

Calculating the minimum bounds of energy consumption for cloud networks

Jens Buysse*, Konstantinos Georgakilas †,
Anna Tzanakaki†, Marc De Leenheer*, Bart Dhoedt *, Chris Develder* and Piet Demeester*

* Ghent University – IBBT, INTEC, Ghent, Belgium

Email: {jens.buysse,chris.develder}@intec.ugent.be

† AIT, Athens, Greece

Email: {koge,atza}@ait.edu.gr

Abstract—This paper is aiming at facilitating the energy-efficient operation of an integrated optical network and IT infrastructure. In this context we propose an energy-efficient routing algorithm for provisioning of IT services that originate from specific source sites and which need to be executed by suitable IT resources (e.g. data centers). The routing approach followed is anycast, since the requirement for the IT services is the delivery of results, while the exact location of the execution of the job can be freely chosen. In this scenario, energy efficiency is achieved by identifying the least energy consuming IT and network resources required to support the services, enabling the switching off of any unused network and IT resources. Our results show significant energy savings that can reach up to 55% compared to energy-unaware schemes, depending on the granularity with which a data center is able to switch on/off servers.

I. INTRODUCTION

It is widely accepted that Internet traffic is growing very fast [1]. As such, the energy consumption of the ICT becomes significant and cannot be neglected any more. Several studies have pointed out that ICT around the world is responsible for up to 10% of the total energy consumption and 2% of global carbon emissions [2].

In the context of Future Internet and cloud computing, integration of IT and network resources in a common infrastructure that supports a large variety of existing and future services also becomes a necessity. Cloud computing entails a system to access a set of computing resources such as computational, data and software services in an on-demand and convenient way, without the end-user interacting with the hardware or service provider [3]. Consequently the network supporting the cloud should be able to bear large data transfers in a fast and reliable way. Given their high data rates and low latency, optical networks based on wavelength division multiplexing (WDM) technology are ideally suited. These considerations motivate us to focus our attention on reducing the energy consumption of integrated optical network and IT infrastructures. It is clear that in order to identify an optimal solution achieving minimum energy consumption, joint consideration of both network and IT resources will be required.

Examining this type of infrastructure, one can identify the following elements to consider from an energy consumption

perspective: optical links, optical switching nodes and data centers. In this work we aim at reducing the overall energy consumption of such an infrastructure by employing two strategies:

- 1) Switching off components when they are in an idle state.
- 2) Exploiting the anycast principle to provision IT requests to the most appropriate data center and compute routes in an energy-efficient way.

Anycast is based on the principle that a user is not concerned with the exact location of the execution of the submitted IT request, as long as the requirements of the service are met. Hence, when operating an infrastructure such as the one described above, the selection of both the destination IT site and the network resources that allow the routing of the IT service from a remote user, can be based on the associated energy consumption. This can be performed by including the relevant energy parameters in the objective of the associated optimization, as opposed to unicast which is less flexible because the destination IT site is known a priori.

The remainder of the paper is organized as follows. In Section II we introduce related work focusing on energy efficiency in optical networks and, in Section III we formulate the power models for the optical network and the IT resources. In Section IV we formally articulate the problem and provide a Mixed Integer Linear Programming (MILP) approach. In Section V we provide a use case scenario providing results and insights. Finally we conclude the paper in Section VI.

II. RELATED WORK

The work in [4] reports a detailed study to estimate the impact of ICT on the environment in general and on energy needs in particular. According to predictions made, it is clear that the pressure on power efficiency in ICT will become more and more prominent in the coming years and needs to be dealt with accordingly.

The authors in [5] have investigated the influence of the availability of switching off network elements under connectivity and Quality of Service (QoS) constraints. Results show that it is possible to reduce the number of active links and nodes up to 25%. This work confirms that there is a network energy optimization to be made by switching off resources.

We extend the principle further by also considering IT power consumption and exploiting the anycast principle.

The work described in [6] provides a comprehensive survey of the most relevant research activities for minimizing energy consumption in telecom networks, with specific emphasis on those employing optical technologies. Energy-minimization opportunities enabled by optical technologies are investigated and classified over different network domains, namely core, metro and access networks.

An investigation of the potential savings achievable through power-aware network design and routing is presented in [7]. The authors have conducted measurements of the power consumption in various configurations of widely used core and edge routers and have explored the potential impact of power-awareness in a set of example networks. Results indicate that power consumption can vary by as much as an order of magnitude, indicating that there may be substantial opportunities for reducing power consumption.

