IEICE TRANS. COMMUN., VOL.E95–B, NO.8 AUGUST 2012

PAPER Special Section on Networking Technologies for Cloud Services

Power Consumption Evaluation of Distributed Computing Network Considering Traffic Locality

Yukio OGAWA[†a)], **Go HASEGAWA**[††], ***Members***, ***and*** **Masayuki MURATA**[††], ***Fellow***

SUMMARY When computing resources are consolidated in a few huge data centers, a massive amount of data is transferred to each data center over a wide area network (WAN). This results in increased power consumption in the WAN. A distributed computing network (DCN), such as a content delivery network, can reduce the traffic from/to the data center, thereby decreasing the power consumed in the WAN. In this paper, we focus on the energy-saving aspect of the DCN and evaluate its effectiveness, especially considering traffic locality, i.e., the amount of traffic related to the geographical vicinity. We first formulate the problem of optimizing the DCN power consumption and describe the DCN in detail. Then, numerical evaluations show that, when there is strong traffic locality and the router has ideal energy proportionality, the system's power consumption is reduced to about 50% of the power consumed in the case where a DCN is not used; moreover, this advantage becomes even larger (up to about 30%) when the data center is located farthest from the center of the network topology.

key words: *WAN, power consumption, distributed computing network, traffic locality, energy saving, energy proportionality*

1. Introduction

With the development of software as a service (SaaS) and cloud computing, computing resources, i.e., servers and storage systems, are increasingly being concentrated in a few huge data centers [1]. Moreover, an enormous amount of data will be created as a result of promoting the information society such as data related to digital video, surveillance cameras, and sensor-based applications [2], [3]. These changes will lead to a massive amount of data traversing a wide area network (WAN), which will increase the power consumed in the WAN. A distributed computing network (DCN) such as a content delivery network, which deploys a caching mechanism and related processing functions at network nodes [4], is a one of the major approaches to complement the concentrated processing in a data center [5]. It not only improves the response time to clients but also reduces the traffic from/to a data center (e.g., in cases of peer-to-peer (P2P) systems [6]–[10]), resulting in the power consumed in the WAN decreasing. This paper focuses on such an aspect of energy-saving through the use of the DCN [11].

When we investigate power efficiency in the WAN, it is important to consider traffic locality resulting from where traffic is supplied and demanded. For example, data collected from a local area for surveillance in a *smart city* [12] could be used mainly in that area and its surrounds. Information originating in and specific to the local area could be accessed more by users in the area and less by users far from it. In such a case, to avoid the round-trip route from the area to a distant data center, localizing the traffic in the area with the aid of a DCN is effective. Our goal is therefore to explore and describe the power efficiency profile of such a DCN, especially considering traffic locality. Note that content supply and demand locations are generally related to the contents themselves and independent of their server's location; therefore, the DCN's server function can be located anywhere to achieve the DCN's power efficiency without affecting the supply and demand distribution.

In recent years, power efficiency in telecommunications networks has been one of the major research issues in this field [13]–[15]. A typical WAN uses Internet protocol (IP) over a wavelength division multiplexing (WDM) optical network. IP routers (often called routers here) are the dominant power consumer in the networks; they consume 90% of the total power in such a network [16], [17]. There are several approaches for developing the network power consumption models and energy-saving techniques; these approaches are classified by their focus on the router whose power is consumed mainly by its chassis, switching fabric, line cards, and ports. The first approach treats a router as one power element. The number of active routers in the network is minimized by turning the router off or putting it to sleep, so that the total power is optimized while satisfying the traffic requirements [18], [19]. The second approach is based on router power consumption being the sum of its individual elements such as chassis and line card. The minimum number of elements are fed power in order to guarantees the traffic requirements [20]. The third approach focuses on the router's ports. Alternative routing for bypassing the active ports eliminates the power consumed by the ports themselves [16], [21], [22] or by their connecting links [23], [24]. The last approach assumes that the router exhibits energy proportionality [25], [26]. For this behavior, the router's power consumption is expressed as a linear function of the router's traffic load (i.e., energy per bit) [27], as a step-increasing function of the traffic load [28], and as the sum of the power consumption in active and idle modes [29].

Furthermore, with energy-proportional behavior assumed, the energy efficiency of distributed content delivery

Manuscript received November 30, 2011.
Manuscript revised March 16, 2012.

[†]The author is with the Telecommunications & Network Systems Division, Hitachi, Ltd., Tokyo, 103-0227 Japan.
[††]The authors are with the Graduate School of Information Science and Technology, Osaka University, Suita-shi, 565-0871 Japan.
a) E-mail: yukio.ogawa.xq@hitachi.com

DOI: 10.1587/transcom.E95.B.2538

architectures have been evaluated in the case of intermediate routers capable of storing popular contents for retrieval throughout the network [30], [31], in the case of distributed servers within an Internet service provider (ISP) compared with those located at a data center [32], and in the case of ISP-controlled home gateways distributing contents in a P2P fashion [33]. These studies aimed to achieve efficient content distribution, so they took into account the popularity of the content, i.e., which content is more likely to be downloaded. However, the content locality, i.e., where the content is created, stored, and downloaded was beyond their scope.

In this paper, we also assume that the router has energy proportionality; furthermore, we focus on evaluating the effect of traffic locality on the DCN's power efficiency as compared with the conventional computing scheme of processing only in the data center. For this purpose, we first formulate the problem of optimizing the DCN power consumption; this network can include a distributed server function providing data compression and caching to reduce the transit traffic through the WAN to a data center. Second, in order to evaluate the DCN power consumption, we describe the details of the network: topology, traffic supply and demand matrix, metric of data center location, and router power consumption model. The traffic matrix is introduced to quantify the traffic locality, i.e., the amount of traffic related to the geographical closeness between the demand and supply areas. Third, by applying a heuristic method, we examine how much the system power consumption is affected by traffic locality and the ratio of download traffic to upload traffic, as well as the data center location, and the router power consumption profile. Note that the system power consumption is assumed to depend on the traffic amount, so we will decrease the system power by reducing the traffic amount rather than by rerouting and grooming the traffic to increase the number of sleeping elements. Furthermore, although the DCN power consumption under a variety of traffic conditions is evaluated, the DCN deals only with a client-server flow model.

The rest of this paper is organized as follows. In Sect. 2, we present an application scenario and introduce a formulation of the power consumption problem. In Sect. 3, we describe the network model. Then, in Sect. 4, we evaluate the effect of traffic locality on the system power consumption. Finally, in Sect. 5, we give conclusions and discuss remaining issues.

2. Application and Power Consumption Models

We first give an application scenario suitable for the DCN to supplement the concentrated processing in a data center. We then construct a model of the system power consumption according to the scenario.

2.1 Application Scenario

A schematic diagram representing an application scenario is shown in Fig. 1. One of the typical applications is a surveillance application system that continuously collects data from widespread sensors such as measuring devices and still & video cameras. The application system comprises a distributed infrastructure including the sensors for collecting data, clients for using data processed on servers, network nodes (called nodes here), a server (with bundled storage) in a data center, and links between them. Each network node is an IP router capable of binding a server function, which we call a *node-attached virtual server* (*NAVS*). The NAVS can be implemented by using a general server device or a service module card added to the router (e.g., [34]). The NAVS might have the same processing functions as a server in the data center; moreover, it provides application-level service in order to reduce the volume of traffic traversing the WAN, as follows.

