{n_h n_t}$ and L_d is the number of links, for example, $L^{m}_{n_h n_t}$, changing from connected status to disconnected one. $L_c \times P_s$ represents the number of msg_i generated in the first case and L_d represents that created in the second case.

The second overhead is msg_f which is sent from n_t to n_c. The amount of msg_f is calculated by

$$C(\text{msg}_\text{f}) = C(\text{msg}_\text{i}) \times \text{NC} + L_\text{d} + L_\text{f} \times R \tag{3}$$

where NC is the number of cooperators, L_f is the amount of $L^{m}_{n_t n_c}$ that has been detected as failed status and R represents the amount of re-sent msg_f per detected failure of n_c. When n_t receives msg_i, n_t forwards this message to some of n_c. These type of messages amount to $C(\text{msg}_i) \times \text{NC}$. In another case, once n_t is not a neighbour of n_c, n_t forwards msg_i to a new n_c, which generates L_d messages. When n_c fails, n_t also sends msg_i to new cooperator nodes. These messages total up to $L_f \times R$.

The third cost is msg_n which is produced when n_c detects failed n_t and notifies the related n_h. The amount of msg_n is calculated by

$$C(\mathrm{msg_n}) = L_f \times A \tag{4}$$

where L_f is similarly used to represent the amount of failed $L^m_{n_c n_t}$ that have been detected by n_c and A is the number of msg_n created when one failed link, from n_c to n_t, is detected. Thus, $L_f \times A$ is the total amount of msg_n.

To sum up, the number of communication messages generated by CFD is defined as

$$C_{\mathrm{CFD}} = C(\mathrm{msg_i}) + C(\mathrm{msg_f}) + C(\mathrm{msg_n}) \tag{5}$$

and the total maintenance cost of failure detection in a MOE without CFD is modelled as

$$M = \sum_{O \in \mathrm{MOE}} F^{\mathrm{O}} \tag{6}$$

where F^{O} is the sum of detection messages generated in the overlay O. On the other hand, the maintenance cost of a MOE with CFD is defined as

$$M'_{\mathrm{CFD}} = F^{\mathrm{m}} + C_{\mathrm{CFD}} \tag{7}$$

where F^{m} is the number of detection messages created in the master overlay. Accordingly, the reduction ratio (RR) of CFD is calculated as

$$\mathrm{RR} = \left(1 - \frac{M'_{\mathrm{CFD}}}{M}\right) \times 100\% \tag{8}$$

4.1.2 Selection criteria and approach: The maintenance cost model shows that the link and the node dynamics are two important factors affecting the RR. The link dynamics mainly relate to the overlay type, which refers to L_d in (2) and (3). For instance, $L^m_{n_t n_c}$ may be removed from the routing table of n_t. In such a case, n_t forwards $\mathrm{msg_f}$ to new n_c. If this event happens frequently, the amount of $\mathrm{msg_f}$ grows dramatically. In many DHT-based overlays, the routing table would be often maintained in a highly dynamic environment. When the master overlay is a DHT-based overlay, the extra cost of $\mathrm{msg_f}$ is large, even when msg_i should be re-sent because of the removal of $L^m_{n_h n_t}$. This case involves at least $1 + \mathrm{NC}$ messages. In addition, the node dynamics makes those efforts of handling the subscription request useless. That is because no sooner had n_h handled a request from n_r than n_h failed. In this case, msg_i, $\mathrm{msg_f}$ and even msg_n are wasted because of the failure of n_r. Hence, node dynamics also affects the reduction ratio of CFD.

Another important factor related to the reduction ratio is the size of the routing table that a node constructs in a master overlay. The larger the routing table is, the higher possibility of handing the common overlay-maintenance will be. An extreme case is that the master overlay constructs a full-connected topology, that is, the parameter K equals N–1. In such a case, the master overlay handles those requests from slave overlays without additional communication messages. Consequently, the size of a routing table may have impact on the reduction ratio of CFD.