An analysis of several designs for green routing algorithms is presented in [8]. The authors formulate the problem as a minimum energy routing optimization, where nodes cannot be switched off (as opposed to our work). It is demonstrated that depending on the topology and traffic matrices, the optimal energy savings can be modest for some scenarios. The authors also counteract the belief that there exists a trade-off between energy-efficient network optimization and performance.

The work presented in this paper extends previous work in two ways. We first develop a generic energy model for the integrated IT and optical network infrastructure where energy is consumed by the optical switching nodes and links as well as the data centers so as to treat network and IT power as part of the whole optimization objective. Secondly we are considering the anycast principle, as opposed to other works where a static traffic matrix is assumed. This allows us to choose the destination sites, in order to decrease the overall power consumption. Thirdly, we benchmark the proposed strategy compared to other traditional routing and allocation schemes in an extensive case study.

III. POWER CONSUMPTION MODELS

We aim to minimize the energy used by both the data centers and the optical network. Therefore it is imperative to rely on models that accurately describe the power consumption of the associated devices. In this work we assume optical switching nodes that are regenerating, wavelength convertible optical cross-connects based on a central optical switching fabric using 3D MEMS (Micro-Electro-Mechanical Systems) switch technology [9] described in detail in Section III-B. For the data center power consumption model we deploy a flexible power estimation framework described in III-A3.

A. IT Power Model

1) *Computer power consumption index*: As a data center can house hundreds or even thousands of servers and storage devices it is obvious that it is quite energy intensive. Apart from the servers, there are several other factors that add to the

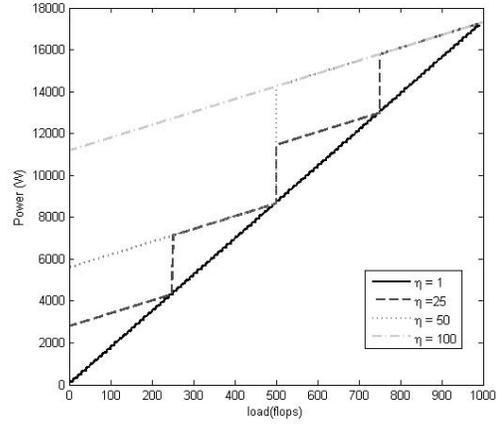


Fig. 1. Power consumption of a Data Center with $\eta \in 1, 25, 50, 100$ with 100 servers with server power characteristics $P_{max} = 118W$ and $P_{idle} = 56.7W$.

power consumption such as backup generators, switching gear, cooling systems, uninterruptible power supplies (UPS) and Power Distribution Units (PDU). To express this extra power consumption for a data center, we use the *computer power consumption index* (the inverse of Power Usage Effectiveness, PUE) which is the fraction of the total energy consumption of the data center to that used by the servers housed in the data center. The authors of [10] have investigated this index for 22 data centers and concluded that this ranges from a very poor 0.33 up to 0.75 where on average it is about 0.5. This means that half of the energy consumption of a data center goes to cooling etc. For OXCs, there is a similar index, which is of the same kind (about 0.5). Therefore, in the rest of the paper, we omit this extra power as it does not modify the relative network vs. IT power.

2) *Power Consumption of a server*: As the authors of [11] have demonstrated, linear eq. 1 produces quite accurate estimations for a server's power consumption, given the load (α), the power in idle state (P_{idle}) and the power when at 100% load (P_{max}). In this work, we express load in flops (FLoating point OPERations per Second). We use the same metric when expressing the capacity of a server (z) i.e. the maximum number of flops it is able to serve.

$$P_{cpu}(\alpha) = P_{idle} + \frac{(P_{max} - P_{idle})}{z} \cdot \alpha \quad (1)$$

3) *Power consumption of a data center*: For this study, we assume a data center which is able to switch off all servers in a certain rack when all servers in that rack are idle. To express the granularity in which a data center can switch off servers, we introduce the parameter η , which corresponds to a number of servers in a rack. For example when $\eta = 1$ the data center can switch on/off each single server, whereas when $\eta = 5$ the data center has to switch on/off a rack with five servers. Note that these servers consume their idle power P_{idle} when turned on. As shown in Section V this parameter directly influences

the routing and IT request allocation scheme.