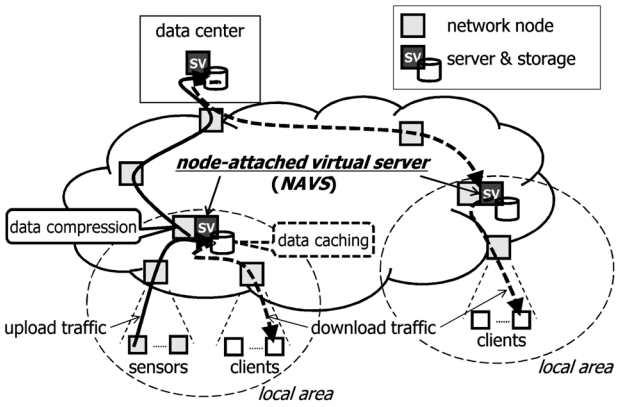


Fig. 1 Traffic flows of application scenario.

When a sensor collects raw data and uploads it to the data center, the sensor sends the data to the nearest NAVS. The NAVS then compresses the received upload data by filtering, re-segmenting, or encoding it and transfers it to the data center server. Furthermore, the NAVS caches the received data in a local storage device, after processing it if necessary, in preparation for future requests from clients. When a client downloads the data from the data center, the client requests the data from the nearest NAVS. If the requested data originated from the local area and the NAVS has cached it, the NAVS sends the cached data to the client. Otherwise, the NAVS downloads it from the data center server and forwards it to the client.

Here, a *local area* for each NAVS is defined as the area where the NAVS collects and delivers data. The total area of the sensors/clients is covered by all NAVSes. Thus, when another NAVS is added, each of the local areas becomes smaller than before.

2.2 System Power Consumption Model

Below, we formulate the problem of optimizing the DCN power consumption. We define the network topology $G = (N, E)$ as consisting of a set of network nodes N and a set of links E. Moreover, we utilize the following notation and definitions, where the traffic volume is measured in bits per

second:

Sets and parameters:

N : set of nodes in the network topology.
N_i^+ : set of neighboring nodes from which node $i \in N$ receives traffic.
N_i^- : set of neighboring nodes to which node $i \in N$ sends traffic.
f_{ij} : traffic volume flowing on the link from node $i \in N$ to node $j \in N$
g_i : traffic volume routed through node $i \in N$; this is equal to the traffic volume processed by the NAVS attached to node i.
t_{sd} : traffic volume flowing from source node $s \in N$ to destination node $d \in N$
f_{ij}^{sd} : traffic volume flowing from source node $s \in N$ to destination nodes $d \in N$ that is routed through the link from node $i \in N$ to node $j \in N$. ($f_{ij}^{sd} \in [0, t_{sd}]$).
$a_{ij}(f_{ij})$: power consumption of the link from node $i \in N$ to node $j \in N$; this is expressed as a function of f_{ij}.
$b_i(g_i)$: power consumption of node $i \in N$; this is expressed as a function of g_i.
$c_i(g_i)$: power consumption of the NAVS attached to node $i \in N$. It is also the power consumption of the data center server receiving traffic from node $i \in N$. This is also expressed as a function of g_i.

Variables:

P : number of NAVSes introduced in the network, each of which is attached to a corresponding node.
$x_{ij} \in \{0, 1\}$: binary variables that take the value of 1 if there is a link from node $i \in N$ to node $j \in N$ and 0 otherwise.
$y_i \in \{0, 1\}$: binary variables that take the value of 1 if a NAVS is attached to node $i \in N$ and 0 otherwise.

To make the model simple, we assume that there is one data center, which has one data center server connected to node $c \in N$. Given the above definition, our power-minimized design model is as follows.

Objective: minimize

$$E = \sum_{i \in N} \sum_{j \in N} a_{ij}(f_{ij}) x_{ij} + \sum_{i \in N} b_i(g_i) + \left(\sum_{i \in N} c_i(g_i) y_i + c_c(g_c) \right) \ (\forall c \in N) \tag{1}$$

Subject to:

$$\sum_{i \in N} y_i = P \tag{2}$$

$$\sum_{j \in N_i^+} f_{ji}^{sd} - \sum_{j \in N_i^-} f_{ij}^{sd} = \begin{cases} -t_{sd} & (\forall i \in N, \forall s \in N, \forall d \in N, i = s) \\ t_{sd} & (\forall i \in N, \forall s \in N, \forall d \in N, i = d) \\ 0 & (\text{otherwise}) \end{cases} \tag{3}$$

$$f_{ij} = \sum_{s \in N} \sum_{d \in N} f_{ij}^{sd} \ (\forall i \in N, \forall j \in N) \tag{4}$$

$$g_i = \begin{cases} \sum_{j \in N_i^+} \sum_{s \in N} \sum_{d \in N} f_{ji}^{sd} & (\forall i \in N, i \neq s) \\ \sum_{j \in N_i^-} \sum_{s \in N} \sum_{d \in N} f_{ij}^{sd} & (\forall i \in N, i = s) \end{cases} \tag{5}$$

Objective (1) states that the system power consumption (E) is modeled as the sum of the power consumptions of all links, of all nodes, and of all NAVSes as well as of the data center server. Constraint (2) states that the number of NAVSes introduced in the network is P. Constraint (3) ensures that the traffic flow is conserved at any node on the path from any source node to any destination node. Constraints (4) and (5) evaluate the total flow routed on each link and through any node, respectively.

The above formulation is a combinatorial optimization problem for finding the locations of P nodes, each of which binds a NAVS, minimizing E in Objective (1). This problem is analogous to the uncapacitated facility location problem [35] and is *NP*-hard to solve optimally.

2.3 Traffic Flow Formulation

In addition to the previous subsection, we also formulate t_{sd} when the traffic is flowing according to the application scenario described in Sect. 2.1. We define additional terms as follows.

N^e : set of nodes accessed from sensors and clients. ($N^e \subseteq N$)
N^v : set of nodes, each of which binds a NAVS. ($N^v \subseteq N$)
n_e : upload traffic volume transmitted from node $e \in N^e$ to NAVS-equipped node $v \in N^v$; this traffic is eventually transferred to the data center.
$m_{ee'}$: download traffic volume, which originated at node $e \in N^e$, received by node $e' \in N^e$.
γ : compression ratio of upload traffic volume at the NAVS attached to node $v \in N^v$. ($0 < \gamma \leq 1$)
$z_{sd} \in \{0, 1\}$: binary variables that take the value of 1 if source node $s \in N$ sends data to destination node $d \in N$ and 0 otherwise.

Additional constraints on Objective (1) are stated below.

