Regular Paper

Relaxing CDN Energy: Considering Peer-Assisted

Mohamad Dikshie Fauzie[1,a)]   Achmad Husni Thamrin[1,b)]   Jun Murai[1,c)]

Abstract:
This paper presents an analysis of a peer-assisted CDN system, where an ISP manages its own CDN and its users participate in a P2P network to assist content delivery. The system consists of a two level of peer where first level can be considered as direct client from CDN (seeders) and second level is peer that get data from first level (leechers). We show that depend on number of peers, upload rate of peer in first level, and capacity of CDN, CDN energy can be reduced when peer in first level contribute to second level. Moreover bigger upload rate of peers in first level, then bigger peers in second level can be served by peers in first level thus reducing CDN work load.

Keywords: CDN, Data Center, Energy, Peer Assisted, P2P

1. Introduction

Streaming content, especially video, represents a significant fraction of the traffic volume on the Internet, and it has become a standard practice to deliver this type of content using Content Delivery Networks (CDNs) such as Akamai and Limelight for better scaling and quality of experience for the end users. For example, Youtube uses Google cache and MTV uses Akamai in their operations.

With the spread of broadband Internet access at a reasonable flat monthly rate, users are connected to the Internet 24 hours a day and they can download and share multimedia content. P2P (peer to peer) applications are also widely deployed. In China, P2P is very popular; we see many P2P applications from China such as PPLive, PPStream, UUSe, Xunlei, etc. Some news broadcasters also rely on P2P technology to deliver popular live events. For example, CNN uses the Octoshape solution that enables their broadcast to scale and offer good video quality as the number of users increases.

From the Internet provider point of view, the presence of so many always-on users suggests that it is possible to delegate a portion of computing, storage and networking tasks to the users, thus creating P2P networks where users can share files and multimedia content. Starting from file sharing protocols, P2P architectures have evolved toward video on demand and support for live events.

Broadband network access helps P2P applications to perform better. xDSL networks are deployed worldwide, and in some countries, such as Japan, even higher bandwidth fiber to the home (FTTH) already exceeds DSL in market penetration. In the coming years, FTTH will be massively deployed by network operators throughout the world. As access bandwidth increases, P2P systems may become more efficient since a peer can contribute much more.

In Peer assisted CDN, users can download content from CDN nodes for from other users or peers the content. A user may cache the content after download to serve requests from other users Due to complexity of behavior of peers, the process should be done in home gateway user where ISP can have control on it.

On the other hand, data center where CDN server placed faces costs for powering the data center. The Uptime Institute, a global data center authority, surveyed 1100 data center owners and operators on 2012, reported that 55% organizations must increase budget 10% than 2011 [1]. 30% of organizations will run out of data center capacity (power, cooling, space, and network) in the end of 2012 [1]. More than 50% organizations surveyed reported that saving energy is major priority [1]. The increases in energy cost and the demand due to growth of traffic urges the data center operators and owners to look for ways to reduce energy usage in the years to come. Although reducing energy consumption can effectively reduce overall cost, this will limit the capacity growth and scalability of the service provisioning. Alternatively, the data center and be revamped by relocating some services to end-host computers or peers. Peers contribute their communication, storage, and computation resources to exchange data and provide services while the data center performs central administration and authentication as well as backend processing. P2P network, formed by peers offer flexibility and scalability in service delivery. Therefore, P2P services can assist and enhance data center. It is not our aim to advocate one system architecture over another. Many issue such as manageability, reliability, and ease of deployment must be taken into account when making high level architectural decisions.
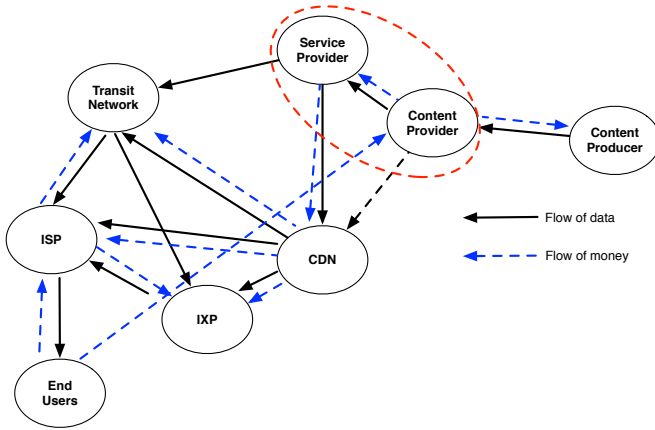
In this paper, we study the energy consumption of hybrid CDN-P2P. It has been known that CDN energy consumption is better than P2P architecture, unfortunately that model only think small