Accordingly, the power consumption of a data center is expressed in eq. 2 (α = total load offered to the data center). In Fig. 1 we have plotted the power consumption of a data center (a stepwise function), with different values for η , depending on the load.

$$P_{DC}(\alpha) = \left\lfloor \frac{\alpha}{z \cdot \eta} \right\rfloor \cdot P_{max} + \eta \cdot P_{idle} + \frac{P_{max} - P_{idle}}{z} \cdot \left(\alpha - \left\lfloor \frac{\alpha}{z \cdot \eta} \right\rfloor \right) \quad (2)$$

B. Network Power Model

The node architecture considered in this work is an optical cross-connect based on an optical switching fabric. The OXC supports N input and N output fibers, each employing a maximum number of wavelengths, W . The total number of ports is the sum of express (through) and add/drop ports. To overcome the limitations that the wavelength continuity constraint imposes in optical networks, we assume full wavelength conversion capability. This is facilitated through the allocation of a wavelength converter at the output of every through switching port based on conventional optoelectronic transponder technology offering at the same time signal regeneration. Moreover, the OXC architecture employed supports the ability to add/drop up to 50% of the total through traffic. As shown in Fig. 2, one transmitter for each add port and one receiver for each drop port is needed. The total power consumption of the node depends on four parts: the switch fabric, the wavelength converters (transponders), the transmission equipment (transmitters (add), receivers (drop)), and the optical amplifiers based on Erbium Doped Fiber Amplifier (EDFA) technology.

Hence, the power consumption of the OXC n is then calculated as follows (all parameters are explained in Section IV):

$$\begin{aligned} P_n &= P_{mems} + P_{oeo} + P_{ampl} + P_{trans} \quad (3) \\ P_{mems} &= \rho_{total} \cdot P_{pair} \\ P_{oeo} &= \rho_{through} \cdot P_p \\ P_{ampl} &= (\omega^+(n) + \omega^-(n)) \cdot P_{edfa} \\ P_{trans} &= \rho_{a/d} \cdot P_{Tx/Rx} \end{aligned}$$

In equation (3) ρ_{total} represents the total number of switch ports and $\rho_{through}$ the the number of express (through) ports. The total power consumption of the optical network can be derived by the addition of the power consumption of all active switching nodes and links as described in detail in [12] and therefore it can be expressed as indicated by eq. 4 including the energy consumption of both optical switching nodes and links. The power consumption of optical links is attributed to the optical amplifiers used to compensate the insertion loss associated with signal transmission over optical fibers. The distance between consecutive amplifiers is referred to as the *fibres span*.

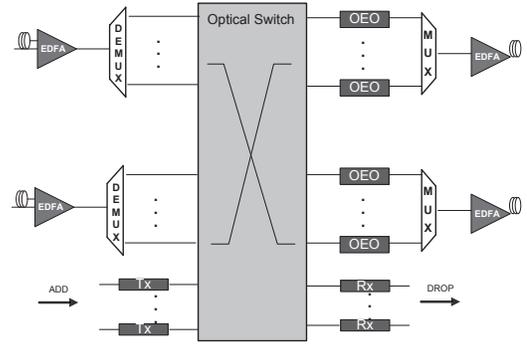


Fig. 2. The OXC architecture, illustrating the power-dissipating elements of the OXC with gray color.

$$P_{netw} = \sum_{n \in V} P_n + \sum_{l \in E} \left\lceil \frac{|l|}{\text{fibres span}} \right\rceil \cdot P_{edfa} \quad (4)$$

IV. FORMAL PROBLEM STATEMENT AND FORMULATION

As opposed to traditional routing schemes, we are not assuming a traffic matrix representing the number of connections requested between each source and destination site. For the offline provisioning of the network we assume a traffic vector expressing a number of desired connections between a source and some server site in the network that is not predetermined. Hence, it is up to the MILP to decide which server site best suites the objective function. Furthermore we assume that every OXC is able to perform full wavelength conversion.

For the IT part, we assign a number of flops each request expresses its need for. For simplicity we assume the same number of flops per request.

The MILP employs the following parameters:

- G = (N, L) , a graph representing an optical network
- N Node set, indexed by $n \in N$
- L The link set, with 1 fibre per link, index by l
- K Request set, indexed by $k \in K$
- k_n Request originating at n
- $z_n \in [0, \infty[$. The maximum number of flops server n can process
- m number of servers in a data center
- η switch on/off granularity of a data center
- $f^k \in [0, \infty[$. The number of flops request k needs
- P_{mems} power consumption of the switching fabric (e.g. MEMS),
- P_{oeo} power consumption attributed for wavelength conversion (e.g. by using the OEO converters)
- P_{ampl} power consumption of all the amplifiers of the OXC
- P_{trans} power consumption required by the transmitters and receivers
- P_{pair} power consumption per input/output port pair [13], 0.107 W
- P_{edfa} power consumption of an EDFA, 13 W
- P_c power consumption of a receiver, 3.5 W
- P_m power consumption of a transmitter, 3.5 W

P_p power consumption of a transponder (OEO converter), 6 W

P_{max} power consumption of a server at full load 118 W

P_{min} power consumption of a server in idle state 56.7 W

We use the following notations:

$\omega^-(n)$ The outgoing fibers of OXC n

$\omega^+(n)$ The incoming fibers of OXC n

$\omega(n) = \omega^-(n) \cup \omega^+(n)$

$|l|$ length of link l

The variables in the MILP are as follows:

$w_l \in [0, \infty[$ number of active wavelengths on $l \in L$. We will assume all links have the same maximum transport capacity (in terms of wavelengths), say $w_l \leq W$ for all $l \in L$

$w_l^k \in \{0, 1\}$ Is 1 if request k is routed over fiber l , 0 otherwise

$f_l \in \{0, 1\}$ Is 1 if fiber l is used, meaning at least one of its wavelengths is activated

$y_n \in \{0, 1\}$ Is 1 if OXC n is powered on, 0 otherwise

$x_n \in \{0, 1\}$ Is 1 if data center n is powered on, 0 otherwise

$b_n^k \in \{0, 1\}$ Is 1 if data center n processes request k

$b_n \in [0, \infty[$ The total number of requests a data center n is processing

$\alpha_n \in [0, \infty[$ Load offered to data center n , in flops.

$\beta_n \in [0, \infty[$ Number of racks which are switched on, in data center n

$\gamma_n \in [0, \infty[$ The power used by all servers at full load in data center n

$d_n \in [0, \infty[$ The power consumption of data center n

$o_n \in [0, \infty[$ The power consumption of OXC n

$s_n \in [0, \infty[$ The number of switched paths in node n

$t_n \in [0, \infty[$ The number of paths terminated in node n

$r_n \in [0, \infty[$ The number of locally processed jobs in node n

A. Objectives

We have implemented four different objectives. The first objective (5) minimizes only the network energy (referred to as N), the second objective (6) minimizes the IT energy and applies shortest path routing (referred to as I), the third objective (7) applies shortest path routing while the last objective (8) minimizes both IT and network energy (referred to as NI). We have used a δ variable in order to achieve shortest path routing in I , as otherwise random routes would be taken. In order to achieve this shortest path routing in I , we need to keep this δ very small therefore we used 0.001.

a) *Objective N:*

$$\min \left(\sum_{n \in N} o_n + \sum_{l \in L} f_l \cdot \left[\frac{l}{\text{fibre span}} \right] \cdot P_{edfa} \right) \quad (5)$$

b) *Objective I:*

$$\min \left(\sum_{n \in N} d_n + \delta \cdot \sum_{l \in L} w_l \cdot |l| \right) \quad (6)$$

c) *Objective SP:*

$$\min \left(\sum_{l \in L} w_l \cdot |l| \right) \quad (7)$$

d) *Objective NI:*

$$\min (N + I) \quad (8)$$

B. Constraints

1) *Network Modeling:* We start by formulating the flow conservations:

$$\sum_{l \in \omega^+(n)} w_l^k - \sum_{l \in \omega^-(n)} w_l^k = \begin{cases} -1 & \text{if } n \text{ is } k\text{'s source} \\ b_n^k & \text{otherwise} \end{cases} \quad n \in N, k \in K. \quad (9)$$

The next set of constraints represent the demand constraints:

$$\sum_{n \in N} b_n^k = 1, \quad \forall k \in K \quad (10)$$

$$b_n = \sum_{k \in K} b_n^k, \quad b_n \leq z_n, \quad \forall n \in N \quad (11)$$

We enforce the network capacity constraints:

$$w_l = \sum_{k \in K} w_l^k, \quad w_l \leq W, \quad \forall l \in L \quad (12)$$

In the next set of constraints we calculate whether a link or an OXC is turned on or not. (M is the node degree of n):

$$f_l \leq w_l, \quad f_l \cdot W \geq w_l, \quad \forall l \in L \quad (13)$$

$$\sum_{l \in \omega(n)} f_l \leq M \cdot y_n, \quad y_n \leq \sum_{l \in \omega(n)} f_l, \quad \forall n \in N \quad (14)$$

We apply these constraints to compute the number of dropped, added and switched paths:

$$r_n = \sum_{K_n \in K} b_n^k, \quad s_n = \left(\sum_{l \in \omega^+(v)} w_l \right) - t_n, \quad \forall n \in N \quad (15)$$

$$t_n = b_n - r_n, \quad \forall n \in N \quad (16)$$

This last constraint represents the power consumption of the OXC, as described in Section III-B.

$$o_n = \sum_{l \in \omega(n)} f_l \cdot P_{edfa} + y_n \cdot P_{mems} + t_n \cdot P_c + (K_n - r_n) \cdot P_m + s_n \cdot P_p, \quad \forall n \in N$$

2) *IT modeling:* We apply demand constraints for the IT requests.

$$\sum_{k \in K} b_n^k \cdot f^k \leq z_n \cdot m, \quad \forall n \in N \quad (17)$$

Next, we check whether a data center is turned on only if it processes requests.

$$x_n \cdot M \geq b_n, \quad x_n \leq b_n, \quad \forall n \in N \quad (18)$$

The next set of constraints are used to compute the power of a data center, depending on η ($\beta \in \mathbb{N}$).