Subject to:

$$z_{ev} \leq y_v \ (\forall e \in N^e, \forall v \in N^v) \tag{6}$$

$$\sum_{v \in N^v} z_{ev} = 1 \ (\forall e \in N^e) \tag{7}$$

$$z_{v'e'} \leq y_{v'} \ (\forall v' \in N^v, \forall e' \in N^e) \tag{8}$$

$$\sum_{v' \in N^v} z_{v'e'} = 1 \ (\forall e' \in N^e) \tag{9}$$

For upload traffic, constraint (6) states that a NAVS is attached to destination node $v \in N^v$ receiving traffic from node $e \in N^e$. Constraint (7) ensures that node $e \in N^e$, accessed from sensors, uploads traffic to one of the nodes, each of which binds a NAVS. Similarly, for download traffic, constraint (8) states that a NAVS is attached to source node $v' \in N^v$ sending traffic to node $e' \in N^e$. Constraint (9) ensures that node $e' \in N^e$, accessed from clients, downloads traffic from one of

the nodes, each of which binds a NAVS.

Then, the end-to-end traffic volume is formulated as follows.

(1) Upload traffic

The upload traffic volume originating from node $e \in N^e$ destined to node $v \in N^v$, which is a node with a NAVS, is

$$t_{ev} = n_e z_{ev} \ (\forall e \in N^e, \forall v \in N^v). \tag{10}$$

When this NAVS-equipped node $v \in N^v$ receives traffic, it uploads γ times the traffic to the node connected to the data center (node $c \in N$). The traffic volume from node $v \in N^v$ to node $c \in N$ is

$$t_{vc} = \gamma \sum_{e \in N^e} n_e z_{ev} \ (\forall v \in N^v, \forall c \in N). \tag{11}$$

(2) Download traffic

Node $e' \in N^e$ downloads all the traffic destined to itself from NAVS-equipped node $v' \in N^v$. The traffic volume from node $v' \in N^v$ to node $e' \in N^e$ is therefore

$$t_{v'e'} = \sum_{e \in N^e} m_{ee'} z_{v'e'} \ (\forall v' \in N^v, \forall e' \in N^e). \tag{12}$$

The traffic $t_{v'e'}$ includes the traffic cached in NAVS-equipped node $v \in N^v$ and that forwarded from the data center. The download traffic that node $e' \in N^e$ receives from the node connected to the data center (node $c \in N$) is all the traffic destined to node $e' \in N^e$ itself except for that stored locally in NAVS-equipped node $v' \in N^v$. The traffic cached in this NAVS-equipped node $v' \in N^v$ is the traffic from all the nodes $e \in N^e$ connected to node $v' \in N^v$ itself; this traffic is $\sum_{e \in N^e} m_{ee'} z_{ev'}$. Therefore, the total traffic volume from node $c \in N$ to node $v' \in N^v$ is

$$t_{cv'} = \sum_{e' \in N^e} \sum_{e \in N^e} m_{ee'} (1 - z_{ev'}) z_{v'e'} \quad (\forall c \in N, \forall v' \in N^v). \tag{13}$$

3. Network and System Model

In this section, in order to evaluate the power consumed by the DCN, we describe the network system in detail.

3.1 Network Topology

We consider the Japan-wide network system depicted in Fig. 2. This network has a hierarchical architecture composed of a core network and edge networks. The core network is a Japan-wide network composed of 47 core nodes, i.e., IP routers, located in the 47 prefectural capitals. The core nodes are connected by 75 links according to the topology of road and railroad networks. Meanwhile, an edge network is a prefecture-wide network where each edge node, which is also an IP router, aggregates sensors/clients and connects to a core node at a distance of 1 hop. In all the edge networks, there are 1194 edge nodes; these are located

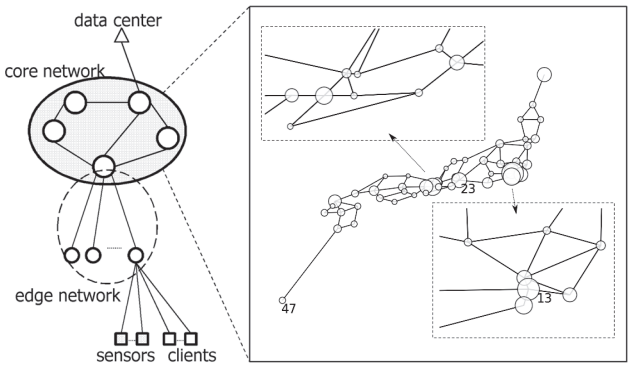


Fig. 2 Network topology (node size proportional to the number of households in the prefecture where it is located (right figure)).

in the areas covered by local governments, i.e., wards, cities, and districts as of 2008. Each of these core and edge nodes is capable of binding a NAVS. In addition, a server in a data center is connected directly to the core node where the data center is located. Note that access networks are not included in this topology, because their energy consumption is hardly affected when the traffic load changes [36].

3.2 Traffic Supply and Demand Matrix

In order to define the metric of traffic locality, we introduce a matrix giving the traffic volumes between source and destination nodes in the network.

(1) Upload traffic

To upload traffic to a data center via a NAVS, each node in an edge network generates a traffic load proportional to the number of households (as of 2008) in the area such as ward, city, and district where the edge node is located; this traffic load is set to n_e ($e \in N^e$), as mentioned in Sect. 2.3.

(2) Download traffic

To define the metric of traffic locality resulting from where traffic is supplied and demanded, we introduce a traffic supply and demand matrix based on a gravity model [37]. This matrix is expressed by $m_{ee'}$ ($e \in N^e, e' \in N^e$), as mentioned in Sect. 2.3, as follows:

$$m_{ee'} = k \frac{A_e A_{e'}}{d_{ee'}^{\beta}} \quad (\forall e \in N^e, \forall e' \in N^e), \tag{14}$$

where k is a proportionality constant, A_e and $A_{e'}$ are factors related to the traffic supply and demand amounts in the areas where the source node $e \in N^e$ and destination node $e' \in N^e$ are located, respectively, $d_{ee'}$ is the distance between these two nodes, β is a parameter for changing the relationship between the distance and traffic amounts exchanged between the two areas. Here, A_e and $A_{e'}$ are assigned to the number of households in the areas, respectively. In addition, distance $d_{ee'}$ is expressed as the hop count between the two nodes. If the source and destination nodes are the same node, the value of $d_{ee'}$ is set to 0.1 for convenience.

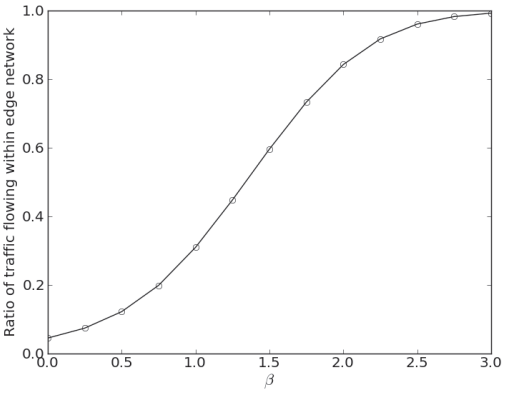


Fig. 3 Ratio of traffic flowing within edge network as a function of β in Eq. (14).

