INTERNATIONAL JOURNAL OF NETWORK MANAGEMENT
Int. J. Network Mgmt 2015; **25**: 375–393
Published online 11 May 2015 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/nem.1891

An economic mechanism for request routing and resource allocation in hybrid CDN–P2P networks

Mehran Garmehi,[1,*†] Morteza Analoui,[1] Mukaddim Pathan[2] and Rajkumar Buyya[3]

[1]*School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran*
[2]*Media Distribution, Telstra Corporation Limited, Melbourne, Australia*
[3]*CLOUDS Lab, Department of Computing and Information Systems, The University of Melbourne, Melbourne, Australia*

SUMMARY

Hybrid content delivery networks (HCDN) benefit from the complementary advantages of P2P (Peer to Peer) networks and CDNs (Content Delivery Network). In order to extend a traditional CDN and enable it to offer hybrid content delivery service, we have modified a traditional domain name system-based request routing mechanism. The proposed scheme relies on an oligopolistic mechanism to balance the load on the edge servers and employs a truthful profit-maximizing auction to maximize the contribution of users in the P2P content delivery. In particular, economics of content delivery in HCDNs is studied, and it is shown that through our request routing mechanism, it is possible to deliver higher quality of service to majority of end-users, increase the net profit of the HCDN provider and decrease the content distribution cost at the same time. Copyright © 2015 John Wiley & Sons, Ltd.

Received 25 August 2014; Revised 16 March 2015; Accepted 30 March 2015

1. INTRODUCTION

In order to offer low-priced, scalable, high-quality and reliable content delivery service, hybrid content delivery networks (HCDNs) blend Content Delivery Network (CDN) and Peer to Peer (P2P) technologies. As illustrated in Figure 1, in HCDNs, users have the option to choose either of the base quality client–server (CS) or the high-quality hybrid (P2P) modes to receive the content [1]. In P2P mode, users have to contribute their upload bandwidth in the P2P content delivery to receive higher quality of service (QoS).

Similar to CDNs, in HCDNs, the content provider pays cost of the content delivery service, based on the service-level agreement (SLA) offered by the (H)CDN provider. In practice, if the user chooses to receive the higher-quality content in P2P mode, not only the content provider pays content delivery fees agreed on the SLA, but also the end-users have to afford a fraction of the content delivery costs in the form of its upload bandwidth. In other words, P2P end-users have to spend a fraction of their network upload capacity on the P2P content delivery process. Therefore, in order to incentivize the end-users to contribute in content delivery optimally, the P2P service must be offered with more quality compared with the CS service [1]. Moreover, the price of the P2P service at the SLA must be kept less than the CS service to convince the content provider to ask for the hybrid content delivery service.

Traditional CDNs optimize their performance by relying on two key mechanisms: replica placement (RP) and request routing (RR) [2,3]. While RP mechanism decides for the locations of replicas, the RR subsystem receives the outcome of RP mechanism and redirects users to the optimum available server. Most existing CDNs employ domain name system (DNS)-based RR mechanism to minimize their

*Correspondence to: Mehran Garmehi, School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran.
†E-mail: Garmehi@iust.ac.ir

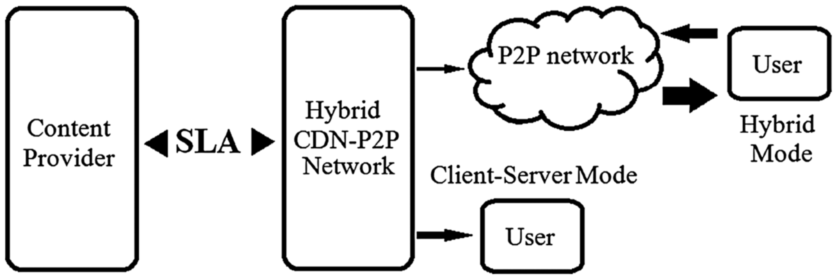


Figure 1. Client–server and peer-to-peer (P2P) content delivery approaches in hybrid content delivery networks (thickness of arrows show traffic intensity). SLA, service-level agreement

content delivery cost [4,5]. DNS-based RR mechanism replies to each request with a list of nearby edge servers. Because all the end-users of the same content demand the same amount of network resources from the CDN perspective, there is no difference among the candidate edge servers listed in the DNS response message; therefore, users are free to choose each of them.

Compared with the RR mechanism in traditional CDNs, the RR mechanism of HCDN requires considering new aspects. First, in HCDN, the P2P content delivery approach can decrease the bandwidth consumption and increase the net profit of the HCDN provider. While the sole economic objective of the RR mechanism in CDNs is to minimize the content distribution costs, in HCDNs, bandwidth contribution of the end-users and different prices of the P2P service (in SLA) change the economic model of the system. Therefore, in designing the RR mechanism of HCDNs, both cost and revenue of the system must be taken into account. Second, the availability and price of resources in P2P subnetworks in HCDNs vary from server to server. Therefore, the selection of the most economic edge server to serve a user is crucial. And last but not the least, the RR mechanism, in addition to determining the address of the best server, must determine the contribution level of users interested to receive the content in P2P mode. In P2P networks, bandwidth allocation is a key challenge [6]. Users of these networks normally behave selfishly [7] and prefer not to contribute their network resources into the P2P content distribution process. This fact faces P2P networks with the free-riding problem. P2P networks employ various techniques to overcome this problem [8]. Therefore, the RR mechanism of the HCDN must develop a coherent policy to motivate the end-users to contribute maximally in P2P content delivery and prevent free riding.

In this paper, we introduce an economic RR mechanism to enable a commercial traditional CDN to offer hybrid content delivery services. The mechanism relies on the mathematical modelling of the HCDN service economy and considers the net profit of HCDN provider as the objective. We have employed a truthful profit-maximizing auction [9] to determine the optimum level of end-user's contribution and an oligopolistic market [10] to redirect the end-users to the most profitable edge server. This economic RR mechanism demands cooperation of edge servers and a traditional DNS-based RR mechanism with a hybrid RR server (HRRS). It relies on communication of a few small messages to determine the price and quantity of production at each edge server. Price and production quantity of each edge server in this domain are equivalent to the contribution amount of end-users in the form of upload bandwidth and the way they are clustered among edge servers.

The major contribution of this paper is an economic RR and resource allocation mechanism for streaming content distribution in HCDNs. This approach extends the RR of traditional CDNs, enabling them to offer hybrid CDN–P2P services. Our proposed mechanism obeys the SLA, maximizes the net profit of hybrid CDN–P2P service provider, increases the contribution of the users and prevents the free-riding problem. Moreover, it is able to be integrated with the existing popular DNS-based RR mechanisms.

The remainder of this paper is organized as follows: We explore related works and compare the problem with neighbouring problems in Section 2. In Section 3, the problem of economic RR in HCDNs is introduced. Section 4 introduces the proposed mechanism. The proposed mechanism is evaluated in Section 5. Finally, we conclude in Section 6.

2. RELATED WORK

Many experimental [1] and theoretical [11] studies show the potentials of HCDNs to cover shortcomings of P2P networks and CDNs. Ha *et al.* compare costs of processing and storage in CDNs, P2P networks and HCDNs [12]. Hai *et al.* give an in-depth understanding of HCDN's effectiveness in [13]. They carry out a detailed performance evaluation based on the deterministic fluid model and provide numeric results of HCDN, comparing with conventional CDN and P2P networks. A comparison of bandwidth consumption of CDNs and HCDNs is presented by Huang *et al.* [11]. They have showed that application of these networks can decrease the bandwidth consumption of a content delivery system by 65%.

In the literature, there are different architectures for HCDNs [14]. These architectures can be classified to P2P-inspired HCDNs and CDN-inspired HCDNs. In the first class [15–17], the architecture of the HCDN is inspired from P2P systems, and the CDN part is employed to fill the missing blocks of the playback buffer. In these networks, the economic model of the system has the most similarity with the traditional P2P networks. Therefore, from economics point of view, P2P-inspired HCDNs have significant difference with commercial CDNs. In these networks, the resource scheduling mechanism mostly considers the social welfare of end-users as the optimization goal [18]. The second class of HCDNs inherits the architecture of traditional CDNs. These networks are able to be deployed as commercial HCDNs. They can extend existing traditional CDNs and are able to tolerate the fact that some end-users might not prefer to install the special hybrid content distribution application software on their machines [1]. End-users equipped with the special hybrid content delivery software are able to receive the content in higher-quality P2P mode [19]. CDN-inspired HCDNs not only inherit the architecture of CDNs, but also their economic model has the maximum similarity with these networks. In the economic cycle of CDN-inspired HCDNs, the objective is to maximize the net profit of the HCDN provider. LiveSky [1] can be referred as one of these architectures and also as the first successful deployment of HCDNs. LiveSky extends an old-fashioned CDN to an HCDN and has succeeded to serve 10 million concurrent requests for streaming videos.

Content delivery networks employ RR and RP mechanisms to optimize their performance [2,20,21] and introduce RP mechanisms for HCDNs. While the RP mechanism decides for the place and number of replicas, the RR mechanism of the HCDN in addition to assigning end-users and edge servers (like CDNs) is needed to incentivize end-users and try to improve their contribution in the P2P content delivery. Although there are several implementations and studies in the field of HCDNs, to the best of our knowledge, there is no work on economics of RR and resource allocation mechanisms in HCDNs, and this work might be the first.