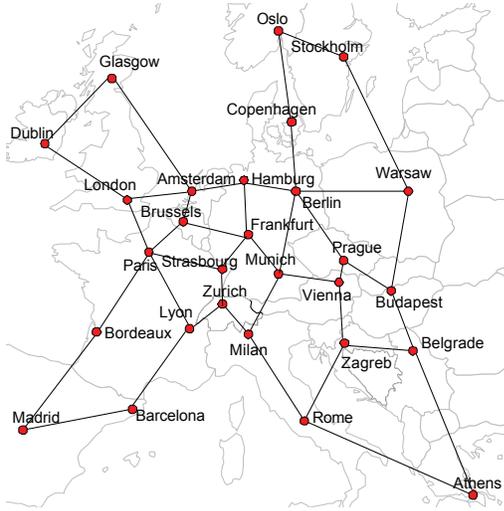


Fig. 3. The reference topology [14] used for obtaining the results.

$$\alpha_n = \sum_k b_n^k \cdot f^k, \quad \gamma_n = \beta_n \cdot z_n \cdot \eta, \quad \forall n \in N(19)$$

$$\gamma_n \leq \alpha, \quad \gamma_n \geq \alpha - (z_n \cdot \eta + 1), \quad \forall n \in N(20)$$

And finally we compute the power of a data center (only if $z_n > 0$) according to Section III-A3.

$$d_n = \beta_n \cdot \eta \cdot P_{max} + \eta \cdot P_{idle} + \frac{(P_{max} - P_{idle})}{z_n} \cdot (\alpha - \gamma_n) \quad (21)$$

C. Complexity

The scalability and complexity of a MILP mainly depends on the number of variables and constraints which are employed. For this MILP the number of variables is $|N| \cdot (11 + |K|) + |L| \cdot (2 + |K|)$ and the number of constraints is $16 \cdot |N| + |K| \cdot (|N| + 1) + 4 \cdot |L|$. As can be observed the MILP is scaling with the number of requests and the number of nodes in the network.

V. USE CASE

The network topology [14] used in this work is the European topology (shown in Fig. 3), which is a result of a joint effort from the IST LION project and the COST 266 action project [15]. We assumed 5 server sites, each with 20 servers (ASUS RS160-E5 Intel Xeon L5420 Processor, 2.50 GHz [16]), located at Berlin, London, Lyon, Vienna and Zurich. For the network we assumed 20 wavelengths per fiber. We generated 10 random demand vectors per demand instance ranging from 10 to 100 requests where each request needed one server ($f^k = z_n$ and 1 wavelength path towards the respective server site). Consequently, the results shown in the graphs represent averages of these 10 demand instances. Note that, as there are five server sites each incorporating 20 servers, 100 IT requests correspond to a load that requires all datacenters to be working at full capacity.

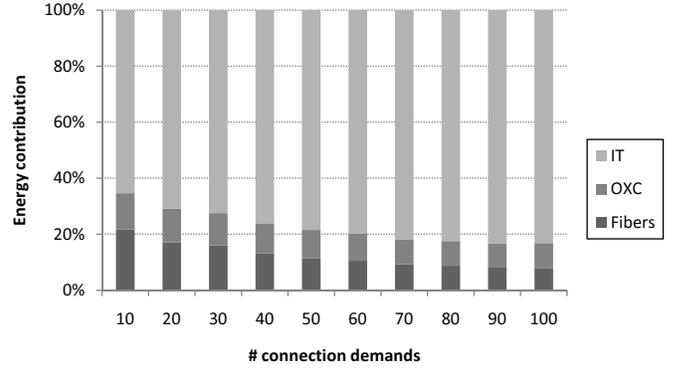


Fig. 4. Relative power consumption of OXCs, links and data centers compared to the total power consumption for $\eta = 1$.