Furthermore, we denote the ratio of the total amounts of download traffic to upload traffic by α. The download traffic that originated at node $e \in N^e$ is as follows:

$$\sum_{e' \in N^e} m_{ee'} = \alpha n_e \quad (\forall e \in N^e). \tag{15}$$

The constant k in Eq. (14) is therefore calculated as follows:

$$k = \frac{\alpha n_e}{\sum_{e' \in N^e} \frac{A_e A_{e'}}{d_{ee'}^{\beta}}} \quad (\forall e \in N^e). \tag{16}$$

Here, β specifies the degree of traffic locality. An example of how β changes the traffic locality is shown in Fig. 3. This figure shows the ratio of traffic flowing within the edge network to the total traffic as a function of β. When β is smaller than 1.0, the traffic flowing within the edge network is less than 30%. On the other hand, when the β is larger than 2.0, this traffic flow is more than 80%.

3.3 Metric of Data Center Location

The location of the data center has a large effect on the power consumed for forwarding the traffic to the data center. We therefore define a metric by using the closeness centrality (C_c) [38] of the node connected to the data center (node $c \in N$) as follows.

$$C_c = \left(\frac{\sum_{i \in N^e} (n_i + m_i) d_{ic}}{\sum_{i \in N^e} (n_i + m_i)} \right)^{-1} \quad (\forall c \in N), \tag{17}$$

where n_i and m_i are traffic volumes uploaded from node $i \in N^e$ and downloaded to it, respectively, and d_{ic} is the distance (i.e., hop count) between node $i \in N^e$ and node $c \in N$. A node that has a higher C_c value receives data via shorter-distance transmission than other nodes. In our network model, the 23th node (which corresponds to Aichi) has the maximum C_c of 0.175 and the 47th node (which corresponds to Okinawa) has the minimum C_c of 0.084 (see Fig. 2).

3.4 Router Power Consumption Model

To describe the power characteristics of a router, we consider steady-state traffic and energy-proportional behavior. As shown in Fig. 4, the router's observed power consumption might be a step-increasing function of traffic load, in which the step size depends on the power consumption of the router's chassis, switching fabric, line cards, and active ports. This power profile (which is $b_i(g_i)$ ($i \in N$) in Sect. 2.2) is approximated as $l_i g_i + base_i$, where l_i is a proportionality factor and $base_i$ is the baseline, i.e., power consumption while the router is idle. The baseline is as much as about 80% of the maximum power consumption in the current router architecture [26]. However, we believe that the baseline will become smaller and that routers will come to have the ideal power-proportional characteristics modeled in studies like [25], [26] in the future, as power saving techniques, e.g., [29], [39]–[41], are developed and applied.

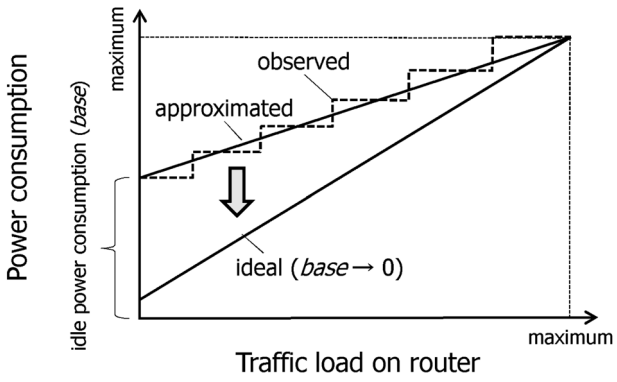


Fig. 4 Router's energy proportionality [26], [41].

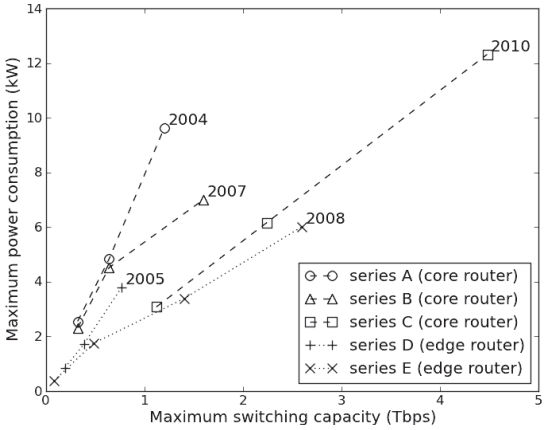


Fig. 5 Maximum power consumption versus maximum switching capacity for several series of routers [42]–[44] (Four-digit numbers are series release years).

Figure 5 plots the maximum switching capacity versus maximum power consumption for several series of core/edge router products; their specifications were gathered from the following web pages [42]–[44]. For routers within the same series, the ratio of router maximum power to maximum capacity is almost the same. We assume that the router's idle power is proportional to the maximum power consumption; this ratio is independent of $i \in N$ and denoted by R_{base}. As a result, the proportionality factor (l_i ($i \in N$))

is the same for all routers in a single series; moreover, the larger routers in the series consume more power in the idle state. When a router has a large R_{base} value, the router consumes much power even if less traffic load is offered. Meanwhile, as R_{base} decreases, the router comes to have perfect energy proportionality.

In addition, since we suppose that the NAVS is implemented by dedicated hardware [45], the power consumption of a NAVS could be close to that of a router. Hence, we assume that the NAVS's power profile ($c_i(g_i)$ ($i \in N$)) is same as the router's ($b_i(g_i)$ ($i \in N$)). Note that the power consumed by a node (i.e., a router) with a NAVS is handled separately as that consumed by a node and that consumed by a NAVS, according to Objective (1).

4. Evaluation

The DCN power consumption (E in Objective (1)) was evaluated by simulating the traffic load on each node of the network system described in Sect. 3. In this section, we first explain the evaluation method and then present detailed results.

4.1 Evaluation Method

To solve the combinatorial optimization problem described in Sect. 2.2 for such a large network, one needs to use an efficient heuristic. We chose to apply the greedy-interchange heuristic [35] to select the locations of P NAVSes minimizing E. Its algorithm is explained as follows.

step 1: The first node to be given a NAVS is found as follows. Starting from the state where all of the nodes do not yet have a NAVS attached, E is calculated when each node has one. For a given node, if NAVS attachment minimizes E, that node is selected as the first node to have the NAVS attached.

step 2: The next node to be given a NAVS is chosen as follows. E is calculated when a NAVS is added to each node that does not have one yet. For a given node, if NAVS attachment minimizes E, that node is selected next for NAVS attachment.

step 3: Relocation of the NAVS to an adjacent node is tested. E is calculated when the NAVS is moved to each adjacent node that does not yet have one. If E is smaller than before the NAVS was moved, the NAVS is moved to the adjacent node. This step is repeated until E becomes larger after NAVS movement, in which case the NAVS is not moved.

step 4: The number of NAVSes is counted. If the number is less than P, go to step 2.

Note that we neglect the first term of Objective (1) when calculating E because the WDM links of an IP-over-WDM network account for only around 10% or less of the total power consumption [16], [17]. In addition, although the traffic route may be changed because of constraints on the available link bandwidth, this consideration is outside the scope of this paper.

The following conditional settings were used for the evaluations conducted in the next subsection.