---

[1]   Keio University Shonan Fujisawa Campus, 5322 Endo, Fujisawa-shi, Kanagawa 252-0882 Japan
[a)]   dikshie@sfc.wide.ad.jp
[b)]   husni@ai3.net
[c)]   jun@wide.ad.jp

Fig. 1 In Complex relationship of entities in Internet, CDN mostly placed in data center near to eyeball ISP. If CDN can not reach eyeball ISP due to business or economic reason, CDN can be placed near to IXP or even inside IXP, and CDN will reach eyeball ISP from peering point inside IXP.

part of problem energy consumption (devices/hardwares). To be fair, we also need to change paradigm that to see if we also can relax power budget of data center by utilizing P2P. If we can move part of computation resource from CDN in data center to P2P, then we can relax power budget of data center for hardware and cooling.

The rest of this paper is organized as follows. Section 2 present motivation of this paper. Section 3 provides an overview of system model, data center model and energy model. In section 3, we will present result and analysis. Section 4 will conclude this paper.

2. Motivation

In the economic supply chain of video traffic, most ISPs get a little revenue from video traffic thus ISP wants to monetize that traffic. the future of CDN business is likely to live deeper into ISP networks, more integrated into and interleaved with ISP infrastructures thus users can get good quality of video. The idea of ISP managed CDN has been proposed in recent years. The complexity of the CDN business encourage ISP to manage their own CDN rather than allow others to run CDNs on their networks.

CDN architectures are host-oriented: content is delivered to end users through host servers that are centrally managed in a few data centers fig.1. The growing of Internet traffic dominated by video, the energy consumption of a host oriented architecture becomes problematic due to over provisioning factor. The idea of utilizing the user's computation power to support ISP operation is not new. The figaro project proposed residential gateway as an integrator of different networks and services, becoming an Internet wide distributed content management for a proposed future Internet architecture. In this case, ISP can offload part of workload on their CDN to user's home gateway. By offloading workload, their CDN can relax energy demand thus relaxing data center power budget or relaxing capacity planning for power.

3. System Description

Figure 2 shows typical CDN with peer assisted. P_s in fig.2
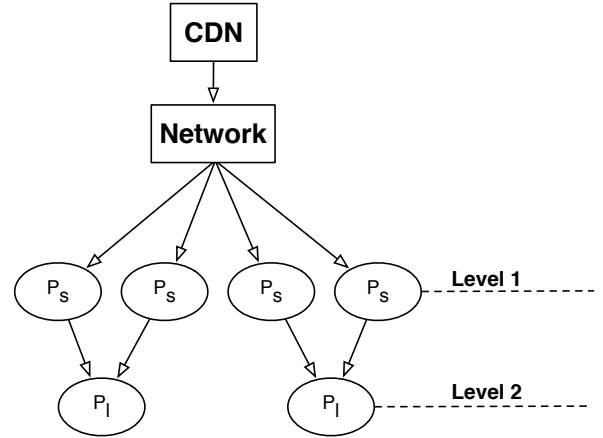
Fig. 2 A typical content delivery architecture with peer assisted.

is peers in level 1 or seeders. Seeders get content directly from CDN. P_l in fig.2 is peers in level 2 or leechers. Leechers get content from seeders. Maximum number of seeders is bounded by maximum CDN's capacity, while maximum number of leechers in level 2 is bounded by number of seeders can support the bitrate. Denote number of seeders is n_s, number of leechers is n_l, ρ is maximum bitrate that supplied by seeders to leechers, and $r = 1$ is video bitrate, therefore we have number of leechers that can be supported by seeders is:

$$\lfloor n_l \rfloor = n_s.\rho \tag{1}$$

Number of seeders that support or upload content to leecher is:

$$n_s^u = n_l.\frac{r}{\rho} \tag{2}$$

The illustration as follows, suppose we have video bitrate $r = 1$, seeder upload rate $\rho = 0.25$, and maximum CDN capacity is $643Mbps$. Maximum number of seeders supported by CDN is $n_s = 643$. Maximum number leechers supported by seeders is $n_l = 160$. Number of seeders that upload content to leechers is $n_s^u = 640$. Therefore we have three seeders that do not need to upload content to leecher.