In the literature, there is a great body of research on the resource allocation mechanism in P2P networks. The resource allocation mechanism of these networks must answer many concerns including scalability, fault tolerance and especially fairness. Contemporary P2P networks mostly rely on microeconomics (e.g. the eMule's crediting system) or game theory (e.g. the BitTorrent's tic-tac-toe strategy) to overcome the selfish behaviour of the end-users, prevent the free-riding problem and improve the fairness among the peers [7]. In this research, we need to solve the free-riding problem in the P2P subnetworks of the HCDN, but two facts make it impractical to rely on the existing resource allocation mechanisms popular in the P2P networks. First, the economic model of the HCDN is completely different from P2P networks. While in P2P networks, the goal of the system is to improve the satisfaction of the end-users (social welfare), in HCDNs, QoS is guaranteed based on the SLA; therefore, the objective of the resource allocation mechanism is to improve the profit of the HCDN provider. The second major difference relies on the fact that peers of P2P networks interact with each other through microeconomic or game-theoretic mechanisms (pay virtual credit or send messages of a game), but in HCDN environment, end-users (peers) must interact with the RR mechanism directed by the HCDN provider. In other words, the receiver of the money and form of the market in the HCDN environment are completely different.

The resource allocation and RR mechanism of the HCDN must consider both economic models of the HCDN architecture and behaviour of the end-users [12]. We rely on two economic mechanisms to select the most beneficial assignment of the end-users and edge servers and to maximize contribution of end-users in hybrid content delivery. In this paper, we extend

DOI: 10.1002/nem

the traditional DNS-based RR mechanism to enable hybrid content delivery in CDNs. Despite shortcomings of the DNS-based RR [4], most of the popular existing CDNs, for example, Akamai, employ this mechanism because of its ease of management, scalability and simplicity [2].

In the theory of economics, the situation in which the market is dominated by small number of sellers is referred to as oligopoly. Oligopolistic markets are based on collusion, and they are able to maximize their net profit even with uncertainty about production costs [22]. Competition of firms in oligopolies might be on price, production quantity or both [10]. Because availability and price of bandwidth at different edge servers are variable [4], we apply an oligopolistic market to determine production quantity among the edge servers introduced by DNS-based RR. The RR mechanism employs an oligopolistic mechanism to choose one of handful near edge servers to serve each user. Because oligopolistic markets, even in symbolic economic model of overlay environments are rare [22], there are few works applying this economic model. Weihong and Li [23] have applied a Cournot oligopoly to propose a control-theoretic solution to adapt the user contribution in P2P file sharing environment.

Request routing mechanism of HCDN, in addition to assigning end-users to edge servers, must optimize contribution of end-users in hybrid content delivery. Bandwidth discovery and allocation in traditional P2P overlay networks are a well-known problem in computer networks literature [24,25]. Because of the high potentials of economic mechanisms and game theory, many of the works in this domain rely on auctions and incentivizing peers [26]. In game theory [27], truthful auctions are divided into two classes: (i) mechanisms relying on truthful auctions and implementing social choice functions, for example, Vickrey–Clarke–Groves (VCG) and Vickrey auctions, and (ii) truthful profit maximization auctions. Because of the economic model of P2P networks, mechanisms of the first group are applied widely [28]. While VCG auctions suffer from several limitations in practice [29], the most successful mechanism in this group of mechanisms is Vickrey second-price auction [30,31]. In this work, features of the problem are shown to be more compatible with one of the mechanisms classified in the second group [9]. There are two categories of profit-maximizing auctions: Bayesian approach and prior-free approximation. The Myerson approach [27] belongs to the first category and is proved to be able to maximize the net profit of digital good auction when distribution of private value of buyers comes from a previously known i.i.d.[1] statistical distribution. In the literature of economic mechanism design, profit-maximizing auctions are usually verified against VCG and Vickrey auctions [32]. Therefore, we apply the Myerson mechanism to implement the solution and verify it against a Vickrey auction.

3. HYBRID CONTENT DELIVERY NETWORK–PEER-TO-PEER NETWORKS

3.1. Architecture of a hybrid content delivery network–peer-to-peer network

An HCDN employs the architecture shown in Figure 2, to receive the content from the content provider and deliver it to the end-users. In this network, the content provider contracts an SLA with the HCDN provider and provides the distribution servers with flow of the content. The content is multicast through the intermediate servers [25] and replicated on edge servers selected by the RP mechanism [21].

As illustrated in Figure 3, the RP mechanism provides the RR mechanism with the list of edge servers containing replicas of the content. Users progressively send their requests to the RR mechanism, one by one. The RR mechanism must provide each user with the address of a suitable server containing a replica of the content. If the user prefers to receive the content in hybrid mode, the RR mechanism and the user must agree on the necessary upload bandwidth contribution of the user.

[1]Independent and identically distributed.

DOI: 10.1002/nem

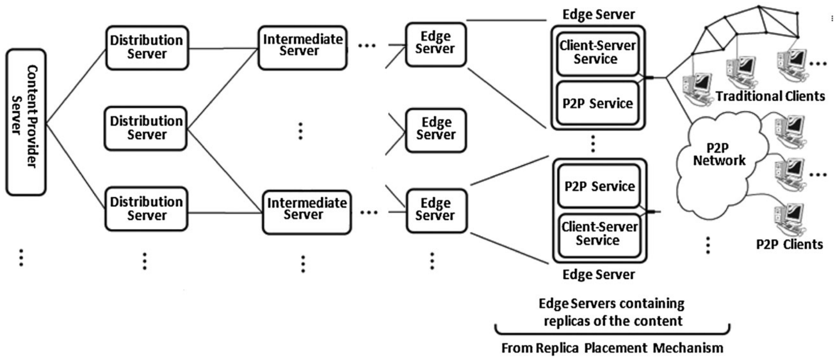


Figure 2. Architecture of a Hybrid CDN-P2P network

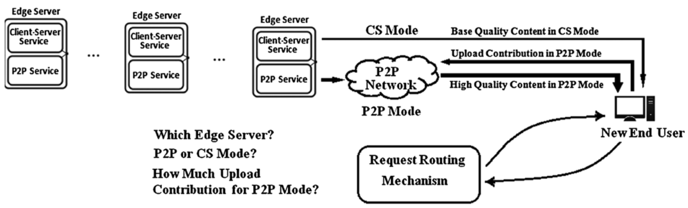


Figure 3. The RR mechanism in HCDN

3.2. Economics of hybrid content delivery network–peer-to-peer networks

Table 1 summarizes the notations used in this paper to model economics and performance of the HCDN. An HCDN contracts SLAs with content providers to distribute the content using client–server and P2P modes. The revenue function in the contracted SLA for distribution of content i between a commercial HCDN provider and the content provider is assumed to be as follows:

$$R_{\mathrm{SLA}i} = P_{\mathrm{CS},i}.\ \mathrm{BW}_{\mathrm{CS},i}.\ N_{\mathrm{totalCS},i} + P_{\mathrm{P2P},i}.\ \mathrm{BW}_{\mathrm{P2P},i}.\ N_{\mathrm{totalP2P},i}\ , \tag{1}$$

where $R_{\mathrm{SLA}i}$ determines the revenue of the HCDN provider according to the SLA_i.$N_{\mathrm{totalCS},i}$ and $N_{\mathrm{totalP2P},i}$ represent the total number of users served with the content i using the CS and P2P approaches, respectively. $P_{\mathrm{CS},i}$ and $P_{\mathrm{P2P},i}$ denote the contracted prices of bandwidth in CS and P2P modes. Bandwidth of the content i in CS and P2P modes are indicated by $\mathrm{BW}_{\mathrm{CS},i}$ and $\mathrm{BW}_{\mathrm{P2P},i}$, respectively. In order to keep the P2P content distribution economic and also prevent the HCDN from becoming a pure CDN, it is assumed that $\mathrm{BW}_{\mathrm{CS},i}.P_{\mathrm{CS},i} > \mathrm{BW}_{\mathrm{P2P},i}.P_{\mathrm{P2P},i} > 0$. Additionally, $\mathrm{BW}_{\mathrm{CS},i} \ll \mathrm{BW}_{\mathrm{P2P},i}$ guarantees that users have enough incentives to contribute in the P2P content delivery. The amount of bandwidth an edge server spends to serve the CS and P2P end-users receiving content i is expressed as follows:

$$\mathrm{BW}_{\mathrm{ES},i} = \mathrm{BW}_{\mathrm{CS,ES},i} + \mathrm{BW}_{\mathrm{P2P,ES},i} = \mathrm{BW}_{\mathrm{CS},i}.\ N_{\mathrm{CS,ES},i} + \alpha_{\mathrm{ES},i}.\ \mathrm{BW}_{\mathrm{P2P},i}.\ N_{\mathrm{P2P,ES},i}, \tag{2}$$

where $\mathrm{BW}_{\mathrm{ES},i}$, $\mathrm{BW}_{\mathrm{CS,ES},i}$ and $\mathrm{BW}_{\mathrm{P2P,ES},i}$ symbolize the total bandwidth consumption, the bandwidth spent for $N_{\mathrm{CS,ES},i}$ CS end-users and the bandwidth spent to serve $N_{\mathrm{P2P,ES},i}$ P2P end-users, respectively. $\alpha_{\mathrm{ES},i}$ represents the effectiveness factor of P2P content delivery for the content i at this edge server. Prices of the bandwidth at different edge servers in practice are variable and represented by $\mathrm{BWP}_{\mathrm{ES}}$. So expenditures or cost function of an edge server is quantified using equation (3) and denoted by $\mathrm{Cost}_{\mathrm{ES},i}$.