- Compression ratio of upload traffic volume at a NAVS (γ in Eq. (11)) was set to a constant value of 0.1.
- The sum of total upload traffic ($\sum_{e \in N^e} n_e$) and total download traffic ($\sum_{e \in N^e} \sum_{e' \in N^e} m_{ee'}$) was set to the estimated volume of Internet traffic in Japan (2.6 Tbps (as of 2011)) [46].
- Routers of the latest series specified in Fig. 5 were selected for nodes in the core and edge networks. Hence, "series C" specifications were used for core nodes and "series E" ones were used for edge nodes. Furthermore, the smallest router model (or the combination of smaller models) was chosen from the series; it satisfied the condition that the maximum possible traffic load on the router was less than 50% of the router's maximum switching capacity regardless of the traffic route selected in the network.
- To avoid an impractical route from an edge node to the data center, the data center was conveniently supposed to provide the services of a NAVS. Hence, in step 1, the location of the first NAVS was fixed to that of the data center and this NAVS was connected to the same node to which the data center server was also connected. Then, if the distance from an edge node to the data center was shorter than that from the edge node to the nearest NAVS outside the data center, the edge node could directly communicate with the data center.

4.2 Evaluation Results

In Sect. 4.2.1 to Sect. 4.2.4, we present evaluation results for when NAVSes are distributed to only core nodes. In these sections, the cycle from step 1 to step 4 explained in the previous subsection is repeated 46 times (i.e., in cases of $P = 2$ to $P = 47$) for one set of evaluations.

4.2.1 Overview of Power Optimization Results

The change in total power consumption (E) versus the number of NAVSes (P) is shown in Fig. 6. Here, the traffic matrix parameters, α and β in Eqs. (14)–(16), were set to 1.0 and 2.0, respectively. The data center was connected to the 13th core node (see Fig. 2); this node corresponds to Tokyo, which sent and received the largest volume in our network system. Each router's power consumption ratio in the idle state (R_{base}) was set to 0.01. In addition, "$P = 1$" means that there was only the data center (containing the first NAVS) and no distributed NAVSes outside the data center. E with only the data center is denoted by E_1 here.

Overall, as more NAVSes were added, E decreased through a reduction in the amount of traffic; it was minimum (around 70% of E_1) when P was equal to 15. Then, it increased steadily with additional base power consumption of the NAVSes. We can say that the number and locations of

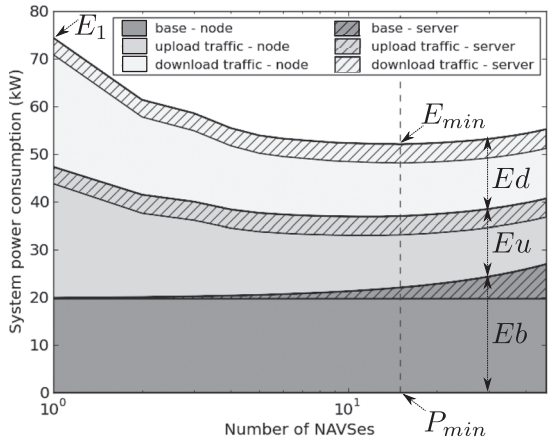


Fig. 6 System power consumption (E) versus number of NAVSes (P).

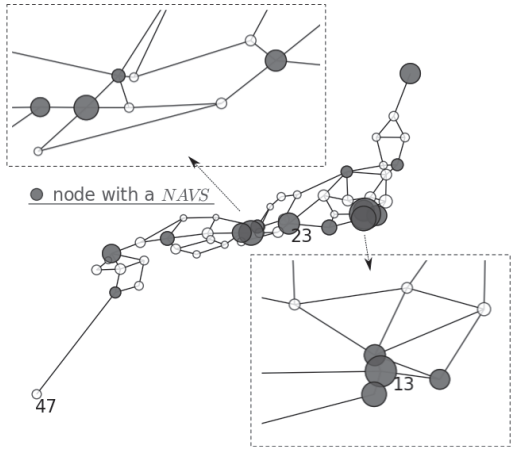


Fig. 7 Location of P_{min} NAVSes at optimized system power consumption (E_{min}) shown in Fig. 6 (Network topology is the same as in Fig. 2).

NAVSes were optimized when E was minimum. Below, the optimized E is denoted by E_{min} and the number of NAVSes providing E_{min} is denoted by P_{min}. Figure 7 gives the location of P_{min} NAVSes at E_{min} shown in Fig. 6. In the optimal state, the NAVSes were attached to the core nodes located in major metropolitan areas where most of the traffic (e.g., about 70% of both all upload and download traffic in this evaluation) was generated and consumed.

In addition, E was composed of three components: power consumed independently of traffic amount (denoted by the term "base") (Eb), that consumed for the upload traffic (Eu), and that consumed for the download traffic (Ed). Each component was the sum of the power consumed by all nodes (excluding that consumed by NAVSes) and that consumed by all servers (i.e., P NAVSes and the data center server), which were related to the second and third terms of Objective (1), respectively.

Most of the power consumed by all of the servers was that consumed by NAVSes. The power consumed by all NAVSes in Eu as well as that in Ed was constant and independent of P because the total amount of upload traffic processed (and also cached) by all NAVSes was constant, and the total number of requests for downloading from all NAVSes remained unchanged regardless of P.

Meanwhile, as P was increased, the power consumed by all nodes in Eu decreased because the average hop count between an edge node and a corresponding NAVS became smaller. At this time, the power consumed by all nodes in Ed was affected by the trade-off between the hop count from an edge node to a NAVS and the amount of traffic cached at the NAVS. This was because, when a NAVS was added, each of the NAVSes covered fewer edge nodes, resulting in less traffic being cached per NAVS and more traffic being downloaded from the data center server. The amount of cached traffic per NAVS, and hence this trade-off as well, was dependent on the degree of traffic locality (β), which will be explained in the next subsection.

4.2.2 Effect of Traffic Locality

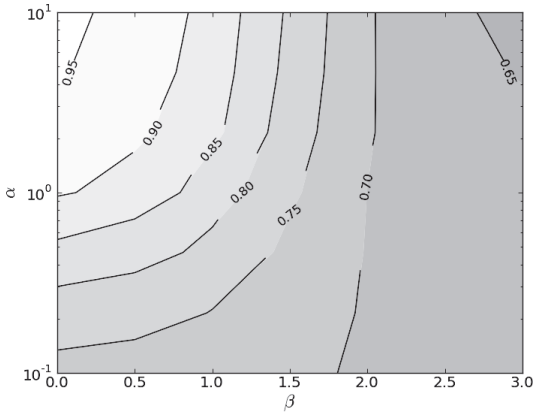
When both the ratio of download traffic to upload traffic and the ratio of traffic flowing within a local area to the total traffic become large, caching at a NAVS becomes more effective and the power consumed by the system thereby becomes smaller. We have therefore analyzed the optimized system power consumption (E_{min}) as a function of traffic matrix parameters (α and β), as shown in Fig. 8. This figure indicates (a) the ratio of E_{min} to E_1 and (b) the number of NAVSes at E_{min} (P_{min}). Note that E_1 was hardly different for the values of α and β. As in the previous subsection, the data center was at the 13th core node and R_{base} was set to 0.01.