In view of this, the important issues in selecting a master overlay are 3-fold: the link dynamics, the node dynamics and the routing table size. Hence, the score-based method of CFD [17] is considered to select the master overlay from a MOE for a better RR. The score function is modelled as

$$s = \left(1 - \frac{K^{\mathrm{m}} \times \left(1 + Z^{\mathrm{m}} + Z^{\mathrm{s}}\right)}{K^{\mathrm{m}} + K^{\mathrm{s}}}\right) \times 100\% \tag{9}$$

where K is the size of the routing table, Z is the degree of the extra cost introduced by the CFD mechanism and the superscript m and s indicate the master and the slave overlays, respectively. This virtual function approximates, in a comparative way, the maintenance cost of a MOE, which takes different combinations of the master and slave overlays into consideration. When the cost is high, the reduction ratio is low as well as the score.

4.2 Cooperative network-proximity estimation (CNPE)

The concept of CNPE in a MOE is to share the network-proximity information in the master overlay. As a result, slave overlays are free to estimate the network-proximity. CNPE develops two key schemes to reduce the redundant network-proximity estimation operations. One is the elimination and the other is the exploration.

The elimination idea is similar to that scheme in CFD. If the requisite network information is maintained by multiple overlays, this information can be shared among these overlays. Hence, when one node in the slave overlay needs to execute the network-proximity estimation operations for reaching a set of nodes that are geographically closed to it, the node could ask the same node in the master overlay via the query/response protocol. This query interface provides a chance to eliminate a number of estimation operations.

Although the elimination approach is useful, the master overlay cannot handle the request from the slave overlays in some situations. That is because the number of nodes requested by the slave overlay may exceed that provisioned by the master overlay. Therefore the intra-overlay protocol tries to solve the problem. A node can explore the required information within the master overlay when the elimination scheme cannot completely handle the query request. This node sends an exploration request to some of its neighbours to retrieve the satisfied network information for slave overlays. As the message cost of sending a request is smaller than that of estimating the network, this exploration approach could improve the RR of CNPE.

4.2.1 Maintenance cost model: This section models the RR of CNPE with the number of network-proximity estimation operations and the overhead of CNPE. The overhead of CNPE only comes from $\mathrm{msg_e}$ introduced by the exploration approach and is calculated by

$$C_{\mathrm{CNPE}} = C(\mathrm{msg_e}) = L_c \times P_e \tag{10}$$

where L_c is the number of created links in slave overlays and P_e is the probability of performing the exploration process.

The maintenance cost of a MOE without CNPE is modelled as

$$M = \sum_{O \in \mathrm{MOE}} E^{\mathrm{O}} \tag{11}$$

where E^{O} presents the number of executed network-proximity

IET Commun., 2014, Vol. 8, Iss. 15, pp. 2676–2683
doi: 10.1049/iet-com.2013.0709

estimation operations. Regarding the maintenance cost of a MOE with CNPE, the cost is defined as

$$M'_{\text{CNPE}} = E^{\text{m}} + B \times C_{\text{CNPE}} \tag{12}$$

where B stands for the overhead weight. Hence, the reduction ratio of CNPE is calculated as

$$\text{RR} = \left(1 - \frac{M'_{\text{CNPE}}}{M}\right) \times 100\% \tag{13}$$

4.2.2 Selection criteria and approach: The first criterion for selecting a suitable master overlay is the size of a routing table. That is because a larger routing table may provide more information for slave overlays. The additional overhead of the exploration approach is further reduced when the master satisfies the needs of slave overlays. The second criterion to select a master overlay concerns the network-proximity estimation. In a MOE, the master overlay should have the accurate estimation approach that satisfies the minimum requirements of the request from other slave overlays.

An interesting phenomenon is that the network dynamics do not have an impact on the selection of the master overlay. That is because the reduction ratio is not affected by the churn rate, as shown in previous experimental results [8]. Accordingly, the network dynamics is not considered while selecting a master overlay.

For CNPE, the score function is refined to decide the better master overlay. The formula is defined as

$$s = \left(1 - \frac{\max(K^{\text{m}},\ K^{\text{s}})}{K^{\text{m}} + K^{\text{s}}}\right) \times 100\% \tag{14}$$

where K, m and s are the same as those in (9). If two scores are the same, the overlay with maximum K should be chosen as the master overlay. Taking the exploration approach into account, the score function can be refined as

$$s = \left(\frac{\min(K^{\text{m}},\ K^{\text{s}}) + S_{\text{plus}}}{K^{\text{m}} + K^{\text{s}}}\right) \times 100\% \tag{15}$$

where S_{plus} is the score related to the exploration approach. S_{plus} is positive if $K^{\text{m}} < K^{\text{s}}$; otherwise, S_{plus} is zero.