A. Network energy aware routing vs. Network+IT energy aware routing

In Fig. 4 we show the distribution of the different power dissipating elements and the extent to which they take part in the total energy consumption for the scenario with $\eta = 1$. In this scheme there are no means to optimize the IT resource allocation, as every IT request can be scheduled to a server without the need to switch on other unnecessary servers. Hence, this is the minimum IT power needed to accommodate the IT load. We notice that, independent of the load, predominant energy consuming resources are the data centers (with their corresponding servers). This result indicates that intelligent IT resource allocation will probably be more beneficial than energy aware allocation of network resources only. This is demonstrated in Fig. 5 where the parameter $1 - P_{netw}^N / P_{netw}^{NI}$ has been plotted. P_{netw}^X is the total network power, for the MILP with objective minimization X . This data represents the extra percentage of network power needed to accommodate for the optimal energy aware IT resource allocation NI , compared to the pure network energy minimization objective N , in order to achieve the reduction of the total energy shown in Fig. 6 (up to 55% for $\eta = 20$). There are two observations in Fig. 5: (i) by allowing a suboptimal solution for the network routing, we enable a general decrease in the overall power consumption due to improved scheduling of IT requests and (ii) for increasing η this extra fraction of network power generally increases (together with the difference in power use). This can be explained by the fact that for bigger η , the use of a server always introduces powering on a whole rack of η servers. Given that IT power consumption is dominant (see Fig. 4) one strives to fully utilize complete racks in the NI optimization, leading to longer (more network power consumption) paths than strictly necessary.

Concluding it is clear that when the intention is to minimize the total energy consumption in an integrated network and IT scenario, the optimal solution should be a careful consideration of combined network and IT resource energy parameters.

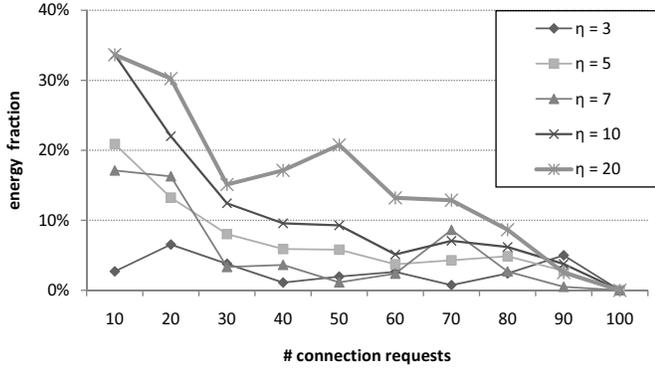


Fig. 5. The percentage of extra power needed to accommodate for intelligent routing in the NI case, compared to the pure network energy optimization case.

B. Comparing the routing schemes with different objectives

As indicated in IV-A we have opted for four different objectives: (i) shortest path routing (SP), (ii) minimization of network energy (N), (iii) minimization of IT energy with SP routing (I) and (iv) minimization of network and IT energy (NI). In Fig. 6 we have shown the total energy consumption values for the different objectives, for different η .

We first focus on the pure network objectives N and SP where we notice a similar power consumption. Both result in a similar total power with a difference of ca. 2% and from a total energy perspective the best objective is N . This observation can be attributed to the fact that the N scheme in most cases selects the closest server site (the same as SP) and therefore both SP and N objectives reach the same IT power. So the only difference in the energy consumption can be attributed to intelligently routing paths in order to allow switching off certain links and nodes and increase the sharing of network resources among paths. This way we can achieve a *network* energy reduction up to 10% compared to SP (depending on the load), but due to the balance between IT and network power (ca. 20% Network power vs. ca. 80% IT power, see Fig. 4), this decrease is translated into a small percentage of the overall energy consumption of the infrastructure (on average 2%). Moreover, this difference in the total energy consumption between Network and SP becomes smaller as η increases, because then the IT resource power consumption becomes even more dominant.

When comparing the IT-aware objectives I and NI , we once more note that there is little or no difference (ca. 2%) for the total energy consumption, with NI always providing the optimal solution. When we differentiate between IT and network energy, we note that there is never a difference between objectives regarding IT power consumption (as was suggested in Section V-A). NI can decrease its network power consumption by 5% to 10% compared to IT-only minimization, by providing energy-efficient routes the same way as already indicated above. However, as this network power decrease only accounts for a small portion of the total energy consumption (see Fig. 4), this decrease is hardly

noticeable (a decrease of 2%). Observing the unused and still available network resources it is seen that there is little or no difference between the two objectives: the free capacity in NI is only 1%-2% lower than that for I . These results clearly indicate that there is practically no penalty in the efficiency of the resource utilization introduced through the reduction of the network power consumption.