When α was set to nearly 0.1, most of the traffic was upload traffic from edge nodes to the data center server via NAVSes. Then, E was almost the sum of Eb and Eu. In this case, lower E_{min} (69 to 73% of E_1) was achieved by deploying data compression in the NAVSes. Thus, E_{min} had little relation to β. In addition, when α was 0.1, P_{min} was approximately constant (14 to 16).

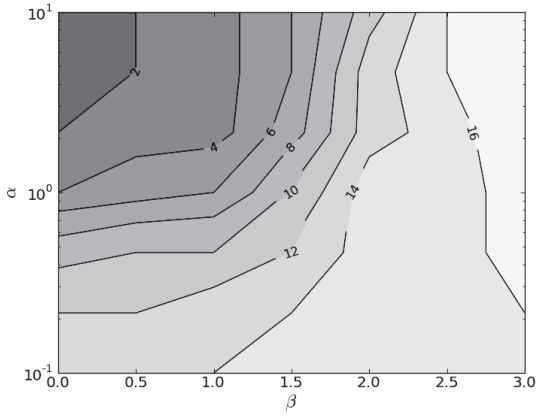
When α became large and approached 10, the contribution of Ed to E also became large and dominant. If β equaled 0.0 in this case, the closeness between the node where the download traffic originated and the node downloading the traffic had no influence on the amount of traffic exchanged between the two nodes. There was less downloaded traffic that originated at edge nodes in the same local area, so caching by the NAVSes made little contribution to the reduction in power consumption. As a result, E_{min} was 96% of E_1 at the point (α,β)=(10,0.0). At this time, P_{min} was two. An extra NAVS resulted in a higher E. On the other hand, when β was 3.0, nearly 100% of the download traffic originated from the local area. This led to the power reduction due to caching being very effective. Consequently, E_{min} was 64% of E_1 at the point (α,β)=(10,3.0). In this case, P_{min} was maximum, for which there were 17 NAVSes.

4.2.3 Effect of Data Center Location

In order to evaluate the effect of data center location, we analyzed the change in E_{min} with the data center's closeness



(a) Ratio of optimized system power consumption (E_{min}) to system power consumption with only data center (E_1)



(b) Number of NAVSes at optimized system power consumption (P_{min})

Fig. 8 Optimized system power consumption (E_{min}) and number of NAVes at E_{min} (P_{min}) as a function of traffic matrix parameters (α and β).

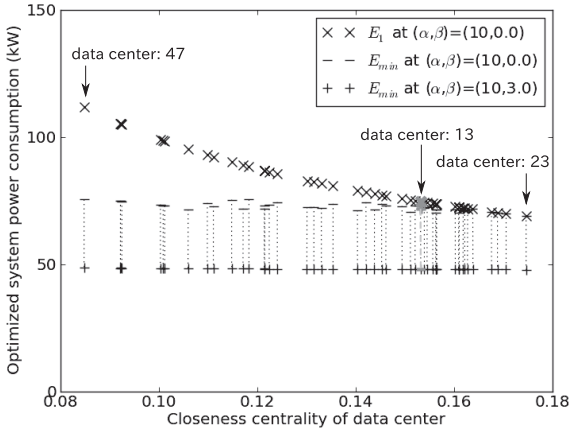


Fig. 9 Change in optimized system power consumption (E_{min}) with data center's closeness centrality (C_c).

centrality (C_c in Eq. (17)), as shown in Fig. 9. Here, R_{base} was set to 0.01, as in the previous subsections. As shown in Fig. 8(a), E_{min} was maximum at the point (α,β)=(10,0.0) and minimum at the point (α,β)=(10,3.0). We have therefore

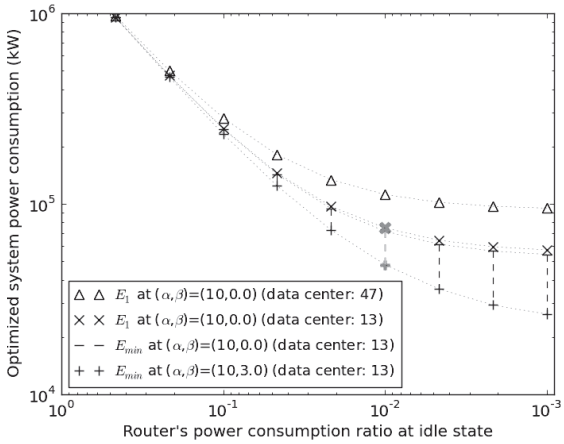


Fig. 10 Optimized system power consumption (E_{min}) versus idle power consumption ratio (R_{base}).

drawn a line connecting these two points for each result in Fig. 9. Furthermore, we have added the point indicating E_1; this was indicated only for (α,β)=(10,0.0), because E_1 was hardly different for the values of α and β.

With only the data center, E_1 largely depended on C_c of the data center. In our network model, C_c was minimum when the data center was connected to the 47th node. Since this 47th node received traffic over a longer distance than others, E_1 in this case was maximum and 1.5 times larger than in the case where the data center was located at the 13th node.

On the other hand, when the NAVSes were distributed, the dependency on the data center location disappeared. When there was no traffic locality (i.e., (α,β)=(10,0.0)), E_{min} was almost the same as E_1 in the case where the data center's C_c was maximum, regardless of the data center location. In this case, P_{min} was two; this meant that a NAVS other than the one in the data center behaved as an alternative to the data center server located far from the center of the network topology. Furthermore, when there was strong traffic locality (i.e., (α,β)=(10,3.0)), E_{min} retained the minimum value regardless of where the data center was located. In this case, P_{min} was nearly constant at 16 or 17.

4.2.4 Effect of Energy Proportionality

Finally, we evaluated the effect of modeling a router's idle power consumption. Figure 10 shows how E_{min} as well as E_1 changed with the ratio of idle power consumption (R_{base}). In this figure, we show two cases for E_1: the data center's C_c was minimum (i.e., connected to the 47th core node) and approximately maximum (i.e., connected to the 13th core node). Moreover, we indicate two cases for E_{min}: in one, E_{min} is maximum (i.e., (α,β)=(10,0.0)) and in the other, it is minimum (i.e., (α,β)=(10,3.0)). We omitted other results because E_1 had little relevance to the values of α and β; furthermore, E_{min} was little dependent on the location of the data center, as mentioned in the previous subsections.

When each node had imperfect energy proportionality,

i.e., a large R_{base} value, each node consumed a lot of base power even if there was less traffic load on it. This made the system power consumption very high. In this case, Eb accounted for a large portion of E. Since E was little dependent on both Eu and Ed, both E_1 and E_{min} were little influenced by the location of the data center or the values of α and β.

On the other hand, as each node had perfect energy proportionality, i.e., a small R_{base} value, the system power consumption became small. This was because each node consumed less base power and the traffic load was not so high for its capacity. In this case, the contribution of both Eu and Ed to E became relatively large. As a result, the effect of the DCN (i.e., the difference between E_1 and E_{min}) and the influence of the data center location, α, and β also became large. When there was strong traffic locality (i.e., (α,β)=(10,3.0)), E_{min} approached 46% of E_1 in the case where the data center was located at the 13th node; furthermore, E_{min} was 28% of E_1 in the case where the data center was located at the 47th node.