5 Experimental results

In this paper, the PeerSim simulator is used to evaluate our proposed strategy. The cycle-based simulation engine is applied in the experiment. Our evaluations consider three kinds of overlays: the unstructured overlay ($O_{\text{unstructured}}$), the ring-based structured overlay (O_{ring}) and the tree-based structured overlay (O_{tree}). Each of these overlays has a parameter K, specifying the number of neighbours.

Our experiments use the following setting, so then considered different MOEs: $K=4$ for $O_{\text{unstructured}}$, $K=2$ for O_{ring} and $K=3$ for O_{tree} with $H=1$ links to other nodes when its parent fails. These diverse environments are simulated through different OCs, listed as $\{O_{\text{unstructured}}, O_{\text{ring}}\}$, $\{O_{\text{unstructured}}, O_{\text{tree}}\}$ and $\{O_{\text{ring}}, O_{\text{tree}}\}$. For each combination of overlays, there are two choices of the master overlay. Thus, six cases are all examined, as listed in Table 2. Owing to the limited length, this paper only depicts some results which are differential to the previous work [8].

Table 2 Overlay combinations

Overlay		Master		
		A	B	C
slave	A		OC III	OC V
	B	OC I		OC VI
	C	OC II	OC IV	

A: unstructured, B: ring, C: tree

5.1 Cooperative failure detection

The performance evaluation of CFD is presented and the RR is adopted as the performance metric. The first expanding experiment is to evaluate the CFD under different scales of MOEs. The scale represents the number of slave overlays. That is, the MOE is composed of five slave overlays when the scale is five. Fig. 2 reveals that when the number of slave overlays increases, the number of redundant detection messages also becomes large. Hence, the experimental result shows that CFD helps with eliminating redundant failure-detection operations, especially in the environment with many slave overlays.

On the other hand, previous experiments [8] show that CFD has performance degradation especially in the highly dynamic environments. Therefore, the aggregator approach is introduced to improve this shortcoming. Fig. 3 depicts that the aggregator method enhances the reduction ratio even more than 20% in the most fluctuating environment. Consequently, the CFD mechanism with the enhanced aggregator method provides a significant performance gain in the environment with the highest churn rate.

5.2 Cooperative network-proximity estimation

To evaluate the performance of the CNPE mechanism, two unstructured overlays are built with $K=4$ and $K=6$. Two MOEs are considered in the following experiments: 'OC A' is the OC with the master overlay $O_{\text{unstructured}}$ $(K=4)$ and 'OC B' is the combination with $O_{\text{unstructured}}$ $(K=6)$ as the master overlay. Regarding the exploration approach, it promotes the success rate of handling the request from a