3.1 Abstracting Data Center Power
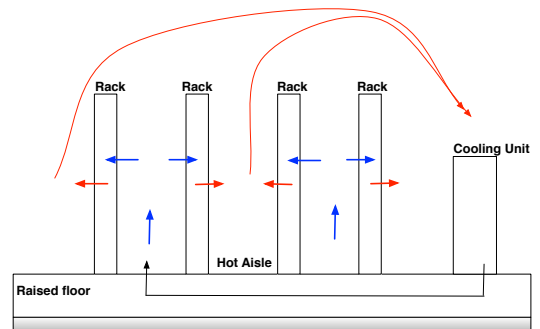
Fig. 3 A typical data center schematic. The cooling infrastructure comprises industry standard racks on a raised floor, through which multiple compressor-driven computer room air conditioning unit circulate cool air through a share plenum.

Many aspects involved in power management in data center

and interactions among power management strategies across subsystems grow more complex. Five distinct subsystems account for most of a data center's power draw: server and storage systems, power conditioning equipment, cooling and humidification systems, networking equipment, and lighting/physical security. In this section, we only focus on relationship between cooling power and hardware temperature.

Data center seek to provision the cooling adequately to extract the heat produce by servers, switches, routers, and other devices. Practically, cooling provisioning are done at facilities level. Data center operator operate cooling facilities based on power ratings of servers, switches, routers, and other devices often with some additional headroom for risk tolerance. The compounded over provisioning of computing and cooling infrastructures can drive up initial and recurring costs.

The cooling cycle of a typical data center operates in the following way: Cooling units operate by extracting heat from the data center and pumping cold air into the room, usually through a pressurized floor plenum. The pressure forces the cold air upward through vented tiles, entering the room in front the hardware Fans draw the cold air inward and through the server; hot air exits through the rear of the server. The hot air rises sometimes with the aid of fans and a ceiling plenum , and is sucked back to the cooling units. The cooling units force the hot air past pipes containing cold air or water. The heat from the returning air transfer through the pipes to the cold substance. The heated substance leaves the room and goes to a chiller and cooling unit fans force the cold air back into the floor plenum. Above process is shown in fig.3

The efficiency of this cycle depends on several factors, including the conductive substance and the air flow velocity, but it is quantified by coefficient of performance (COP). The COP is the ratio of heat removed (Q) to the amount of work necessary (W) to remove that heat:

$$COP = \frac{Q}{W} \tag{3}$$

Therefore, the work necessary to remove heat is inversely proportional to the COP. A higher COP indicates a more efficient process, requiring less work to remove a constant amount of heat.

However, the COP for a cooling cycling is not constant, increasing with the temperature of the air the cooling unit pushes into the plenum. COP value empirically can be computed using [2]:

$$COP(T) = 0.0068.T^2 + 0.0008.T + 0.458 \tag{4}$$

where $T = T_{sup} + T_{adj}$ and $T_{adj} = T_{safe}^{in} - T_{max}^{in}$.
T_{sup} is temperature supply by cooling unit and T_{adj} is temperature difference between maximum safe hardware inlet temperature (T_{safe}^{in}) and the maximum observed hardware inlet temperature (T_{max}^{in}). If T_{adj} is negative, it indicates that a hardware inlet exceeds maximum safe temperature thus we need to lower T_{sup} to bring the hardwares back below the system redline level.

Cooling cost can be calculated as [2] :

$$C = \frac{Q}{COP(T)} \tag{5}$$
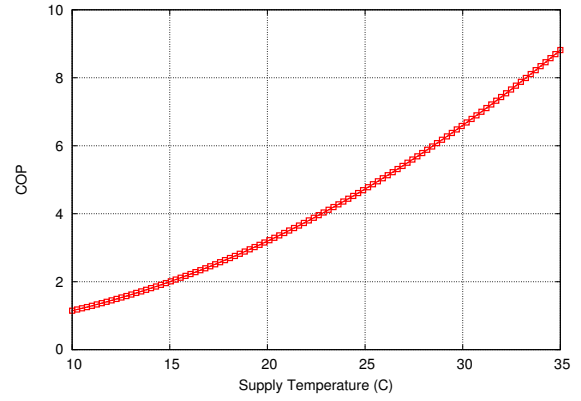
Fig. 4   COP curve for the chilled water cooling units from HP Lab utility data center. As the target temperature of the air the cooling unit pumps into the floor plenum increases, the COP increases.