In Fig. 6 we show the total power consumption graphs for $\eta \in [7, 10, 20]$. We see that with increasing η , the difference between the pure network objectives (SP and N) and the objectives incorporating IT power parameters (I and NI) increases. This can be explained by the fact that the requirement to switch on a rack of servers, when allocating an IT request to an element of that rack, increases the penalty brought on by IT-unaware routing which is too large compared to the potential network savings. We conclude that the selection of the allocation scheme could depend on the granularity of set of servers (e.g. a rack) that can jointly be turned on/off. If η is rather small ($\eta \in [1, 2, 3]$), the only considerable optimization is one for the network power. If the absolute minimum energy consumption is targeted, N is to be favored. If some tolerance is allowed, SP will yield an acceptable solution in a shorter time frame. As η increases ($\eta \geq 4$), the IT-unaware routing introduces a high IT related energy penalty (up to 55% for $\eta = 20$) and the combined resource allocation scheme needs to be considered if the absolute minimum is requested, while I will yield acceptable results if some margin is allowed.

VI. CONCLUSION AND FUTURE WORK

Energy considerations in ICT are becoming of significant importance, as it is shown that ICT is responsible for about 10% of the global energy consumption. Therefore this work addresses the energy efficient operation of integrated network and IT infrastructures in the context of cloud computing. By allowing to switch off several IT and network elements and by exploiting the anycast principle, we propose an energy-efficient routing and IT allocation algorithm, using MILP. Results gathered from a use case on an European topology, demonstrated that the predominant energy consuming resources are the servers installed in the data centers, as they are responsible for ca. 80% of the total power consumption. If only the network energy consumption is taken into account in deciding to which IT server site requests are allocated, considerable energy waste may be introduced. More specifically, comparing joint minimization of both network and IT energy provides energy savings of the order of 3%-55% compared to the network energy minimization only approach, depending on the ability of a data center to switch on/off a set of servers (e.g. a rack). On the other hand, pure network-energy minimization allows energy savings of the order of 1-2% of the total energy budget compared to shortest path routing (i.e. energy-unaware). Future work includes investigating the effect of choice of server site locations on the energy savings and creating a more scalable method of computation (e.g. column generation, heuristics, etc.).

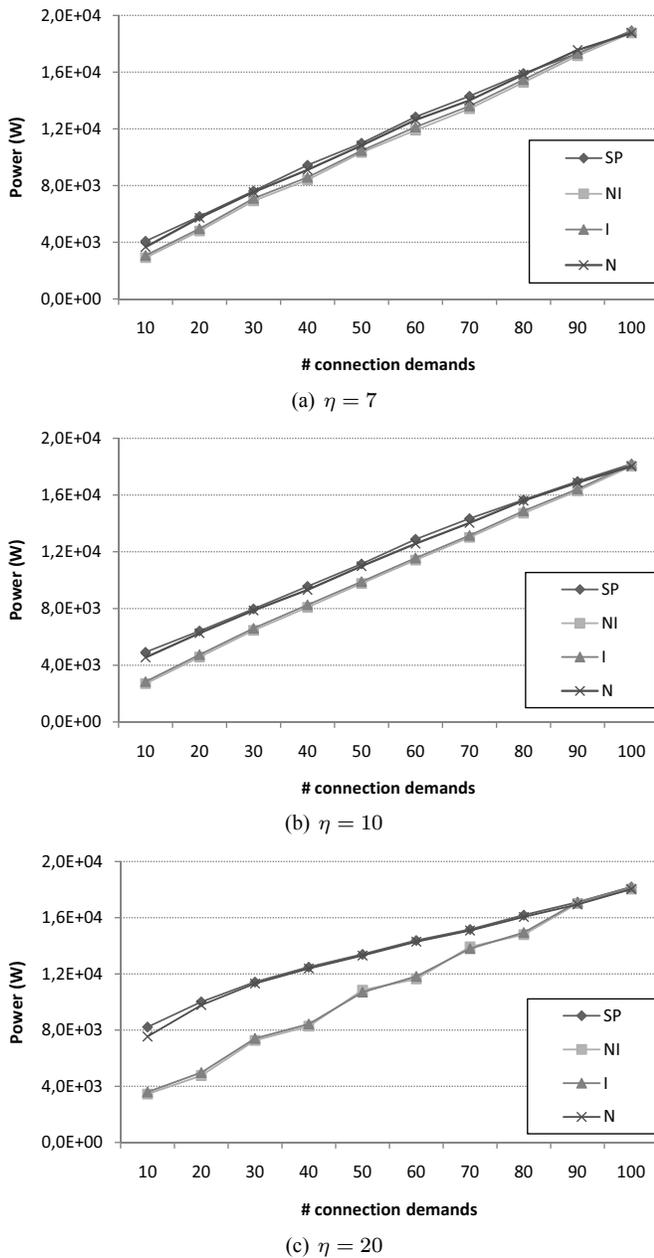


Fig. 6. Total power consumption, for each optimization objective (SP, NI, I and N) for $\eta = 7, 10, 20$

ACKNOWLEDGMENT

This work was carried out with the support of the GEYSERS (FP7-ICT-248657) project funded by the European Commission through the 7th ICT Framework Program. Jens Buysse is supported by the research institute IWT. C. Develder and M. De Leenheer are supported by the Research Foundation Flanders (FWO) as post-doctoral fellows. All computations were carried out using the Stevin Supercomputer Infrastructure at Ghent University.