$$\mathrm{Cost}_{\mathrm{ES},i} = \mathrm{BW}_{\mathrm{ES},i}.\ \mathrm{BWPrice}_{\mathrm{ES}}. \tag{3}$$

DOI: 10.1002/nem

Table 1. Summary of notations

Notation	Description
$R_{\mathrm{SLA}i}$	Revenue of the hybrid CDN provider based on SLA_i
$N_{\mathrm{totalCS},i}$	Total number of users served with content i using CS approach
$N_{\mathrm{totalP2P},i}$	Total number of users served with content i using P2P approach
$P_{\mathrm{CS},i}$	Price of bandwidth in CS mode for content i
$P_{\mathrm{P2P},i}$	Price of bandwidth in P2P mode for content i
P_ratio	Pricing ratio ($P_{\mathrm{P2P},i}/P_{\mathrm{CS},i}$)
$\mathrm{BW}_{\mathrm{CS},i}$	Bandwidth of content i in CS mode
$\mathrm{BW}_{\mathrm{P2P},i}$	Bandwidth of content i in P2P mode
B_ratio	Ratio of P2P stream bandwidth to CS stream bandwidth ($\mathrm{BW}_{\mathrm{P2P},i}/\mathrm{BW}_{\mathrm{CS},i}$)
$\mathrm{BW}_{\mathrm{ES},i}$	Total bandwidth consumption for distribution of content i at the edge server ES
$\mathrm{BW}_{\mathrm{CS,ES},i}$	Bandwidth spent for distribution of the content i in CS mode at the edge server ES
$\mathrm{BW}_{\mathrm{P2P,ES},i}$	Bandwidth spent for distribution of the content i in P2P mode at the edge server ES
$N_{\mathrm{CS,ES},i}$	Number of end-users receiving content i in CS mode from edge server ES
$N_{\mathrm{P2P,ES},i}$	Number of end-users receiving content i in P2P mode from edge server ES
$\alpha_{\mathrm{ES},i}$	Effectiveness factor of distribution of content i in P2P mode at edge server ES
$\mathrm{BWP}_{\mathrm{ES}}$	Price of bandwidth at edge server ES
$\mathrm{Cost}_{\mathrm{ES},i}$	Cost of distribution of content i at edge server ES
$\mathrm{Replicas}_i$	Set of edge servers containing replicas of the content i
$\mathrm{RPCost}_{\mathrm{SLA}i}$	Costs spent by RP mechanism for replication of content i
$U_{\mathrm{SLA}i}$	Profit or utility of the hybrid CDN–P2P network due to SLA_i
U_{User}	Utility of the user due to receiving the content i in either CS or P2P modes
$P_{\mathrm{User},D}$	Price of the user’s download bandwidth
$P_{\mathrm{User},U}$	Price of the user’s download bandwidth
β_{User}	Ratio of the user’s upload to download bandwidth in P2P mode
$\mathrm{P2P}_{\mathrm{ES},i}$	Set of all P2P end-users receiving their service from ES
$\mathrm{CS}_{\mathrm{ES},i}$	Set of all CS end-users receiving their service from ES
$\mathrm{DLC}_{\mathrm{User}}$	Capacity of user’s downlink
$\mathrm{ULC}_{\mathrm{User}}$	Capacity of user’s uplink
$\mathrm{Max\_Contribution}_{\mathrm{User}}$	Maximum contribution of the user in P2P mode
PRF(BW)	Preference function of user
TRR(i,User)	Set of local edge servers containing replicas of content i introduced by the traditional request routing mechanism
$\mathrm{OptES}_{\mathrm{User}}$	Selected optimum ES to serve user

The cost function of content distribution at the edge servers in HCDN for SLA_i can be represented as follows:

$$\mathrm{Cost}_{\mathrm{SLA}i} = \sum_{\mathrm{ES} \in \mathrm{Replicas}_i} \mathrm{Cost}_{\mathrm{ES},i} + \mathrm{RPCost}_{\mathrm{SLA}i}\,, \tag{4}$$

where the set $\mathrm{Replicas}_i$ denotes all the edge servers containing a replica of the content i and $\mathrm{RPCost}_{\mathrm{SLA}i}$ indicates costs spent by the RP mechanism for replication of this content. As we are about to design an RR mechanism, $\mathrm{Replicas}_i$ and $\mathrm{RPCost}_{\mathrm{SLA}i}$ are considered constant in this work. The profit or utility of the HCDN is symbolized by $U_{\mathrm{SLA}i}$ and can be obtained by

$$U_{\mathrm{SLA}i} = R_{\mathrm{SLA}i} - \mathrm{Cost}_{\mathrm{SLA}i}. \tag{5}$$

3.3. Economic modelling of end-user’s behaviour in hybrid content delivery network–peer-to-peer networks

We consider the end-users as rational agents trying to maximize their own utility with preference on high-quality content. We assume that the preference is a function of the bit rate of the content they receive, denoted by PRF(BW). Equation (6) gives the private utility function of a user for receiving the content in CS or P2P modes and describes how the end-users decide to choose the traditional low-quality CS service or join the P2P network and make contribution providing the requested upload bandwidth.

DOI: 10.1002/nem

$$U_{\text{User}} = \text{Max}\begin{cases} \text{PRF}\left(\text{BW}_{\text{User,CS}}\right) \; - P_{\text{User},D}.\text{BW}_{\text{CS}} & \text{(CS mode)} \\ \text{PRF}\left(\text{BW}_{\text{User,P2P}}\right) - P_{\text{User},D}.\text{BW}_{\text{P2P}} - \beta_{\text{User}}.P_{\text{User},U}.\text{BW}_{\text{P2P}} & \text{(P2P mode)} \end{cases}. \tag{6}$$

U_{User} denotes the expected utility of a user due to receiving the content i in CS or P2P modes. $P_{\text{User},D}$ and $P_{\text{User},U}$ are private functions for pricing the download and upload capacities of the user's link, and β_{User} is the ratio of upload to download bandwidth the user has to contribute to receive the high-quality P2P content from ES. Equation (7) describes the relation between $\alpha_{\text{ES},i}$ and β_{User}

$$\alpha_{\text{ES},i} = 1 - \frac{\sum_{\text{User} \in \text{P2P}_{\text{ES},i}} \beta_{\text{User}}}{\left|\text{P2P}_{\text{ES},i}\right|}. \tag{7}$$

Here, $\text{P2P}_{\text{ES},i}$ is the set of all the P2P end-users receiving their service from ES. Considering equation (6) and capacity of the uplink and downlink of the user denoted by DLC_{User} and ULC_{User}, the maximum contribution of an end-user in P2P mode denoted by $\text{Max\_Contribution}_{\text{User}}$ can be calculated as follows:

$$\text{Max\_Contribution}_{\text{User}} = \begin{cases} 0 & \text{(in CS mode)} \\ \text{Min}\left(\beta_{\text{User}}.P_{\text{User},U}(\text{BW}_{\text{P2P}}), \text{ULC}_{\text{User}}\right) & \text{(in P2P mode)} \end{cases}$$
$$\beta_{\text{User}}.P_{\text{User},U}(\text{BW}_{\text{P2P}}) = \text{PRF}\left(\text{BW}_{\text{User,P2P}}\right) - \text{PRF}\left(\text{BW}_{\text{User,CS}}\right) + P_{\text{User},D}(\text{BW}_{\text{CS}}) - \text{Price}_{\text{User},D}(\text{BW}_{\text{P2P}}). \tag{8}$$

This equation describes that first, in CS mode, the maximum contribution of an end-user, $\text{Max\_contribution}_{\text{User}}$, will be zero. This case happens only when the downlink of the end-user does not have enough capacity to pass the high-quality P2P content or when the end-user's machine is not equipped with the hybrid content delivery software. Second, in P2P mode, the end-user finds it economic to contribute as much as $\beta_{\text{User}}.P_{\text{User},U}(\text{BW}_{\text{P2P}})$ to receive the higher-quality content. When the end-user's uplink capacity is greater than $\beta_{\text{User}}.P_{\text{User},U}(\text{BW}_{\text{P2P}})$, the end-user's maximum contribution equals this number; otherwise, the uplink capacity (ULC_{User}) will be returned as the maximum possible contribution. $\text{Max\_contribution}_{\text{User}}$ is a private value of the user, and the HCDN has no direct way to find it.

3.4. Problem statement

In order to introduce the economic RR mechanism, we employ the notation used in economic mechanism design theory, illustrated in Figure 4 [33]. In this domain, the problem is introduced by specifying two parameters Θ and F, the environment space and goal function, respectively [33]. Environment space of the problem is the set of all instances of the problem. An instance of the problem or $\theta \in \Theta$ consists of current state of the system and features of the new user. As the outcome of the problem is the maximum profit, the outcome space, Z, is single dimension and is a subset of R^+. In this notation, F describes the goal function of the mechanism, F: $\Theta \rightarrow Z$. Therefore, the mechanism designer is requested to design a mechanism to produce the outcome of $F(\theta)$ for each $\theta \in \Theta$.