Where Q is amount of power the servers and hardwares consume. COP(T) is our COP at $T = T_{sup} + T_{adj}$. Currently we assume a uniform T_{sup} from each cooling units due to the complications introduced by non-uniform cold air supply.

3.2   Energy Model

In this paper, our goal is to provide a general view and a fair comparison of the energy consume by a CDN and hybrid CDN-P2P architecture. To do so, we designed a series of model and perform an analysis. Our energy model is similar to the models used in [3]. Network model for energy is assumed to be flat network as shown in fig.2.

For each component, we consider two energy measurement: baseline energy consumption and energy consumed per work unit δ. The baseline energy is the energy consumed to keep the device on. Even when there is no traffic, the device still consume energy. The energy consumption for a single request in CDN server as:

$$E_d = E_s \tag{6}$$

while the energy consumption for a single request in network as:

$$E_r = d_s.E_r \tag{7}$$

where d_s is number of hops or the path length. We define δ_s and δ_r work-induced energy consumed per additional bit transferred by a server and router. We can express these per-bit work induced consumptions as follows:

$$\delta_s = \frac{(S_{max} - S_{base})}{M_s} \tag{8}$$

S_{base} is a server's baseline power consumption. S_{max} is a server's power when operating at maximum capacity. M_s is the maximum capacity in bit per second for a server. Same formulation also applied for network component.

$$\delta_r = \frac{(R_{max} - R_{base})}{M_R} \tag{9}$$

R_{max} is a router power when operating at maximum capacity. R_{base} is a router's baseline power consumption. M_R is maximum capacity in bit per second for a router.

Substituting eq.8, and eq.9 to eq.6 and eq.7, we can rewrite eq.6 and eq.7 as follows:

$$E_d = \delta_s.B + E_{s-baseline}$$
$$E_r = d_s.\delta_r.B + E_{r-baseline} \tag{10}$$

Finaly total energy is:

$$E_t = E_d + E_r$$
$$E_t = \delta_s.B + E_{s-baseline} + d_s.\delta_r.B + E_{r-baseline} \tag{11}$$

where B is size of file to be transferred.

Considering cooling energy, we can substitute eq.5 to eq.11, therefore total energy consumed:

$$\hat{E}_t = E_t.\left(1 + \frac{1}{COP(T)}\right) \tag{12}$$

4.   Result and Analysis

For numerical simulation, we use work induced energy consumed per second from [3]. For CDN server baseline, we use server idle power $E_{s-baseline}$ = 290 watt and router baseline power $E_{r-baseline}$ = 750 watt. We use that both values from [3]. The dummy network model used in this numerical simulation is shown in fig.5. The data flows from CDN to network in this case router then arrive in seeders. If seeders need to upload data to leechers then, it will flows through router then arrive in leechers.

Parameters values used in this numerical simulation

Table 1   Numerical Simulation Parameters.

| Symbol	Value
δ_s	$5.2.10^{-8}$ (J/b)
δ_r	$8.0.10^{-9}$ (J/b)
$E_{r-baseline}$	750 watt
$E_{s-baseline}$	290 watt
B	500MB
d_s	1
N_s^u	[0.25, 0.5, 0.75, 1] Mbps
N	[100, .., 1000]
T	[20, 25]
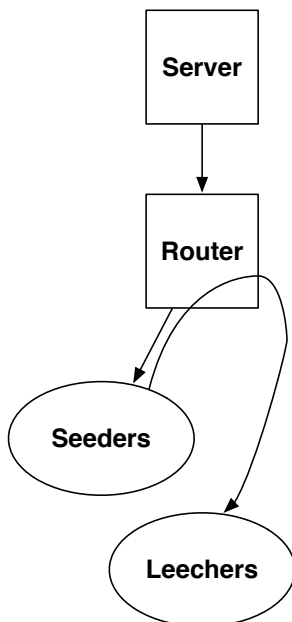
Fig. 5   Dummy network