4.2.5 Results for Another Application Scenario

The DCN could have another application scenario in which a NAVS can communicate with other NAVSes as well. Consider the situation in which a client requests data from a NAVS (called NAVS A) and NAVS A does not have the data; this data is cached in another NAVS (called NAVS B) which lies closer to NAVS A than the data center does. Then, NAVS A downloads it not from the data center server but from NAVS B. We evaluated this case with the same settings as in Sect. 4.2.2. The difference from Fig. 8(a) was maximum at the point (α,β)=(1.0,0.5), although E_{min} at this point was reduced to 97% of that in the case of Fig. 8(a) for the given network topology.

We also investigated the situation in which a NAVS could be attached to a node in the edge network. In this case, with the same settings as in Sect. 4.2.2, the evaluation results showed that, when there was strong traffic locality (i.e., (α,β)=(10,3.0)), E_{min} in this case approached 93% of that in the case of Fig. 8(a). The power reduction was not so large because all edge nodes could reach the core network in a single hop for the given topology.

5. Conclusion

This paper focused on the energy-saving aspect of the DCN, especially considering traffic locality. Through numerical evaluations of a Japan-wide network model, we revealed the following results. When most of the traffic was upload traffic to the data center, the power consumption of the DCN had little relation to traffic locality. On the other hand, as the download traffic became dominant, the dependence on traffic locality also became large. When there was strong traffic locality, the power consumed with the DCN was reduced by up to about 30% of that consumed with only the data center. We note that, although the NAVS's power model and specifications affect the total power consumption (E), the above results should hold as long as the power consumption of routers predominates in E. In this paper, the DCN's topology was fixed. The energy efficiency related to traffic locality described above could hold for similar static network topologies. However, a dynamic topology change, e.g., dynamic WDM path setting bypassing an intermediate IP router in an IP-over-WDM network, will have an impact on the DCN's power consumption. In the future, we would therefore like to investigate a power-efficient network topology.

Acknowledgment

This work was supported in part by "Research and Development on Secure Cloud Networking Technologies (Intelligent Distributed Processing Technologies)" of Ministry of Internal Affairs and Communications, Japan.

References

[1] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A berkeley view of cloud computing," Technical Report of EECS Department, University of California, Berkeley, no.UCB/EECS-2009-28, Feb. 2009.

[2] J. Gantz, "The diverse and exploding digital universe: An updated forecast of worldwide information growth through 2011," White Paper of International Data Corporation, March 2008.

[3] Cisco Systems, "Visual networking index." http://www.cisco.com/en/US/netsol/ns827/networking_solutions_sub_solution.html, June 2011.

[4] A.M.K. Pathan and R. Buyya, "A taxonomy and survey of content delivery networks," Technical Report, GRIDS-TR-2007-4, Grid Computing and Distributed Systems Laboratory, The University of Melbourne, Australia, Feb. 2007.

[5] A. Greenberg, J. Hamilton, D.A. Maltz, and P. Patel, "The cost of a cloud: Research problems in data center networks," SIGCOMM Computer Communication Review, vol.39, pp.68–73, Dec. 2008.

[6] H. Wang, J. Liu, B. Chen, K. Xu, and Z. Ma, "On tracker selection for peer-to-peer traffic locality," Proc. IEEE Tenth International Conference on Peer-to-Peer Computing (P2P 2010), pp.1–10, Aug. 2010.

[7] Y. Liu, L. Guo, F. Li, and S. Chen, "A case study of traffic locality in Internet P2P live streaming systems," Proc. 29th IEEE International Conference on Distributed Computing Systems (ICDCS 2009), pp.423–432, June 2009.

[8] C. Tian, X. Liu, H. Jiang, W. Liu, and Y. Wang, "Improving bittorrent traffic performance by exploiting geographic locality," Proc. IEEE Global Telecommunications Conference (GLOBECOM 2008), pp.1–5, Dec. 2008.

[9] R. Bindal, P. Cao, W. Chan, J. Medved, G. Suwala, T. Bates, and A. Zhang, "Improving traffic locality in bittorrent via biased neighbor selection," Proc. 26th IEEE International Conference on Distributed Computing Systems (ICDCS 2006), July 2006.

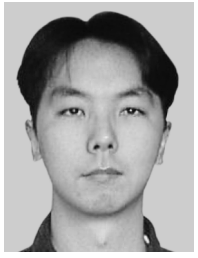
[10] T. Karagiannis, P. Rodriguez, and K. Papagiannaki, "Should Internet service providers fear peer-assisted content distribution?," Proc. 5th ACM SIGCOMM Conference on Internet Measurement (IMC 2005), Oct. 2005.

[11] Y. Ogawa, G. Hasegawa, M. Murata, and S. Nishimura, "Performance evaluation of distributed computing environment considering power consumption," IEICE Technical Report, IN2009-172, March 2010. (in Japanese).