REFERENCES

- [1] "Usage and population statistics," Internet World Stats, <http://www.internetworldstats.com/stats.htm>, Tech. Rep., 2011.
- [2] "An inefficient truth - executive summary," Global Action Plan, www.globalactionplan.org.uk, Tech. Rep., 2007.
- [3] M. Peter and G. Timothy, "The NIST definition of cloud computing," Institute of Standards and Technology, http://csrc.nist.gov/publications/drafts/800-145/Draft-SP-800-145_cloud-definition.pdf, Tech. Rep., Jan 2011.
- [4] M. Pickavet, W. Vereecken, S. Demeyer, P. Audenaert, B. Vermeulen, C. Develder, D. Colle, B. Dhoedt, and P. Demeester, "Worldwide energy needs for ICT- the rise of power-aware networking," in *Proc. of the 2nd International Symposium on Advanced Networks and Telecommunication Systems, 2008. ANTS 08.*, 2008, pp. 1-3. [Online]. Available: <http://dx.doi.org/10.1109/ANTS.2008.4937762>
- [5] L. Chiaraviglio, M. Mellia, and M. Neri, "Energy-aware networks: Reducing power consumption by switching off network elements," in *FEDERICA-Phosphorus tutorial and workshop (TNC2008)*, May 2008.
- [6] Z. Yi, P. Chowdhury, M. Tornatore, and B. Mukherjee, "Energy efficiency in telecom optical networks," *IEEE Communications Surveys Tutorials*, vol. 12, no. 4, pp. 441-458, 2010.
- [7] J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsang, and S. Wright, "Power awareness in network design and routing," in *Proc. of the 27th Conference on Computer Communications. INFOCOM 2008.*, Apr 2008, pp. 457-465.
- [8] A. Bianzino, C. Chaudet, F. Larroca, D. Rossi, and J. Rougier, "Energy-aware routing: A reality check," in *Proc. of 3rd Int. Workshop on Green Commun. (GreenComm3)*, in conjunction with Globecom 2010, Miami, FL, USA, Dec 2010.
- [9] M. Makoto, "Analyzing power consumption in optical cross-connect equipment for future large-capacity optical networks," *Journal of Networks*, vol. 5, no. 11, pp. 1-4, 2010.
- [10] S. Greenberg, M. Evan, T. Bill, R. Peter, and M. Bruce, "Best practices for data centers: Lessons learned from benchmarking 22 data centers." *Proc. of the ACEEE Summer Study on Energy Efficiency in Buildings in Asilomar, CA. ACEEE*, vol. 3, pp. 76-87, 2006.
- [11] X. Fan, W. Weber, and L. Barroso, "Power provisioning for a warehouse-sized computer," in *Proc. of the 34th annual international symposium on Computer architecture, SCA 07.* New York, NY, USA: ACM, 2007, pp. 13-23.
- [12] A. Tzanakaki, K. Katrinis, T. Politi, A. Stavdas, M. Pickavet, P. Van Daele, D. Simeonidou, M. J. O. Mahony, A. Slaviša, L. Wosinska, and P. Mont, "Power considerations towards a sustainable pan-european network," in *Proc. of the National Fiber Optic Engineers Conference (NFOEC/OFC 2011)*, Optical Society of America, Mar 2011.
- [13] A. Slaviša, "Analysis of power consumption in future high-capacity network nodes," *Journal of Optical Communication Networks*, vol. 1, no. 3, pp. 245-258, 2009. [Online]. Available: <http://jocn.osa.org/abstract.cfm?URI=JOCN-1-3-245>
- [14] S. De Maesschalck, D. Colle, I. Lievens, M. Pickavet, P. Demeester, C. Mauz, M. Jaeger, R. Inkret, B. Mikac, and J. Derkacz, "Pan-european optical transport networks: An availability-based comparison," *Photonic Network Communications*, vol. 5, no. 3, pp. 203-225, May 2003.
- [15] A. Kuchar, "Achievements of cost 266 action and further prospects in research of advanced infrastructure for photonic networks," in *6th International Conference on Transparent Optical Networks, 2004.*, vol. 1, Jul 2004, pp. 37 - 42 vol.1.
- [16] S. P. E. Corporation, "Specpower," SPEC, http://www.spec.org/power_ss/2008/, Tech. Rep., 2008.