Figure 6 shows energy usage for CDN server, router, total energy consumption for CDN scenario (without peer assisted). We plot energy consumption for CDN server, router, and total energy for every T or COP coefficient value. In CDN only scenario, the energy consumption is increase as number of client increase. This is natural. Increasing of clients imply to increasing of traffic from CDN server through router to clients.
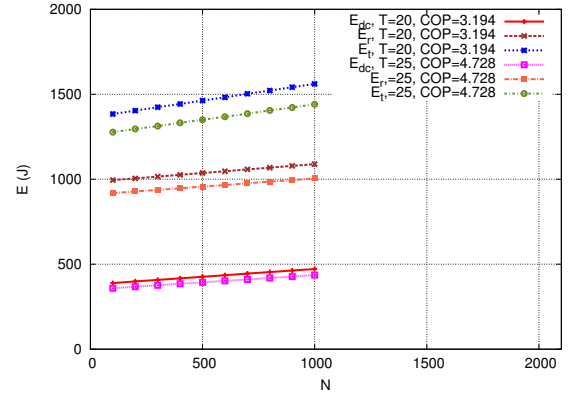
Fig. 6   0.
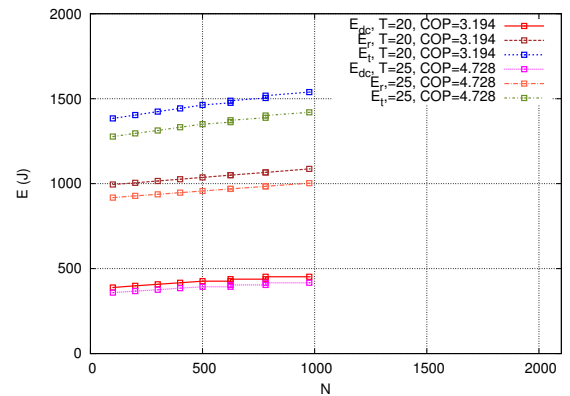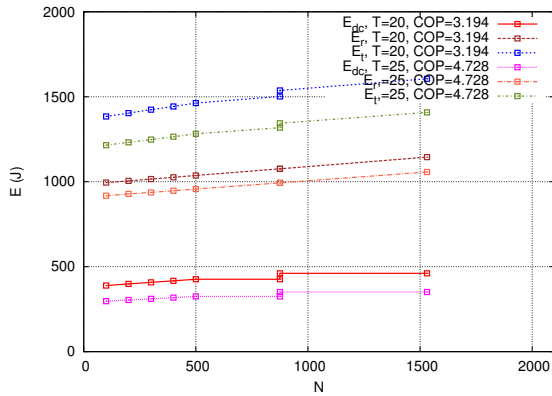

Fig. 7   1.
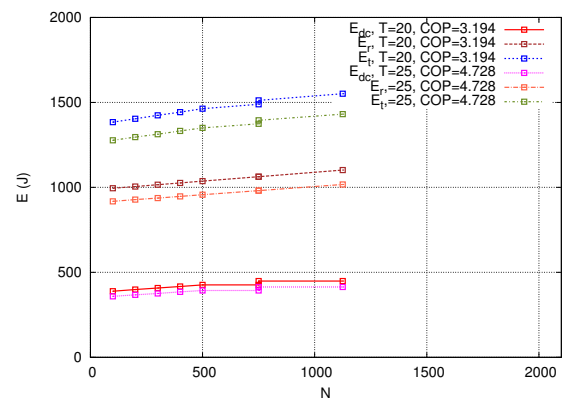
Fig. 8   2.
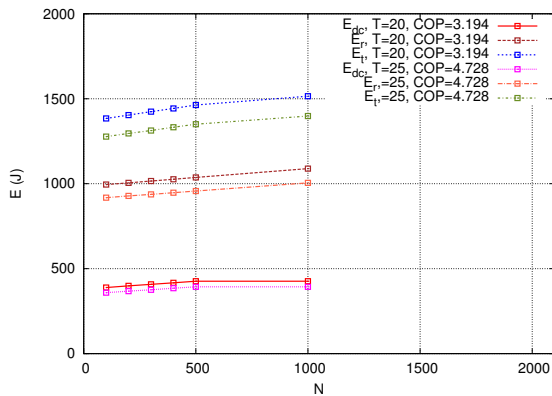
Fig. 9 3.

Fig. 10 3.