[12] J.M. Hernández-Muñoz, J.B. Vercher, L. Muñoz, J.A. Galache,

M. Presser, L.A.H. Gómez, and J. Pettersson, The future Internet, ch. Smart cities at the forefront of the future Internet, pp.447–462, Springer-Verlag, Berlin, 2011.
[13] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti, "Energy efficiency in the future Internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures," IEEE Communications Surveys Tutorials, vol.13, no.2, pp.223–244, May 2011.
[14] Y. Zhang, P. Chowdhury, M. Tornatore, and B. Mukherjee, "Energy efficiency in telecom optical networks," IEEE Communications Surveys Tutorials, vol.12, no.4, pp.441–458, Nov. 2010.
[15] H. Mellah and B. Sanso, "Review of facts, data and proposals for a greener Internet," Proc. 6th International Conference on Broadband Communications, Networks, and Systems (BROADNETS 2009), pp.1–5, Sept. 2009.
[16] G. Shen and R. Tucker, "Energy-minimized design for IP over WDM networks," IEEE/OSA J. Optical Communications and Networking, vol.1, no.1, pp.176–186, June 2009.
[17] W. Vereecken, W. Van Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester, "Power consumption in telecommunication networks: Overview and reduction strategies," IEEE Commun. Mag., vol.49, no.6, pp.62–69, June 2011.
[18] L. Chiaraviglio, M. Mellia, and F. Neri, "Reducing power consumption in backbone networks," Proc. IEEE International Conference on Communications (ICC 2009), pp.1–6, June 2009.
[19] B. Sanso and H. Mellah, "On reliability, performance and Internet power consumption," Proc. 7th International Workshop on Design of Reliable Communication Networks (DRCN 2009), pp.259–264, Oct. 2009.
[20] J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsiang, and S. Wright, "Power awareness in network design and routing," Proc. IEEE International Conference on Computer Communications (INFOCOM 2008), pp.457–465, April 2008.
[21] M. Gupta and S. Singh, "Greening of the Internet," Proc. 2003 ACM conference on Applications, technologies, architectures, and Protocols for Computer Communications (SIGCOMM 2003), pp.19–26, Aug. 2003.
[22] W. Van Heddeghem, M. De Groote, W. Vereecken, D. Colle, M. Pickavet, and P. Demeester, "Energy-efficiency in telecommunications networks: Link-by-link versus end-to-end grooming," Proc. 14th Conference on Optical Network Design and Modeling (ONDM 2010), pp.1–6, Feb. 2010.
[23] H. Imaizumi, T. Nagata, G. Kunito, K. Yamazaki, and H. Morikawa, "Power saving technique based on simple moving average for multichannel ethernet," Proc. 14th OptoElectronics and Communications Conference (OECC 2009), pp.1–2, July 2009.
[24] W. Fisher, M. Suchara, and J. Rexford, "Greening backbone networks: Reducing energy consumption by shutting off cables in bundled links," Proc. 1st ACM SIGCOMM workshop on Green networking (Green Networking 2010), pp.29–34, Aug. 2010.
[25] L.A. Barroso and U. Hölzle, "The case for energy-proportional computing," Computer, vol.40, no.12, pp.33–37, Dec. 2007.
[26] P. Mahadevan, P. Sharma, S. Banerjee, and P. Ranganathan, "A power benchmarking framework for network devices," Proc. 8th International IFIP-TC 6 Networking Conference (NETWORKING 2009), pp.795–808, May 2009.
[27] J. Baliga, R. Ayre, K. Hinton, and R. Tucker, "Green cloud computing: Balancing energy in processing, storage, and transport," Proc. IEEE, vol.99, no.1, pp.149–167, Jan. 2011.
[28] F.M. Ramos, R.J. Gibbens, F. Song, P. Rodriguez, J. Crowcroft, and I.H. White, "Reducing energy consumption in IPTV networks by selective pre-joining of channels," Proc. 1st ACM SIGCOMM workshop on Green networking (Green Networking 2010), pp.47–52, Aug. 2010.
[29] S. Nedevschi, L. Popa, G. Iannaccone, S. Ratnasamy, and D. Wetherall, "Reducing network energy consumption via sleeping and rate-adaptation," Proc. 5th USENIX Symposium on Networked Systems Design and Implementation (NSDI 2008), pp.323–336, April 2008.
[30] K.C. Guan, G. Atkinson, D. Kilper, and E. Gulsen, "On the energy efficiency of content delivery architectures," Proc. 4th International Workshop on Green Communications (GreenComm 2004), GreenComm4, June 2011.
[31] U. Lee, I. Rimac, and V. Hilt, "Greening the Internet with content-centric networking," Proc. 1st International Conference on Energy-Efficient Computing and Networking (e-Energy 2010), pp.179–182, April 2010.
[32] A. Feldmann, A. Gladisch, M. Kind, C. Lange, G. Smaragdakis, and F.J. Westphal, "Energy trade-offs among content delivery architectures," Proc. 9th Conference on Telecommunications Internet and Media Techno Economics (CTTE 2010), pp.1–6, June 2010.
[33] V. Valancius, N. Laoutaris, L. Massoulié, C. Diot, and P. Rodriguez, "Greening the Internet with nano data centers," Proc. 5th international conference on Emerging networking experiments and technologies (CoNEXT 2009), pp.37–48, Dec. 2009.
[34] C. Francalanci, P. Giacomazzi, and A. Poli, "Cost-performance optimization of application- and context-aware distributed infrastructures," IEEE Trans. Syst. Man. Cybern. A, Syst. Humans, vol.39, no.6, pp.1200–1213, Nov. 2009.
[35] K. Hidaka and H. Okano, "Simulation-based approach to the warehouse location problem for a large-scale real instance," Proc. 29th Conference on Winter Simulation (WSC'97), pp.1214–1221, Dec. 1997.
[36] J. Baliga, R. Ayre, W. Sorin, K. Hinton, and R. Tucker, "Energy consumption in access networks," Proc. Conference on Optical Fiber Communication/National Fiber Optic Engineers Conference (OFC/NFOEC 2008), pp.1–3, Feb. 2008.
[37] Y. Zhang, M. Roughan, N. Duffield, and A. Greenberg, "Fast accurate computation of large-scale ip traffic matrices from link loads," Proc. 2003 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS 2003), pp.206–217, June 2003.
[38] V. Latora and M. Marchiori, "A measure of centrality based on the network efficiency," eprint arXiv:cond-mat/0402050, Feb. 2004.
[39] G. Ananthanarayanan and R.H. Katz, "Greening the switch," Proc. 2008 conference on Power aware computing and systems (HotPower 2008), pp.7–7, Dec. 2008.
[40] IEEE, "IEEE P802.3az energy efficient ethernet task force." http://www.ieee802.org/3/az/index.html, June 2011.
[41] M. Yamada, T. Yazaki, N. Matsuyama, and T. Hayashi, "Power efficient approach and performance control for routers," Proc. IEEE International Conference on Communications (ICC 2009), pp.1–5, June 2009.
[42] "ALAXALA Networks Corporation." http://www.alaxala.com/, June 2011.
[43] "Cisco Systems, Inc." http://www.cisco.com/, June 2011.
[44] "Juniper Networks, Inc." http://www.juniper.net/, June 2011.
[45] M. Hiraiwa, H. Masukawa, and S. Nishimura, "Hitachi's involvement in networking for cloud computing," Hitachi Review, vol.59, no.5, pp.206–212, Dec. 2010.
[46] Ministry of Internal Affairs and Communications, "Estimate of Internet traffic in Japan." http://www.soumu.go.jp/main_content/000109282.pdf, March 2011. (in Japanese).

Yukio Ogawa received his M.S. degree in Science from Nagoya University, Japan, in 1994, and Ph.D. degree in Information Science and Technology from Osaka University, Japan, in 2012. He joined the Hitachi Central Research Laboratory in Japan, in 1994. He is currently a senior engineer in the Telecommunications & Network Systems Division, Hitachi, Ltd. in Japan. His research interests include network architectures for enterprise and carrier systems. He is a member of IEEE.



Go Hasegawa received the M.E. and D.E. degrees in Information and Computer Sciences from Osaka University, Osaka, Japan, in 1997 and 2000, respectively. From July 1997 to June 2000, he was a Research Assistant in the Graduate School of Economics, Osaka University. He is now an Associate Professor at the Cybermedia Center, Osaka University. His research is in the area of transport architecture for future high-speed networks and overlay networks. He is a member of IEEE.

Masayuki Murata received the M.E. and D.E. degrees in Information and Computer Science from Osaka University, Japan, in 1984 and 1988, respectively. In April 1984, he joined the Tokyo Research Laboratory, IBM Japan, as a Researcher. From September 1987 to January 1989, he was an Assistant Professor with the Computation Center, Osaka University. In February 1989, he moved to the Department of Information and Computer Sciences, Faculty of Engineering Science, Osaka University. In April 1999, he became a Professor at the Cybermedia Center, Osaka University, and since April 2004, he has been with the Graduate School of Information Science and Technology, Osaka University. He has had more than five hundred papers published in international and domestic journals and conferences. His research interests include computer communication network architecture, performance modeling, and evaluation. He is a member of IEEE and ACM. He became chair of the IEEE COMSOC Japan Chapter in 2009. He is also working at the National Institute of Information and Communications Technology (NICT) as Deputy of the New-Generation Network R&D Strategic Headquarters.