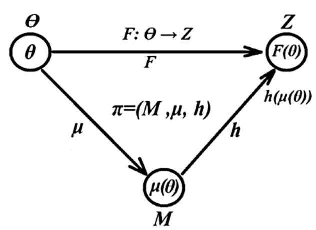


Figure 4. Notations used to introduce an economic mechanism

An economic mechanism is denoted by $\pi = (M, \mu, h)$ and consists of three elements, a message space, symbolized by M, an equilibrium message correspondence, denoted by μ, and an outcome function, denoted by h [33], where μ: $\Theta \rightarrow M$ and h: $M \rightarrow Z$. The message space M consists of the messages available for communication. The group equilibrium message correspondence μ associates with each environment, θ, the set of messages, $\mu(\theta)$, that are equilibrium or stationary messages for all the agents. The outcome function h translates messages into outcomes. Thus, the mechanism $\pi = (M, \mu, h)$ when operated in an environment θ leads to the outcome $h(\mu(\theta))$ in Z. If it is the case for all environments, θ, in the given space, Θ, the mechanism π leads to a desirable outcome in that environment, then we say that π realizes F on Θ.

In this paper, an instance of the problem θ is described as follows:

- SLA_i for distribution of content i.
- The user and its private maximum contribution.
- Set of local edge servers containing replicas of the requested content introduced by traditional RR mechanism, $\mathrm{TRR}(i, \mathrm{End\_User})$, and their current state including: $\mathrm{BWP}_{\mathrm{ES}}$, $\mathrm{BW}_{\mathrm{CS,ES},i}$, $\mathrm{BW}_{\mathrm{P2P,ES},i}$, $N_{\mathrm{CS,ES}}$, $N_{\mathrm{P2P,ES}}$ and $\alpha_{\mathrm{ES},i}$.

Objective function, F, is defined to maximize

$$F(\theta) = \begin{cases} \left(P_{\mathrm{CS},i} - \mathrm{BWP}_{\mathrm{ES}}\right).\mathrm{BW}_{\mathrm{CS},i} & (\text{CS mode}) \\ P_{\mathrm{P2P},i}.\mathrm{BW}_{\mathrm{P2P},i} - (1-\beta_{\mathrm{User}}).\ \mathrm{BW}_{\mathrm{P2P},i}.\ \mathrm{BWP}_{\mathrm{ES}} & (\text{P2P mode}) \end{cases}. \tag{9}$$

The only variables in this objective function are selected edge server and amount of the end-user's contribution. Therefore, in order to maximize $F(\theta)$, the mechanism must tune two variables, β_{User} and ES.

4. THE PROPOSED MECHANISM

4.1. Economic foundations of the request routing mechanism

In order to solve the problem defined earlier, we need to answer two questions for a new user. First, if the user is interested in receiving the content in P2P mode, how much must it pay (contribute) in the form of upload bandwidth to join the P2P service? Second, among the near servers containing replicas of the requested content, which server must take the responsibility to serve the user?

Hybrid CDN–P2P network has the sole responsibility of distribution of the content. Therefore, in this domain, HCDN is able to leverage its monopoly power to determine the optimum contribution level and select the best server. Based on microeconomics theory of monopolistic markets, a monopoly firm is a price maker [10]; therefore, the optimum outcome of the problem can be theoretically calculated by finding the optimum monopoly solution [34]. However, knowing that revenue and cost function of the HCDN, equations (1) and (4), depend on the private utility function of the end-users, equation (6), it is obvious that solving the problem using a straightforward monopoly approach is not trivial.

Users interested in the high-quality P2P content do not naturally reveal their maximum contribution level. If the RR mechanism finds it economically beneficial to serve the new user in P2P mode, the best strategy is to charge the user the highest possible upload bandwidth. In other words, the maximum possible value of β_{User} maximizes the objective function, equation (9). This strategy does not depend on the selected server. In order to determine β_{User}, we employ a profit-maximizing auction.

Because of the fact that CDNs are believed to have precise vision of the underlying network [35], under supervision of the administrator (when the number of requests is small) or automatically (when number of requests is large), it is possible to acquire the statistical distribution of private value of users (calculating probability density function (PDF) and cumulative distribution function (CDF) from previous observations) and apply the Myerson equation, equation (10), to determine the optimum price (requested bandwidth). The Myerson approach is proved to be both incentive compatible[2] and profit

[2] In mechanism design, a process is incentive compatible if all of the participants fare best when they truthfully reveal any private information asked for by the mechanism.

DOI: 10.1002/nem

Table 2. Message space (M) employed to implement the HRR mechanism

Message name	Source	Destination	Content	Comments
Hybrid content distribution request	End-user	Hybrid RR server (HRRS)	(Object ID, address of user and desired upload contribution)	The employed auction, Myerson, is truthful, so the user has to reveal its truthful maximum value of desired upload contribution
Traditional RR request	HRRS	Traditional RR server	(Object ID and address of user)	Exactly same as DNS request
Traditional RR response	Traditional RR server	HRRS	(Address of edge server_1, …, address of edge server_n)	Exactly same as DNS response, it returns address of n edge servers $\in \text{Replicas}_{\text{Content } i}$
Expected net profit request	RRHS	Edge server	(Object ID, address of user and determined upload contribution)	The HRRS calculates determined upload contribution using truthful desired upload contribution announced by the user and the Myerson equation
Expected net profit response	Edge server	HRRS	(Expected net profit, mode)	Each edge server applies equation (5) to calculate its expected net profit due to serving the new user with determined upload contribution either in P2P or CS modes
Hybrid content distribution response	HRRS	End-user	(Address of edge server, mode and determined upload contribution)	HRRS selects the address of the edge server with the maximum expected net profit and replies the end-user with the address, mode and determined upload contribution

DNS, domain name system; RR, request routing; P2P, peer to peer; CS, client–server.

maximizing and employs a take-it-or-leave-it offer to the bidder (end-user) [27].

$$\phi(b) = b - \frac{1 - \text{CDF}(b)}{\text{PDF}(b)} \rightarrow \text{Opt}(b) = \phi^{-1}(0) \tag{10}$$

In this equation, $b = \beta_{\text{User}}.\text{BW}_{\text{P2P},i}$, where b symbolizes the private value of the user. The Myerson function is denoted by $\phi(b)$. CDF(b) and PDF(b) represent the CDF and PDF of private value of users, respectively. $\text{Opt}(b) = \phi^{-1}(0)$ gives the optimum price of the auction or the optimum contribution of the end-user.

While the Myerson approach determines the amount of bandwidth contribution for end-users, we need another economic mechanism to assign an end-user to the most profitable edge server. In order to select the best server among, TRR(i,User), a simple oligopoly is applicable. Oligopolistic markets are based on collusion and are able to increase their benefit to maximum, keeping the prices as much as the monopoly price [10]. Foundation of these markets is based on cooperation of a small group of sellers splitting the market and acting as price setters. It is proved that the equilibrium for an oligopoly is Nash equilibrium[3] [10]. As mentioned earlier, the price of P2P service, $\beta_{\text{User}}.\text{BW}_{\text{P2P},i}$, is determined in equation (10), and the competition of firms, edge servers, would be only on quantity. For the new end-user, the optimum edge server denoted by $\text{OptES}_{\text{User}}$ among the candidate edge servers, TRR(i, User), obtained based on equation (9) and is expressed as follows:

$$\text{OptES}_{\text{User}} = \underset{ES}{\text{argmax}}\ (F(\theta)) : \{\text{ES}|\text{ES}\in\text{TRR}(i, \text{User})\} \tag{11}$$

4.2. Mechanism design

In this section, the HCDN RR (HRR) is introduced formally as an economic mechanism. The set of messages, M, employed in HRR is introduced in Table 2.

[3]Nash equilibrium: a situation in which economic actors are interacting with one another, each choose their best strategy given the strategies that all the other actors have chosen.

DOI: 10.1002/nem

Figure 5 illustrates the flow of $\mu(\theta)$, in HRR. In step 1, HRR receives the request of the user in the form of a hybrid content distribution request message. This message, in addition to the ID of the content and address of the user, contains its maximum desired contribution level to receive the content in P2P mode. The user is assumed to be rational and aware that an incentive-compatible auction is employed to reveal its private value. Thus, the dominant strategy for the user is to declare truthfully its maximum upload contribution [27] quantified by equation (8).

Hybrid CDN–P2P network RR begins when an end-user requests for the content i. First, the HRRS, in Figure 5, receives the hybrid content distribution request message and performs the following operations as step 2: First, it determines the optimum bandwidth contribution, based on equation (10), and decides to accept or reject the user's request. Second, HRR sends address of the user and the ID of the requested content through a traditional RR request message to the traditional DNS-based mechanism (traditional RR server).

In step 3, HRRS receives the traditional RR response (TRRRS) message. For rejected requests (end-users with desired contribution less than determined contribution level), HRR forwards the address of nearby edge servers obtained from TRR to the end-user to receive the content in CS mode (jumps to step 6). But for accepted requests in step 4, using expected net profit request messages, HRRS, requests the edge servers introduced in TRRRS message to determine their expected profit if they serve the new user with determined bandwidth contribution. In step 5, edge servers respond HRRS with expected net profit response messages and determine their maximum expected net profit to serve the new end-user in addition to the respective content distribution mode. Finally, in step 6, HRRS selects the server with maximum expected net profit and through a hybrid content distribution response message provides the user with the address of the selected edge server, the preferred content distribution mode and the requested bandwidth contribution.

It is assumed that $h(\mu(\theta))$ is being calculated using equation (9). Moreover, the mechanism acquires the contribution level from the profit-maximizing auction and selects the most profitable edge server and content delivery mode through equation (11). Therefore, it can be concluded that the mechanism is able to maximize the objective function, equation (9). In other words, the mechanism realizes the objective function $F(\theta)$.

