Distributed Energy Efficient Clouds Over Core Networks

Ahmed Q. Lawey, Taisir E. H. El-Gorashi, and Jaafar M. H. Elmirghani

***Abstract*—In this paper, we introduce a framework for designing energy efficient cloud computing services over non-bypass IP/WDM core networks. We investigate network related factors including the centralization versus distribution of clouds and the impact of demand, content popularity and access frequency on the clouds placement, and cloud capability factors including the number of servers, switches and routers and amount of storage required in each cloud. We study the optimization of three cloud services: cloud content delivery, storage as a service (StaaS), and virtual machines (VMS) placement for processing applications. First, we develop a mixed integer linear programming (MILP) model to optimize cloud content delivery services. Our results indicate that replicating content into multiple clouds based on content popularity yields 43% total saving in power consumption compared to power un-aware centralized content delivery. Based on the model insights, we develop an energy efficient cloud content delivery heuristic, DEER-CD, with comparable power efficiency to the MILP results. Second, we extend the content delivery model to optimize StaaS applications. The results show that migrating content according to its access frequency yields up to 48% network power savings compared to serving content from a single central location. Third, we optimize the placement of VMs to minimize the total power consumption. Our results show that slicing the VMs into smaller VMs and placing them in proximity to their users saves 25% of the total power compared to a single virtualized cloud scenario. We also develop a heuristic for real time VM placement (DEER-VM) that achieves comparable power savings.**

***Index Terms*—Cloud computing, content delivery, energy consumption, IP/WDM, popularity, StaaS, virtual machines.**

I. Introduction

CLOUD computing exploits powerful resource management techniques to allow users to share a large pool of computational, network and storage resources over the Internet. The concept is inherited from research oriented grid computing and further expanded toward a business model where consumers are charged for the diverse offered services [1]. Cloud computing is expected to be the main factor that will dominate the future Internet service model [2] by offering a network based rather than desktop based users applications [3].

Virtualization [4] lies at the heart of cloud computing, where the requested resources are created, managed and removed flexibly over the existing physical machines such as servers, storage and networks. This opens the doors towards resource consolidation that cut the cost for the cloud provider and eventually, cloud consumers. However, cloud computing elastic management and economic advantages come at the cost of increased concerns regarding their privacy [5], availability [6] and power consumption [7]. Cloud computing has benefited from the work done on datacenters energy efficiency [7]. However, the success of the cloud relies heavily on the network that connects the clouds to their users. This means that the expected popularity of the cloud services has implications on network traffic, hence, network power consumption, especially if we consider the total path that information traverses from clouds storage through its servers, internal LAN, core, aggregation and access network up to the users' devices. For instance, the authors in [8] have shown that transporting data in public and sometimes private clouds might be less energy efficient compared to serving the computational demands by traditional desktop.

Designing future energy efficient clouds, therefore, requires the co-optimization of both external network and internal clouds resources. The lack of understanding of this interplay between the two domains of resources might cause eventual loss of power. For instance, a cloud provider might decide to migrate virtual machines (VMs) or content from one cloud location to another due to low cost or green renewable energy availability, however, the power consumption of the network through which users data traverse to/from the new cloud location might outweigh the gain of migration.

The authors in [9] studied the design of disaster-resilient optical datacenter networks through integer linear programming (ILP) and heuristics. They addressed content placement, routing, and protection of network and content for geographically distributed cloud services delivered by optical networks. In [10] mixed integer linear programming (MILP) models and heuristics are developed to minimize delay and power consumption of clouds over IP/WDM networks. The authors of [11] exploited anycast routing by intelligently selecting destinations and routes for users traffic served by clouds over optical networks, as opposed to unicast traffic, while switching off unused network elements. A unified, online, and weighted routing and scheduling algorithm is presented in [12] for a typical optical cloud infrastructure considering the energy consumption of

Manuscript received August 25, 2013; revised November 20, 2013 and January 13, 2014; accepted January 14, 2014. Date of publication January 19, 2014; date of current version February 17, 2014. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), INTERNET (EP/H040536/1) and STAR (EP/K016873/1), and from King Abdulaziz University Deanship of Scientific Research (DSR), Gr/9/33. The first author would like to acknowledge his PhD scholarship awarded by the Iraqi Ministry of Higher Education and Scientific Research.

A. Q. Lawey and T. E. H. El-Gorashi are with the School of Electronic and Electrical Engineering, University of Leeds, Leeds, Yorkshire LS2 9JT, U.K. (e-mail: elaql@leeds.ac.uk; t.e.h.elgorashi@leeds.ac.uk).

J. M. H. Elmirghani is with the School of Electronic and Electrical Engineering, University of Leeds, Leeds, Yorkshire LS2 9JT, U.K. and also with the Department of Electrical and Computer Engineering, King Abdulaziz University, Jeddah, Saudi Arabia (e-mail: j.m.h.elmirghani@leeds.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2014.2301450

the network and IT resources. In [13], the authors provided an optimization-based framework, where the objective functions range from minimizing the energy and bandwidth cost to minimizing the total carbon footprint subject to QoS constraints. Their model decides where to build a data center, how many servers are needed in each datacenter and how to route requests.