5.   Related Work

Content Distribution Networks with peer assist have been successfully deployed on the Internet, such as Akamai [4] and LiveSky [5]. The authors of [4] conclude from two real world traces that hybrid CDN-P2P can significantly reduce the cost of content distribution and can scale to cope with the exponential growth of Internet video content. Yin et al. [5] described LiveSkye as commercial operation of a peer-assisted CDN in China. LiveSky solved several challenges in the system design, such as dynamic resource scaling of P2P, low startup latency, ease of P2P integration with the existing CDN infrastructure, and network friendliness and upload fairness in the P2P operation. Measurement from LiveSky showed that LiveSky can save bandwidth around 40% [5]. The author in [6] and [7] proposed mechanisms to minimize CDN server bandwidth to make the content distribution cheap. They designed different peer prefetching policies of video on demand system in surplus mode while ensuring user quality of experience. A similar analysis has been done in [8] for live video streaming system where the authors proposed different limited peer contribution policies to reduce CDN bandwidth requirement and eventually off the distribution process from CDN to P2P system. An ISP friendly rate allocation schemes for a hybrid CDN-P2P video on demand system in [9]. These technique try to minimize CDN server bandwidth while reducing ISP unfriendly traffic and maximizing peer prefetching. Load on CDN server has been shown to be reduced using this approach while reducing cross ISP traffic. Above studies were performed for video on demand or live video streaming. While video is the most

popular content, the systems can be also for other type contents. Moreover while content based services are growing, energy consumption of a content distribution system has not been analyzed.

The idea that utilize ISP controlled home gateway to provide computing and storage services and adopts managed peer-to-peer model is presented in [10]. Valancius et al. [10] show that distributing computing platform in NaDa (Nano Data Center) save at least 20-30% energy compare to traditional data centers. The saving in NaDa comes from underutilizing home gateway, avoidance of cooling costs, and the reduction of network energy consumption as a result of demand and service co-localization.

The comparison between CDN architecture and peer-to-peer architecture are discussed in [11] and [12]. Both authors in [11] and [12] agree that CDN architecture is more energy saving compare to peer-to-peer architecture. Another interesting study of energy efficient in content delivery architectures is presented by Guan et al. [13]. by Guan et al. [13] comparing energy efficient of CDN architecture and CCN architecture. The authors in [13] conclude that CCN is more energy efficient in delivering popular content while the approach with optical bypass is more energy efficient in delivering infrequent accessed content.

To the best of our knowledge, the study of energy in that considering peer-to-peer in CDN architecture is presented in [14]. Mandal et al. [14] mentioned that hybrid CDN-P2P systems can reduce a significant amount energy in the core network around 20-40% less energy. The authors are focus on energy consumption of different algorithms for minimizing server bandwidth based on popularity of content. By maximizing peer request, the authors can minimizing CDN workload. The authors only considered energy consumption of hardware and does not include overhead inside data center. Our work is quite different, we take number of peers with static content and dynamic content (popularity) and add overhead of data center which power of cooling cause by temperature of hardware.

6.   Conclusion and Future Work

In this paper, we compare energy consumption between the peer-assisted CDN architecture and CDN architecture. Integrating peer to peer capability to assist the existing CDN has a great potential to save energy consumption. In this study, we show that event without explicitly considering energy consumption while assigning content, the peer assisted CDN saves significant energy. Although the energy savings depend on number of request (number of clients) and the content popularity, 20-30% energy can be saved for a moderately popular content. Additionally, by downloading the contents from the peers, especially of high popular content, the server load is decreased significantly. The energy saving potential of peer assisted CDN architecture presents as a strong reason for the ISP to reduce energy usage.

References

[1]   Institute, U.: the 2012 Uptime Institute Data Center Industry Survey. Available on http://uptimeinstitute.com/2012-survey-results.
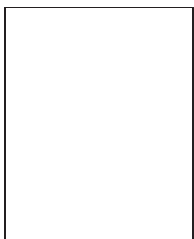[2]   Moore, J., Chase, J., Ranganathan, P. and Sharma, R.: Making

scheduling" cool": temperature-aware workload placement in data centers, *Proceedings of the annual conference on USENIX Annual Technical Conference*, pp. 5–5 (2005).

[3]   Nedevschi, S., Ratnasamy, S. and Padhye, J.: Hot data centers vs. cool peers, *Proceedings of the 2008 conference on Power aware computing and systems*, HotPower'08, Berkeley, CA, USA, USENIX Association, pp. 8–8 (online), available from ⟨http://dl.acm.org/citation.cfm?id=1855610.1855618⟩ (2008).