In HRR mechanism, HRRS sends requests to the traditional DNS-based RR mechanism frequently. Therefore, the best possible place for the HRRS servers seems to be on the same physical machine or the same network of presence of the DNS-based traditional RR. If the HRR finds all the servers listed in TRRRS message saturated, it queries higher level DNS-based RR for potential servers. This case happens only when all the local edge servers are saturated; thus, it harms locality of traffic but improves reliability of the system and increases the potential of the system to overcome flash crowds.

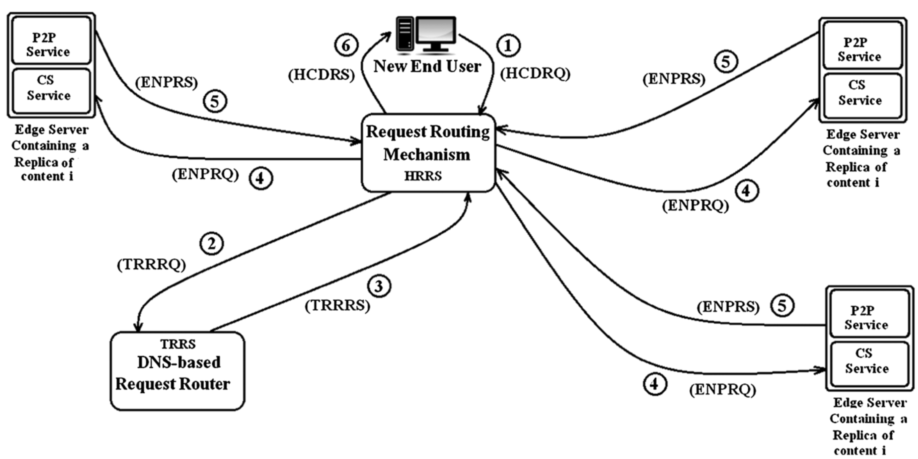


Figure 5. Request Routing Mechanism of Hybrid CDN-P2P network (HRR mechanism). P2P, peer to peer; CS, client–server; ENPRS, expected net profit response; ENPRQ, expected net profit request; TRRRQ, traditional request routing request; DNS, domain name system; TRRRS, traditional request routing response; HRRS, hybrid request routing server; HCDRQ, high content distribution request

DOI: 10.1002/nem

Figure 5 depicts that HRR, as a distributed mechanism, uses 10 messages to route a request. Because messages transferred for steps 4 and 5 are sent concurrently, only six round-trip times are sufficient to route a request in HRR.

5. PERFORMANCE EVALUATION

5.1. Experimental setup

We have employed a custom-written simulation test bed to study the performance of HRR. The discrete event simulation is written in C++, compiled using Borland C++ 5.02 compiler, and runs on a Windows machine with two 3.2 GHz processing cores and 4 GB of RAM. Because RR mechanism in HRR demands the strategic decision of the end-users, and it is not possible to have thousands of real users simultaneously, their behaviour is also simulated and fed to the test bed. Although the HRR mechanism is theoretically able to solve the problem in any environment, maximum effort has been made to set the parameters as real as possible. Some of the parameters of the environment (e.g. sequence and source address of requests and location of edge servers and replicas) are set using the real log files of a traditional CDN during the month of May 2013, operating in Australia. Other parameters are assigned based on approximate distribution of parameters in the geographical domain of the referenced CDN [36].

Among the parameters of end-users, the source IP address and the sequence of requests are extracted from the CDN's trace log files. Because measurement of different features of users is not in domain of this research, some simplifying assumptions are made to set their private values. The upload and download capacities and their respective prices are assumed as random variables having normal distribution with standard deviation equal to one-fourth mean. Values of preferences of users are also set using the same approach.

A typical 1000 s streaming content with bandwidth of 500 kbps (medium quality 240p with H.264 encoding) is considered as the CS service [37]. Bandwidth of P2P content is assumed as double of CS service, 1 Mbps (480p with H.264 encoding) [37]. Price of bandwidth for CS service is extracted from the SLA of the CDN, and in order to keep the SLA incentive compatible, the price of bandwidth for P2P service is assumed as one-fourth of CS bandwidth price. Considering equation (1), cost of P2P service per user in the SLA for the content provider is half of the CS service. It is assumed that 10 edge servers have replicas of the stream. These servers are able to distribute the content in both P2P and CS

Table 3. Summary of parameters used in experiments

Parameter	Value	Comment
$BW_{CS,i}$	0.5 Mbps	Bit rate of an average quality video
$BW_{P2P,i}$	1 Mbps	Bit rate of a good quality video
$P_{CS,i}$	\$80/Gbps	Price of traditional CDN service in the SLA
$P_{P2P,i}$	\$20/Gbps	To keep the SLA incentive compatible, the price of CS service is assumed to be two times of the P2P service price per user
$BWPrice_{Edge\_Server}$	[60, 76] \$/Gbps	Network bandwidth price at different edge servers
$Price_{User,D}$	\$100/Gbps	Average price of download bandwidth, a normal distribution with standard deviation of one-fourth average is used in the simulation
$Price_{User,U}$	\$200/Gbps	Average price of upload bandwidth, a normal distribution with standard deviation of one-fourth average is used in the simulation
DLC_{User}	4 Mbps	Average of download bandwidth in geographical domain of the referenced CDN
ULC_{User}	1 Mbps	Average of upload bandwidth in geographical domain of the referenced CDN

CDN, content delivery network; SLA, service-level agreement; P2P, peer to peer; CS, client–server.

DOI: 10.1002/nem

modes. One HRRS server is responsible to perform RR for this network. Table 3 represents values of parameters used in experiment.

In our experiments, a traditional DNS-based RR returns address of three nearest servers for each request. The administrator at the beginning of content delivery service assigns values of the PDF and CDF of the private values of end-users. After that, the number of requests became enough, and before bootstrapping of P2P network, the HRRS gradually calculates the mean and variance of bids to have good estimates of PDF and CDF functions used in the Myerson mechanism. In order to keep the experiments comparable, churn rate of end-users is not considered. In other words, the users enter the system in an accumulative manner to enable us to study the behaviour of the HRR increasing the load without worrying about the churning rate. As number of replicas compared with number of end-users is negligible, we have excluded replication cost from cost calculations. When the simulation starts, users request for the content, one by one, receive the hybrid content distribution response message and make strategic decision on joining P2P network or receiving CS content.

5.2. Experiments

In order to study the performance of HRR, we have studied its outcome in four different experiments. First, we have studied the behaviour of HRR under a real workload. Next, we have compared outcome of HRR with other strategies. At the third experiment, we have studied the effect of parameters of the contracted SLA in the behaviour of HRR and end-users. The last experiment is devoted to studying the performance and economics of hybrid content distribution under future conditions.

5.2.1. Normal operation

In the first experiment, 2000 requests are submitted to the system. Distribution of the end-users among edge servers and economics of HRR are reported in Table 4.

The two last rows of the table show the cost and revenue of the system assuming a pure CS content delivery and enable us to compare the economics of the HCDN with a traditional CDN. The last

Table 4. Summary of results for serving 2000 end-users

Edge server ID	1	2	3	4	5	6	7	8	9	10	Total
Bandwidth price ($/Gbps)	76	60	62	64	74	66	68	76	70	68	68
Number of end-users	70	478	446	227	117	213	130	66	122	131	2000
Number of P2P end-users	61	391	376	190	106	188	105	58	103	104	1682
Number of CS end-users	9	87	70	37	11	25	25	8	19	27	318
α	0.16	0.16	0.156	0.176	0.187	0.159	0.15	0.182	0.142	0.185	0.164
Bandwidth spent by the server (Mbps)	14.27	106.1	93.92	52.02	25.3	42.37	28.25	14.57	24.15	32.8	433.8
Bandwidth delivered to end-users (Mbps)	65.5	434.5	411	208.5	111.5	200.5	117.5	62	112.5	117.5	1841
Bandwidth contribution of end-users (Mbps)	51.23	328.4	317.07	156.47	86.15	158.12	89.25	47.425	88.35	84.7	1407.2
Cost in hybrid mode ($)	1.075	6.375	5.825	3.325	1.875	2.8	1.925	1.1	1.7	2.225	28.225
Revenue in hybrid mode ($)	1.58	11.3	10.33	5.28	2.55	4.75	3.1	1.48	2.825	3.15	46.35
Cost in pure CDN mode ($)	2.65	14.35	13.825	7.275	4.325	7.025	4.425	2.5	4.275	4.45	65.1
Revenue in pure CDN mode ($)	2.8	19.125	17.85	9.075	4.675	8.525	5.2	2.65	4.875	5.25	80

CDN, content delivery network; CS, client–server; P2P, peer to peer.

DOI: 10.1002/nem

column of the Table 4 shows aggregated outcome of HRR mechanism for all the edge servers. Based on these results, we can see that although most of the users prefer to receive the content with higher quality using P2P network and contribute in content distribution, 19.05% of them (381/2000) strategically decided to received the CS content. These CS end-users did not find it possible (14.9% of them due to their downlink capacity) or economic (4.15% of them due to their preference function and upload bandwidth price) to contribute in P2P content delivery and receive higher-quality content. Considering the average value of α reported at the last column of Table 4, the HCDN affords only 16.4% of bandwidth delivered in P2P network. The HCDN has spent 433.8 Mbps upload capacity, P2P end-users have provided 1407.2 Mbps bandwidth and the system succeeded to deliver a total of 1841 Mbps bandwidth to 2000 end-users.