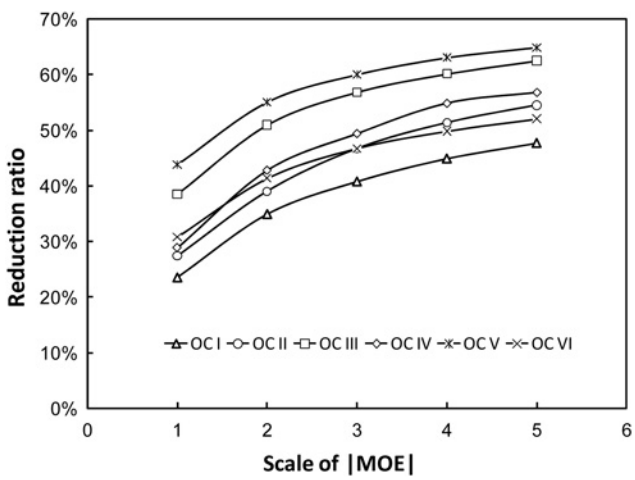

Fig. 2 *RR when ST = 11.6 m and N = 1000*

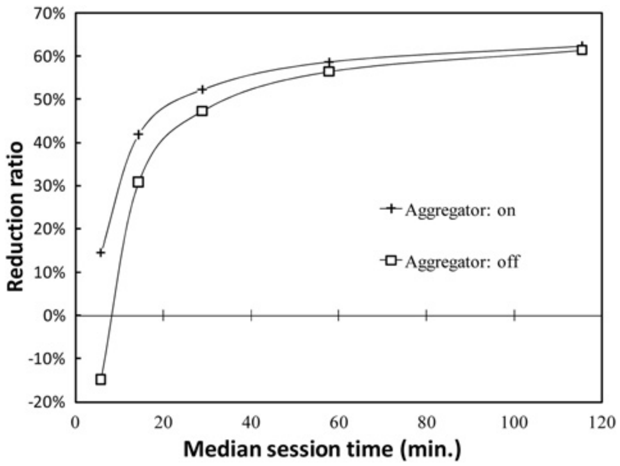

Fig. 3 *RR of 'OC III' when N = 10 000*

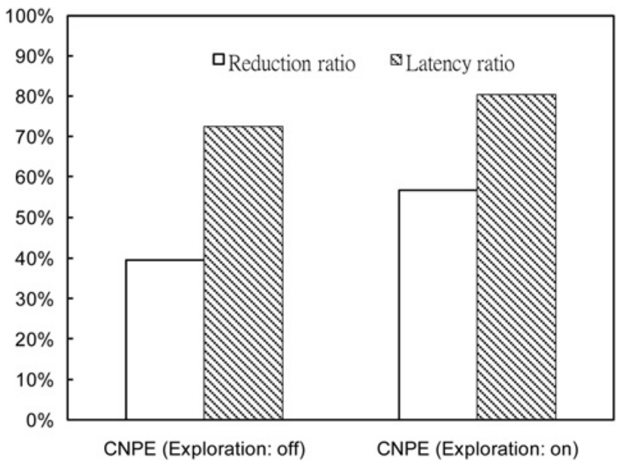

Fig. 4 *Effect of the exploration method on average latency for 'OC A'*

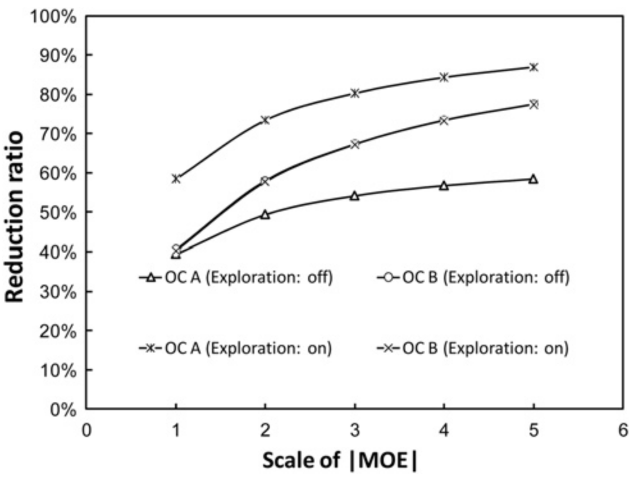

Fig. 5 *RR of varied MOE sizes*

slave overlay. Although the exploration procedure generates extra overhead on sending msg_e, the message size of msg_e is relatively smaller than that of network-proximity estimation. The size ratio of the exploration message to the network-proximity estimation message is set to B (10% in this experiment).

To evaluate the exploration approach, the RR and the latency ratio are both adopted as metrics. The latency ratio is calculated as dividing the average latency of neighbours in the slave overlay by that in the original overlay. For example, the latency ratio equals to 90% represents that the slave overlay improves the network-proximity estimation by 10%. Fig. 4 reveals that the average latency of exploration approach is larger than that without the exploration approach. A trade-off can be found that the reduction ratio can be improved by nearly 20%. Next, the varied scale of a MOE is also evaluated in this experiment. As expected in Fig. 5, the reduction ratio gains when the number of slave overlays grows. That is because CNPE is able to share the information of network state with more slave overlays to reduce a higher amount of maintenance operations.