[4]   Huang, C., Wang, A., Li, J. and Ross, K. W.: Understanding hybrid CDN-P2P: why limelight needs its own Red Swoosh, *Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, NOSSDAV '08, New York, NY, USA, ACM, pp. 75–80 (online), DOI: 10.1145/1496046.1496064 (2008).

[5]   Yin, H., Liu, X., Zhan, T., Sekar, V., Qiu, F., Lin, C., Zhang, H. and Li, B.: LiveSky: Enhancing CDN with P2P, *ACM Trans. Multimedia Comput. Commun. Appl.*, Vol. 6, No. 3, pp. 16:1–16:19 (online), DOI: 10.1145/1823746.1823750 (2010).

[6]   Huang, C., Li, J. and Ross, K. W.: Can internet video-on-demand be profitable?, *SIGCOMM Comput. Commun. Rev.*, Vol. 37, No. 4, pp. 133–144 (online), DOI: 10.1145/1282427.1282396 (2007).

[7]   Huang, C., Li, J. and Ross, K.: Peer-assisted VoD: Making internet video distribution cheap, *Proc. of IPTPS*, Vol. 7 (2007).

[8]   Xu, D., Kulkarni, S. S., Rosenberg, C. and Chai, H.-K.: Analysis of a CDN–P2P hybrid architecture for cost-effective streaming media distribution, *Multimedia Systems*, Vol. 11, No. 4, pp. 383–399 (2006).

[9]   Wang, J., Huang, C. and Li, J.: On ISP-friendly rate allocation for peer-assisted VoD, *Proceedings of the 16th ACM international conference on Multimedia*, MM '08, New York, NY, USA, ACM, pp. 279–288 (online), DOI: 10.1145/1459359.1459397 (2008).

[10] Valancius, V., Laoutaris, N., Massoulié, L., Diot, C. and Rodriguez, P.: Greening the internet with nano data centers, *Proceedings of the 5th international conference on Emerging networking experiments and technologies*, ACM, pp. 37–48 (2009).

[11] Baliga, J., Hinton, K. and Tucker, R. S.: Energy consumption of the Internet, *Optical Internet, 2007 and the 2007 32nd Australian Conference on Optical Fibre Technology. COIN-ACOFT 2007. Joint International Conference on*, IEEE, pp. 1–3 (2007).

[12] Feldmann, A., Gladisch, A., Kind, M., Lange, C., Smaragdakis, G. and Westphal, F.-J.: Energy trade-offs among content delivery architectures, *Telecommunications Internet and Media Techno Economics (CTTE), 2010 9th Conference on*, IEEE, pp. 1–6 (2010).

[13] Guan, K., Atkinson, G., Kilper, D. and Gulsen, E.: On the Energy Efficiency of Content Delivery Architectures, *Communications Workshops (ICC), 2011 IEEE International Conference on*, pp. 1–6 (online), DOI: 10.1109/iccw.2011.5963557 (2011).

[14] Mandal, U., Lange, C., Gladisch, A. and Mukherjee, B.: Should ISPs adopt hybrid CDN-P2P in IP-over-WDM networks: An energy-efficiency perspective?, *Advanced Networks and Telecommuncations Systems (ANTS), 2012 IEEE International Conference on*, pp. 6–11 (online), DOI: 10.1109/ANTS.2012.6524219 (2012).

Mohamad Dikshie Fauzie was born in 1976. He received a bachelors degree and a master's degree from Institute of Technology Bandung, Indonesia. He is currently a Ph.D candidate at Keio University's Shonan Fujisawa Campus.

Achmad Husni Thamrin is Assistant Professor at Keio University. He is a graduate of Keio University, Graduate School of Media and Governance (Ph.D 2005, MMG, 2002). His research interests include multicast, Internet over broadcast media, and peer-to-peer networks.
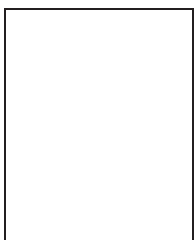
Jun Murai was born in March 1955 in Tokyo. Graduated Keio University in 1979, Department of Mathematics, Faculty of Science and Technology. He received his M.S. for Computer Science from Keio University in 1981, and received his Ph.D. in Computer Science, Keio University in 1987. He specializes in computer science, computer network, and computer communication. He is currently Dean of the Faculty of Environment and Information Studies, Keio University since October 2009. Former Vice-President of Keio University from May 2005 to May 2009. He was an Executive Director of the Keio Research Institute at SFC, Keio University from 1999 to 2005.