In this experiment, although the price of service (in the SLA) for a P2P end-user receiving double bandwidth is half of a CS end-user, HRR has succeeded to increase the net profit of pure CDN mode from (80–65.1 = \$14.9) to (46.35–28.25 = \$18.125) in hybrid mode. This observation reveals the win–win situation for the content provider who has to pay \$46.35 instead of \$80, and the HCDN earns \$18.125 instead of \$14.9 and also for 83.6% (1682/2000) of end-users who received the content with double quality. It is also observable that servers with lower bandwidth price, that is, servers 2 and 3, succeeded to serve more users, and it is easy to find out that when number of P2P end-users at an edge server increases, it becomes able to absorb more end-users and succeeds to offer the P2P service to more end-users.

5.2.2. *Comparison of hybrid request routing with other strategies*

In order to study the performance of HRR, we have implemented three other RR strategies. These strategies include the following:

- Non-oligopolistic RR with the Myerson auction. The Myerson auction is moved to edge servers, and users are free to choose one of handful nearby edge server strategically.
- Oligopolistic RR with Vickrey auction. Instead of the Myerson approach, a Vickrey auction receives requests and determines the contribution level using second-price auction.
- Pure CDN where the P2P content delivery is shut down and all requests are served in CS mode.

As shown in Figure 6, HRR is able to provide more net profit compared with all other approaches. This experiment implies that both maximization of end-user's contribution and economical balance of the load on edge servers are necessary to reach the maximum profit for the HCDN. Application of the Myerson mechanism even without oligopoly gives more net profit compared with Vickrey truthful auction. This observation verifies that Vickrey and all the classes of VCG truthful auctions are not able to maximize the net profit and just benefit social welfare of the end-users [27]. In fact, class of VCG mechanisms is able to reveal the private values of end-users, but they are not able to benefit from this information, maximize the net profit and maximize the total payment of end-users (in the form of upload bandwidth). As expected, the lowest outcome belongs to pure CS strategy.

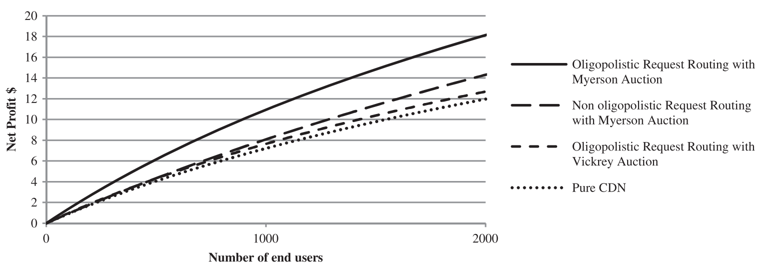


Figure 6. Comparison of net profit in different mechanisms varying number of end-users.

DOI: 10.1002/nem

5.2.3. Effect of parameters of service-level agreement in hybrid request routing

The third experiment is devoted to study the effect of variation of parameters of the SLA in economics of content distribution. These parameters are varied in two scenarios: price of P2P service is varied, while the price of CS service is kept constant, and bandwidth of the P2P content is varied against the bandwidth of CS content, while the bandwidth of CS service is kept steady. Two thousand requests are submitted to the system, and effect of the variation in the SLA parameters on both economics of the hybrid content distribution and decisions made by the end-users is reported.

Figure 7 illustrates the effect of variation of price ratio, $P\_\mathrm{ratio} = P_{\mathrm{P2P}}/P_{\mathrm{CS}}$, on the number of requests served in P2P mode when P_{CS} is kept constant. Effects of this variation on cost of the content distribution for the content provider (revenue of HCDN), the cost of content distribution for HCDN and the net profit of HCDN are illustrated in Figure 8.

Figures 7 and 8 represent that variation of the P2P traffic price against price of the CS service does not affect the contribution of end-users. When the price of the P2P service is below a threshold ($0.21 \times P_{\mathrm{CS}}$), all the requests are served in CS mode. Therefore, the cost of content distribution for the content provider (revenue of the HCDN) and cost of the content distribution for HCDN have remained constant. After the threshold, P2P content distribution becomes economic for HCDN, and as a result of contribution of end-users, content distribution cost for HCDN drops. Effect of this sudden decrease of content distribution cost on the revenue function is also visible. After the threshold, the revenue and the net profit of HCDN increase linearly. So, below this threshold, the HCDN loses its incentive to distribute the content in P2P mode, and just above this threshold, the content is delivered in hybrid mode. Practically, this threshold can determine the minimum price of P2P service in hybrid content delivery.

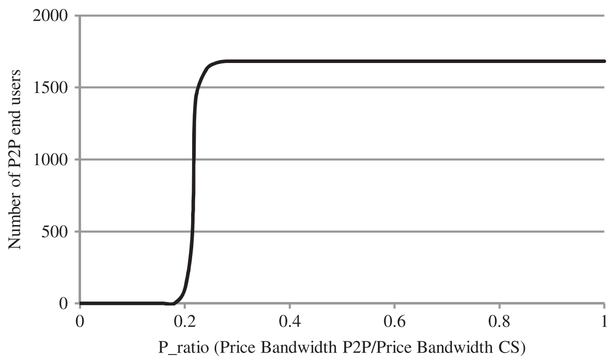


Figure 7. Effects of variation of Price of P2P bandwidth on number of P2P end users

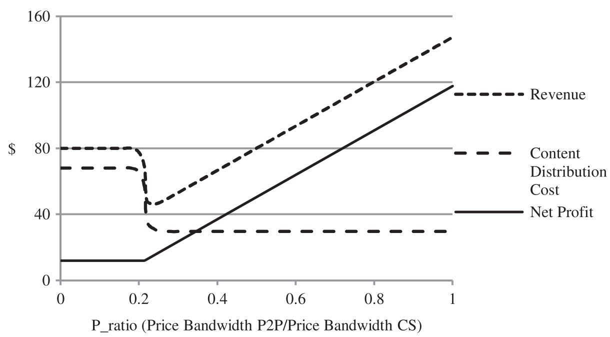
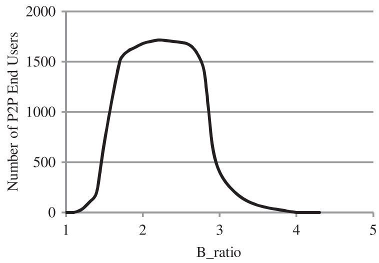
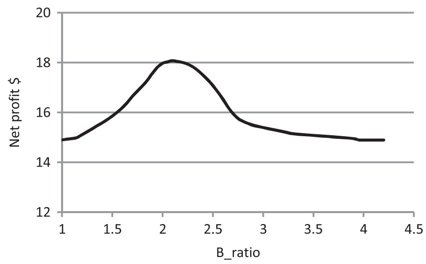


Figure 8. Effects of variation of Price of P2P bandwidth on economics of HCDN

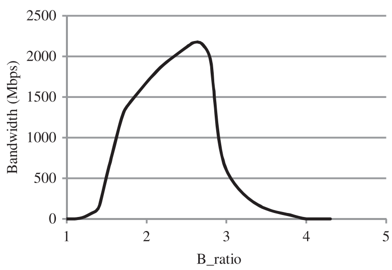
DOI: 10.1002/nem



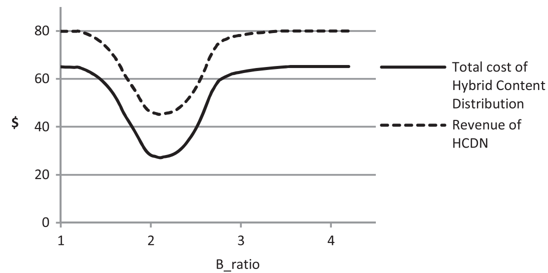
a. Number of P2P end users varying Bandwidth of P2P stream



c. Net profit of HCDN varying Bandwidth of P2P stream



b. Bandwidth contribution of P2P end users varying Bandwidth of P2P stream



d. Revenue and cost of HCDN varying Bandwidth of P2P stream

Figure 9. (a) Number of peer-to-peer (P2P) end-users varying bandwidth of P2P stream, (b) bandwidth contribution of P2P end-users varying bandwidth of P2P stream, (c) net profit of hybrid content delivery network–peer-to-peer network (HCDN) varying bandwidth of P2P stream, and (d) revenue and cost of HCDN varying bandwidth of P2P stream

Figure 9(a–d) illustrates effects of the variation of bandwidth ratio (B_ratio = BW_{P2P}/BW_{CS}) on the economics and performance of the HCDN.

Figure 9(a) depicts that increasing the quality of P2P content motivates more end-users to contribute their bandwidth into the P2P networks and receive P2P content. With settings given in Table 3, when the bandwidth of the P2P content is 2.235 times more than CS content, the number of P2P end-users and the net profit of HCDN are maximized as illustrated in Figure 9(a and c, respectively). At this point also, the content distribution cost for the content provider (revenue of HCDN) is minimized. Figure 9(b) shows that maximum total bandwidth contribution of end-users does not necessarily happen when the number of P2P end-users is maximized. Instead, when the quality of P2P content increases, end-users with limited upload capacity become unable to satisfy requested upload contribution of the HCDN, but end-users with high-upload capacity find it beneficial to contribute more and receive higher-quality content. Increasing B_ratio forces the HCDN to afford a bigger fraction of P2P content delivery bandwidth and decreases its net profit. Because of the fact that upload capacity of the end-users is limited, with further increase of B_ratio, the HCDN loses its incentive to deliver the P2P service, its net profit decreases and it makes the HCDN offer just CS service, like a traditional CDN.