5.3 Selection of master overlay

This section demonstrates the effectiveness of master selecting. To evaluate the selection method for CFD, "MOE A" is built with a six-overlay environment composed of $O_{\text{unstructured}}$ $(K=4)$, $O_{\text{unstructured}}$ $(K=8)$, O_{ring} $(K=2)$, O_{ring} $(K=6)$, O_{tree} $(K=3,\ H=2)$ and O_{tree} $(K=5,\ H=3)$. Fig. 6 reveals the best and the worst RR. The RRs of random selection and score-based selection are also presented. The results show that the score-based function can achieve nearly the best RR.

On the other hand, a 'MOE B' composed of $O_{\text{unstructured}}$ $(K=K_1)$ and $O_{\text{unstructured}}$ $(K=K_2)$ is constructed to evaluate

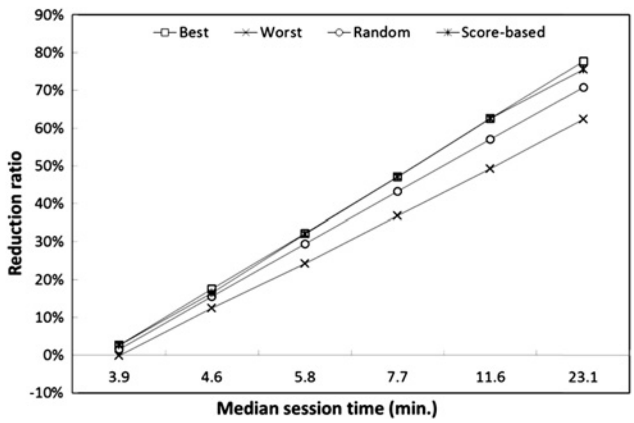

Fig. 6 *Selection approach of CFD*

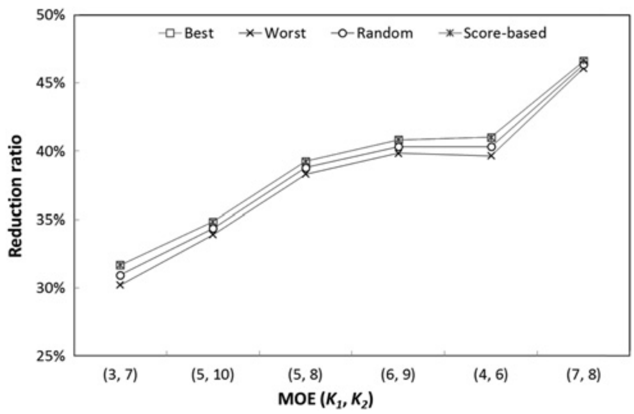

Fig. 7 *Selection approach of CNPE*

doi: 10.1049/iet-com.2013.0709

the score function for CNPE. Since the RR of CNPE is not affected by the churn rate, the score function is not evaluated under varied STs; instead, the parameter K is varied. Fig. 7 depicts the result of score function. The result reveals that the RR is significantly improved as the degree of covered K increases. As so, the RR of pair (4, 6) is higher than that of pair (5, 10), in which the coverage ratio is 66.7 and 50%, respectively.

6 Conclusions and future works

For the sake of handling the common overlay maintenance in diverse MOEs, this paper conducts a cooperative strategy on eliminating a considerable amount of redundant overlay-maintenance operations. A generic master–slave model and corresponding interactions are proposed to realise the cooperative strategy. Two mechanisms (CFD and CNPE) are introduced based on the event-driven (subscription/notification) and information-retrieval (query/response) overlay maintenance. To evaluate CFD and CNPE, experiments consider some complex environments with different parameters and OCs. Results show that our proposed strategy and the master–slave model significantly trim the redundant overlay maintenance overhead. In some cases, the RR is more than 60%.

This work is a good start to illustrate how the proposed approach can be applied to some common operations. The failure-detection and the network-proximity estimation are taken as examples. Other types of maintenance operations are remaining as future discussions. A comprehensive understanding on supporting an automatic selection of the master overlay and an optimised solution for a cooperative strategy are also interesting. Furthermore, as cloud computing is servicing a variety of cloud applications, realising our proof of concepts based on open-source cloud platform is being considered as future work.