5.2.4. The performance of hybrid request routing under future conditions

Because of the increasing trend of the end-user access bandwidth and streaming video quality, major competitors of content delivery market are getting ready for distribution of very high bandwidth content. In future, it would be common to have the streaming video content with quality as high as 1080p, 2160p or even 4320p. In addition, in the future, the average access bandwidth of end-users will be improved several times. In this test, the performance of HRR mechanism is studied under future conditions.

Considering the year 2018 as the future reference point, the average access bandwidth of the end-users, their bandwidth price and parameters of the SLA are scaled considering the global [38] and local [39] forecasts of the network performance and economics. Based on these forecasts, the average access bandwidth is assumed to be 42.5 Mbps [38]. The price of content delivery

DOI: 10.1002/nem

Table 5. The performance of HRR under future conditions

Encoding standard	VP9			h.264		
Resolution of CS content	720p	1080p	1440p	480p	720p	1080p
Resolution of P2P content	1080p	1440p	2160p	720p	1080p	1440p
Bandwidth of CS content (Mbps)	1.378	2.182	6.522	1.0	2.5	4.5
Bandwidth of P2P content (Mbps)	2.182	6.522	10.4	2.5	4.5	8.0
Number of P2P end-users	1841	1735	1689	1928	1845	1733
Number of CS end-users	159	265	311	72	155	267
α	0.168	0.185	0.162	0.169	0.178	0.173
Bandwidth spent by the HCDN (Gbps)	0.894	2.672	4.873	0.887	1.865	3.600
Bandwidth delivered to end-users (Gbps)	4.236	11.894	19.593	4.892	8.690	15.065
Bandwidth contribution of end-users (Gbps)	3.342	9.222	14.719	4.005	6.824	11.456
Cost in hybrid mode ($)	20.263	60.556	110.47	20.095	42.281	81.599
Revenue in hybrid mode ($)	32.623	90.857	171.193	34.053	65.683	124.466
Net profit in hybrid mode ($)	12.359	30.3	60.716	13.957	23.402	42.867
Cost in pure CDN mode ($)	62.469	98.917	295.664	45.333	113.33	204.00
Revenue in pure CDN mode ($)	73.493	116.373	347.84	53.333	133.33	240.00
Net profit in pure CDN mode ($)	11.024	17.456	52.176	8	20	36
Improvement of the net profit in HCDN compared with traditional CDN	37.9%	33.3%	35.4%	40.9%	35.9%	34.4%

service and bulk bandwidth are scaled down with respect to the improvement in the average access bandwidth. In order to choose the video qualities of the future HCDN, a wide spectrum of video qualities from 720p to 2160p is used. It is assumed that the video is encoded with Google's new VP9 video codec [40] and popular h.264 codec. As it can be apparently seen in Table 5, qualities of CS and P2P content are set as (720p and 1080p), (1080p and 1440p) and (1440p and 2160p) for VP9 codec and (480p and 720p), (720p and 1080p) and (1080p and 1440p) for h.264 codec. In this test, other parameters of the environment, like the topology of the underlying network, total number of end-users and the duration of the streaming content, are kept consistent with those of the previous experiments. Table 5 reports the performance and economics of the HCDN architecture with HRR mechanism under new conditions and for different video qualities. The results of the experiment show that HRR has always succeeded to benefit from the bandwidth contribution of the end-users (α). As a result, the content provider always has to spend less (compared with the traditional CDN economy), the HCDN provider has a positive and considerably large profit margin and a big fraction of end-users enjoy the stream of video with higher quality. The last three rows of Table 5 report the cost and revenue of a traditional CDN economy and compare the net profit of HCDN architecture with the traditional CDN. This experiment vividly shows that the HRR is able to keep the HCDN architecture more profitable compared with the pure CDN economy even in the future.

5.3. *Practical implications*

Our experiments depict that, employing an economic RR and using the hybrid content delivery approach, it is possible to benefit all the three sides of the content distribution economy. In other words, the economic RR mechanism enables the HCDN to provide a big fraction of end-users with higher-quality content, increase the profit of HCDN provider and decrease the expenditures of content provider at the same time. Therefore, there is a high economic potential for CDNs to offer HCDN services and improve their profit.

These experiments also show that the parameters of the SLA between HCDN and the content provider must be adjusted with special care. In other words, although the price of P2P bandwidth must be less than the price of CS bandwidth, it must be more than a special threshold to incentivize the HCDN to offer the hybrid service. This threshold affects from the private value of end-users, and in this study, we showed that the minimum value of P_{P2P} at the SLA must be at least $0.21 * P_{CS}$.

DOI: 10.1002/nem

Bandwidth of P2P content must be large enough to incentivize the end-users to contribute in the P2P content delivery. We showed that in this environment, for a wide spectrum of the P2P QoS, the economic RR for hybrid content delivery can increase the profit significantly. Also, we observed that in this domain, the net profit of HCDN compared with the total bandwidth contribution of P2P end-users influences more from population of the P2P end-users. In other words, increase of quality of P2P service can maximize total bandwidth contribution of the end-users, but this maximum point does not necessarily maximize the net profit of the HCDN because at this point the HCDN has to spend more bandwidth to incentivize a relatively small group of end-users to contribute in the P2P content delivery.

In order to verify the scalability of HCDN architecture and HRR mechanism, some simulation studies performed. In these experiments parameters of the underlying network scaled considering the forecasts of global network growth. In addition in these experiments the effect of new video encoding standards on the performance of HRR studied. Results vividly show that HRR mechanism is able to guaranty the performance and economics of content distribution in HCDN, even in the future.

6. CONCLUSION AND FUTURE WORK

We introduced a new RR mechanism for the streaming content distribution in HCDNs. Design of the mechanism performed considering both performance and economic constraints. Maximization of the net profit of the CDN owner is set as the optimization goal, and we showed that the problem consists of two isolated decisions, the optimum requested bandwidth contribution and the optimum assignment of end-users to edge servers.

Experimental results demonstrate that application of HRR makes it possible to serve a big fraction of end-users with higher-quality content, improve the net profit of HCDN provider and at the same time decrease the expenditures of content provider. In order to evaluate the performance of HRR, some other strategies including Vickrey second-price auction were implemented, and experiments show that compared with all other strategies, HRR has the best outcome.

Effect of the SLA parameters in the economics and performance of hybrid content distribution was studied, and we have also showed that parameters of the SLA must be determined by considering incentives of the content provider, the HCDN provider and end-users.

Scalability of the HCDN architecture and HRR mechanism was studied under future conditions, and we showed that HRR remains effective even when the access bandwidth improves and the quality of the video stream increases several times.

Content delivery networks optimize their performance by relying on two key mechanisms: RR and RP. In the future, we will consider RP mechanism for streaming content of hybrid CDN–P2P architecture. Because the HCDN serves the users with either CS or P2P mode, the RP mechanism of an HCDN has a fundamental difference with RP in traditional CDNs. This mechanism in addition to challenges of RP in traditional CDNs will optimize two sets of edge servers (CS and P2P) to replicate the content in CS and P2P modes.

REFERENCES

1. Hao Y, Liu X, Zhan T, Sekar V, Qiu F, Lin C, Zhang H, Li B. Design and deployment of a hybrid CDN–P2P system for live video streaming: experiences with LiveSky, in *Proceedings of the 17th ACM international conference on Multimedia*. ACM, 2009, pp. 25–34.
2. Pathan M, Buyya R. A taxonomy of CDNs. In Content Delivery Networks. Springer: Berlin Heidelberg, 2008; ch. 2, pp. 33–77.
3. Pallis G, Vakali A. Insight and perspectives for content delivery networks. *Communications of the ACM* 2006; **49**(1): 101–106.
4. Buyya R, Pathan M, Vakali A. Content Delivery Networks. Springer: Berlin, 2008.
5. CardelliniV, Colajann M, Yu PS. DNS dispatching algorithms with state estimators for scalable Web-server clusters. *World Wide Web* 1999; **2**(3): 101–113.