7 References

1 Zhuang, S.Q., Geels, D., Stoica, I., Katz, R.H.: 'On failure detection algorithms in overlay networks'. Proc. IEEE INFOCOM, Miami, FL, USA, March 2005, pp. 2112–2123

2 Rowstron, A., Druschel, P.: 'Pastry: scalable, decentralized object location, and routing for large-scale peer-to-peer systems'. Proc. Springer-Verlag Int. Conf. Middleware, Heidelberg, Germany, November 2001, pp. 329–350

3 Castro, M., Druschel, P., Hu, Y.C., Rowstron, A.: 'Exploiting network proximity in distributed hash tables'. Proc. Springer-Verlag Int. Workshop FuDiCo, Italy, June 2002, pp. 52–55

4 Rhea, S., Geels, D., Roscoe, T., Kubiatowicz, J.: 'Handling churn in a DHT'. Proc. USENIX Annual Technical Conf., Boston, MA, USA, June 2004, pp. 127–140

5 Zhang, X.Y., Zhang, Q., Zhang, Z., Song, G., Zhu, W.: 'A construction of locality-aware overlay network: mOverlay and its performance', *IEEE J. Sel. Areas Commun.*, 2004, **22**, (1), pp. 18–28

6 Maniymaran, B., Bertier, M., Kermarrec, A.-M.: 'Build one, get one free: leveraging the coexistence of multiple P2P overlay networks'. Proc. IEEE ICDCS, Toronto, Ontario, Canada, June 2007, pp. 33–40

7 Lin, S., Taïani, F., Blair, G.S.: 'Exploiting synergies between coexisting overlays'. Proc. Springer-Verlag Int. Conf. DAIS, Lisbon, Portugal, June 2009, pp. 1–15

8 Hsu, C.-J., Chung, W.-C., Lai, K.-C., Li, K.-C., Chung, Y.-C.: 'A novel approach for cooperative overlay-maintenance in multi-overlay environments'. Proc. IEEE Int. Conf. CloudCom, Indianapolis, IN, USA, November 2010, pp. 81–88

9 Grace, P., Coulson, G., Blair, G., *et al.*: 'GRIDKIT: pluggable overlay networks for grid computing'. Proc. Springer-Verlag Int. Symp. DOA, Agia Napa, Cyprus, October 2004, pp. 1463–1481

10 Mao, Y., Loo, B.T., Ives, Z.G., Smith, J.M.: 'MOSAIC: unified declarative platform for dynamic overlay composition'. Proc. ACM Int. Conf. CoNEXT, Madrid, Spain, December 2008, pp. 1–12

11 Jiang, W., Chiu, D.-M., Lui, J.C.S.: 'On the interaction of multiple overlay routing', *Perform. Eval.*, 2005, **62**, (1–4), pp. 229–246

12 Kwon, M., Fahmy, S.: 'Synergy: An overlay internetworking architecture'. Proc. IEEE ICCCN, San Diego, CA, USA, October 2005, pp. 401–406

13 Cooper, B.F.: 'Trading off resources between overlapping overlays'. Proc. Springer-Verlag Int. Conf. Middleware, Melbourne, Australia, November 2006, pp. 101–120

14 Wu, C., Li, B., Li, Z.: 'Dynamic bandwidth auctions in multioverlay P2P streaming with network coding', *IEEE Trans. Parallel Distrib. Syst.*, 2008, **19**, (6), pp. 806–820

15 Gummadi, K., Gummadi, R., Gribble, S., Ratnasamy, S., Shenker, S., Stoica, I.: 'The impact of DHT routing geometry on resilience and proximity'. Proc. ACM Int. Conf. SIGCOMM, Karlsruhe, Germany, August 2003, pp. 381–394

16 Ratnasamy, S., Francis, P., Handley, M., Karp, R., Schenker, S.: 'A scalable content-addressable network'. Proc. ACM Int. Conf. SIGCOMM, San Diego, CA, USA, August 2001, pp. 161–172

17 Hsu, C.-J., Chung, W.-C., Lai, K.-C., Li, K.-C., Chung, Y.-C.: 'Cooperative failure detection in multi-overlay environments', *J. Internet Technol.*, 2011, **12**, (2), pp. 259–267