DOI: 10.1002/nem

6. Wang W, Li B. Market-driven bandwidth allocation in selfish overlay networks, in Proceedings of INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies, 2005; pp. 2578–2589.
7. Michal F, Chuang J. Overcoming free-riding behavior in peer-to-peer systems. *ACM SIGecom Exchanges* 2005; **5**(4): 41–50.
8. Ramayya K, Smith MD, Tang Z, Telang R. The impact of free-riding on peer-to-peer networks, in Proceedings of IEEE 37th Annual Hawaii International Conference on In System Sciences, Hawaii, 2004; pp. 1–10.
9. Myerson RB. Optimal auction design. *Mathematics of Operations Research* 1981; **6**(1): 58–73.
10. Gans J, King S, Mankiw GN. Principles of Microeconomics. Second Edition, Cengage Learning: Melbourne, 2012.
11. Huang C, Wang A, Li J, Ross KW. Analysis of a CDN–P2P hybrid architecture for cost-effective streaming media distribution. *Multimedia Systems* 2006; **11**(4): 383–399.
12. Ha I, Wildman SS, Bauer JM. P2P, CDNs, and hybrid networks: the economics of Internet video distribution. *International Telecommunications Policy Review* 2010; **17**(4): 1–22.
13. Hai J, Li J, Li Z, Bai X. Performance evaluation of content distribution in hybrid CDN–P2P network, in Proceedings of IEEE Future Generation Communication and Networking, FGCN08, 2008; pp. 188–193.
14. ZhiHui L, Wang Y, Yang Yang R. An analysis and comparison of CDN–P2P hybrid content delivery system and model. *Journal of Communications* 2012; **7**(3): 232–245.
15. Lu ZH, Gao XH, Huang SJ, Huan Y. Scalable and reliable live streaming service through coordinating CDN and P2P, in proceedings of IEEE 17th International Conference on Parallel and Distributed Systems (ICPADS), 2011; pp. 581–588.
16. Florescu B, Andreica MI. Towards a peer-assisted content delivery architecture, in Proceedings of 18th International Conference on Control Systems and Computer Science (CSCS), 2011; pp. 521–528.
17. Bronzino F, Gaeta R, Grangetto M, Pau G. An adaptive hybrid CDN/P2P solution for content delivery networks, in Proceedings of IEEE Visual Communications and Image Processing (VCIP), 2012; pp. 1–6.
18. Chao H, Chen M, Xing C, Xu B. EUE principle of resource scheduling for live streaming systems underlying CDN–P2P hybrid architecture. *Peer-to-Peer Networking and Applications* 2012; **5**(4): 312–322.
19. Yin H, Liu X, Zhan T, Sekar V, Qiu F, Lin C, Zhang H, Li B. LiveSky: enhancing CDN with P2P. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)* 2010; **6**(3): 1–16.
20. Khalaji FK, Analoui M. Replica placement algorithms in hybrid CDN–P2P architectures, in IEEE Sixth International Conference on Telecommunications, 2012; pp. 771–775.
21. Garmehi M, Analoui M, Pathan M, Buyya R. An economic replica placement mechanism for streaming content distribution in hybrid CDN–P2P networks. *Elsevier Computer Communications* 2014; **52**: 60–70.
22. Cramton P, Palfrey TR. Cartel enforcement with uncertainty about costs. *International Economic Review* 1990; **31**: 17–47.
23. Wang W, Li B. To play or to control: a game-based control-theoretic approach to peer-to-peer incentive engineering, in Springer, In Quality of Service—IWQoS 2003, Berlin Heidelberg, 2003; pp. 174–192.
24. Manzato DAG, da Fonseca NLS. Incentive mechanisms for cooperation peer-to-peer networks. In Handbook of Peer-to-peer Networking. Springer: Heidelberg, 2010; pp. 631–660.
25. Garmehi M, Analoui M. A distributed mechanism for economic management of transmission infrastructure in hybrid CDN–P2P networks. *Journal of Economic Computation and Economic Cybernetics Studies and Research (ECECSR)* No. 3/2014; 287–308.
26. Xiao S, Dhaliwal SK. Incentive mechanisms in P2P media streaming systems. *IEEE Internet Computing* 2010; **14**(5): 74–81.
27. Noam N, Roughgarden T, Tardos E. Algorithmic Game Theory. Cambridge University Press: Cambridge, 2007.
28. Oksana L, Lu H, Wang HX. Incentive schemes in peer-to-peer networks. *The BE Journal of Theoretical Economics* 2009; **9**(1): 1–20.
29. Rothkopf MH. Thirteen reasons why the Vickrey–Clarke–Groves process is not practical. *Operations Research* 2007; **77**(2): 191–197.
30. Chuan W, Li Z, Qiu X, Lau F. Auction-based P2P VoD streaming: incentives and optimal scheduling. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)* 2012; **8**(1): 1400–1422.
31. Yang G, Chen Y, Liu KJ. A cheat-proof game-theoretic framework for cooperative peer-to-peer video streaming, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012; pp. 2297–2300.
32. Hartline JD, Roughgarden T. Simple versus optimal mechanisms, in Proceedings of 10th ACM conference on Electronic Commerce, 2009; pp. 225–234.
33. Hurvics L, Reiter S. Designing Economic Mechanisms. Cambridge University Press: Cambridge, 2006.
34. Kartik H, Krishnan R, Chuang J, Choudhary V. Pricing and resource allocation in caching services with multiple levels of quality of service. *Management Science* 2005; **51**(12): 1844–1859.
35. Su A-J, Choffnes DR, Kuzmanovic A, Bustamante FE. Drafting behind Akamai: inferring network conditions based on CDN redirections. *IEEE/ACM Transactions on Networking* 2009; **17**(6): 1752–1765.
36. Cnet. Australia's average internet speed, 2013, Jul 12th [Online]. http://www.cnet.com/au/news/australias-average-internet-speed-now-4-7mbps/
37. YouTube. Live streaming guide, live encoder settings, bitrates and resolutions, 2014, Feb. 9th [Online]. https://support.google.com/youtube/answer/2853702?hl=en
38. Index, Cisco Visual Networking. The zettabyte era—trends and analysis. Cisco white paper, 2013.
39. Kenny R, Broughton T. Domestic bandwidth requirements in Australia, A forecast for the period 2013-2023, Technical Report, Communication Chambers, 2014.
40. Felina A. 2015, Apr. 8th animafelina.it. [Online]. http://www.animafelina.it/test-video-quality-720p-1080p-1440p-2160p-max-bitrate-which-compresses-youtube/

DOI: 10.1002/nem

AUTHORS' BIOGRAPHIES

Mehran Garmehi received his BS degree in computer engineering in Shahid Beheshti Tehran University, Tehran and the M.S. degree in computer architecture major from Sharif University of Technology Tehran. He is currently a PhD. Candidate in Computer Networks research laboratory at Iran University of Science and Technology under supervision of Dr Analoui. He has been a lecturer in the Faculty of Engineering, Bojnurd University, Iran. He has authored and co-authored a number of research papers in internationally recognized journals and conferences. His main research interests include content distribution and delivery, caching and replication in computer networks. Email address: Garmehi@iust.ac.ir

Dr. Morteza Analoui received a B.A. degree in Electrical Engineering from Iran University of Science and Technology (IUST) in 1983 and a Ph.D. degree in Electrical Engineering from Okayama University, Japan, in 1989. He is an Associate University Professor (and past Dean) in the College of Computer Engineering and the director of Computer Network Research Lab in IUST. He has been an assistant professor in the Faculty of Engineering, Tarbiat Modaress University, 1994 and in the Engineering Faculty of Okayama University from 1990 to 1993. His research interests include Networking (High Speed, Virtual Social, Virtual Economic, CDN, P2P), Optimization Theories and Techniques. Dr. AnaLoui has authored over 150 publications, including journal articles, book chapters, and conference papers. Email address: Analoui@iust.ac.ir

Dr. Mukaddim Pathan is a Senior Media Specialist at Telstra Corporation Limited, the largest Telecommunications and Media Company in Australia. He also holds the position of an honorary fellow at the University of Melbourne and serves as a guest lecturer at Deakin Univeristy, RMIT University and the University of Melbourne. Prior to joining Telstra, he worked as a Research Scientist at Australian Government's research organization, CSIRO. He holds a PhD in Computer Science from the University of Melbourne, Australia. His research interests include resource management, origin storage, load balancing, and coordination policies in wide-area distributed systems such as CDNs and Cloud computing. He is a key contributor in Telstra's Next Generation CDN and media infrastructure, delivering live streaming, Video-on-Demand, online gaming and lifestyle content for Foxtel, BigPond and enterprise customers in Australia. He is the editor of the books Content Delivery Networks and Advanced Content Delivery, Streaming, and Cloud Services. He has authored and co-authored a number of research papers in internationally recognized journals and conferences. He is involved in the organization of several international conferences and serves as a reviewer and/or editorial board member of a few renowned journals. His website: http://www.cloudbus.org/cdn/mpathan Email address: mukaddim.pathan@team.telstra.com

Dr. Rajkumar Buyya is a Fellow of IEEE, Professor of Computer Science and Software Engineering, Future Fellow of the Australian Research Council, and Director of the Cloud Computing and Distributed Systems (CLOUDS) Laboratory at the University of Melbourne, Australia. He is also serving as the founding CEO of Manjrasoft, a spin-off company of the University, commercializing its innovations in Cloud Computing. He has authored over 450 publications and five text books including "Mastering Cloud Computing" published by McGraw Hill and Elsevier/Morgan Kaufmann, 2013 for Indian and international markets respectively. He also edited several books including "Cloud Computing: Principles and Paradigms" (Wiley Press, USA, Feb 2011). He is one of the highly cited authors in computer science and software engineering worldwide (h-index=87, g-index=176, 38000+ citations). Microsoft Academic Search Index ranked Dr. Buyya as the world's top author in distributed and parallel computing between 2007 and 2012. Software technologies for Grid and Cloud computing developed under Dr. Buyya's leadership have gained rapid acceptance and are in use at several academic institutions and commercial enterprises in 40 countries around the world. Dr. Buyya has led the establishment and development of key community activities, including serving as foundation Chair of the IEEE Technical Committee on Scalable Computing and five IEEE/ACM conferences. These contributions and international research leadership of Dr. Buyya are recognized through the award of "2009 IEEE TCSC Medal for Excellence in Scalable Computing" from the IEEE Computer Society TCSC. Manjrasoft's Aneka Cloud technology developed under his leadership has received "2010 Frost & Sullivan New Product Innovation Award" and recently Manjrasoft has been recognised as one of the Top 20 Cloud Computing companies by the Silicon Review Magazine. He is currently serving as Co-Editor-in-Chief of Journal of Software: Practice and Experience, which was established 40+ years ago. For further information on Dr. Buyya, please visit his cyberhome: www.buyya.com

DOI: 10.1002/nem