In [14] we built a MILP model to study the energy efficiency of public cloud for content delivery over non-bypass IP/WDM core networks. The model optimizes clouds external factors including the location of the cloud in the IP/WDM network and whether the cloud should be centralized or distributed and cloud internal capability factors including the number of servers, internal LAN switches, routers, and amount of storage required in each cloud. This paper extends the work by (i) studying the impact of small content (storage) size on the energy efficiency of cloud content delivery (ii) developing a real time heuristic for energy aware content delivery based on the content delivery model insights, (iii) extending the content delivery model to study the Storage as a Service (StaaS) application, (iv) developing a MILP model for energy aware cloud VM placement and designing a heuristic to mimic the model behaviour in real time.

The remainder of this paper is organized as follows. Section II briefly reviews IP/WDM networks and clouds. In Section III, we introduce our MILP for content delivery, discuss its results and propose the DEER-CD real time heuristic. Section IV extends the model of Section III to study Storage as a Service. In Section V, a MILP for VMs based cloud is introduced and a heuristic (DEER-VM) is proposed. Finally, Section VI concludes the paper.

II. Clouds in IP/WDM Networks

The IP/WDM network consists of two layers, the IP layer and the optical layer. In the IP layer, an IP router is used at each node to aggregate data traffic from access networks. Each IP router is connected to the optical layer through an optical switch. Optical switches are connected to optical fiber links where a pair of multiplexers/demultiplexers is used to multiplex/demultiplex wavelengths [15]. Optical fibers provide the large capacity required to support the communication between IP routers. Transponders provide OEO processing for full wavelength conversion at each node. In addition, for long distance transmission, EDFAs are used to amplify the optical signal on each fiber. Fig. 1 shows the architecture of an IP/WDM network.

Two approaches can be used to implement the IP/WDM network, namely, lightpath bypass and non-bypass. In the bypass approach, lightpaths are allowed to bypass the IP layer of intermediate nodes eliminating the need for IP router ports, the most power consuming devices in the network, which significantly reduces the total network power consumption compared to the non-bypass approach. However, implementing such an approach involves many technical challenges such as the need for long reach, low power optical transmission systems, limitations with the loss of electronic processing and as such the advantages electronic processing provides at intermediate nodes in terms of grooming, shared protection [16], and deep packet inspection. On the other hand, the forwarding decision in the non-bypass approach is made at the IP layer; therefore, the incoming lightpaths go through OEO conversion at each intermediate node. The non-bypass approach is implemented in most of the current IP/WDM networks. In addition to the simple implementation, the non-bypass approach allows operators to perform traffic control operations such as deep packet inspection and other analysis measures.

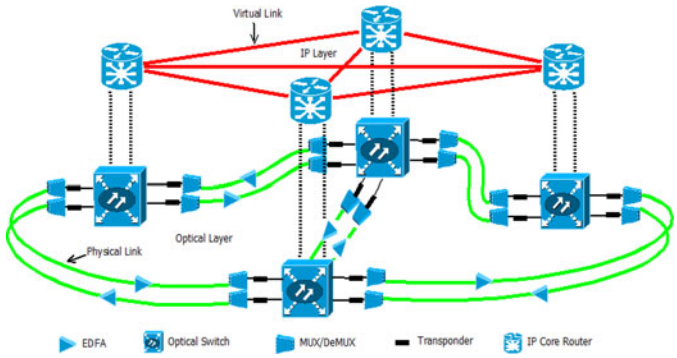

Fig. 1. IP/WDM Network.

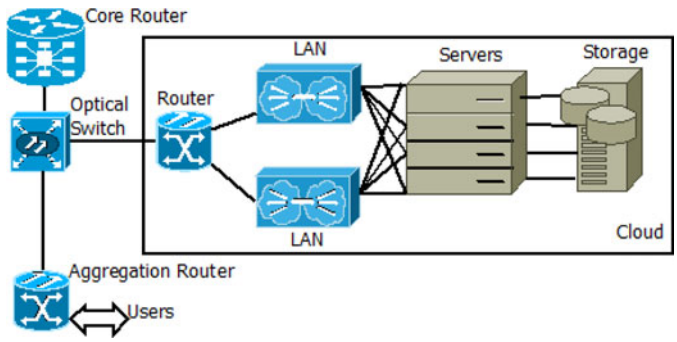

Fig. 2. Cloud architecture.

A number of papers in the literature have investigated the energy efficiency of IP/WDM networks. The authors in [15] have shown that the lightpath bypass approach usually reduces the power consumption compared to the non-bypass approach as bypassing the IP layer at intermediate nodes reduces the number of router ports, the major power consumers in IP/WDM networks. In [17], the authors focused on reducing the CO_2 emission of backbone IP/WDM networks by introducing renewable energy sources. In [18], a MILP model is developed to optimize the location of datacentres in IP/WDM networks as a means of reducing the network power consumption. In [19], energy-efficient IP/WDM physical topologies are investigated considering different IP/WDM approaches, nodal degree constraints, traffic symmetry and renewable energy availability. In this paper we evaluate our proposed cloud models over a non-bypass IP/WDM network as non-bypass is still the most widely implemented approach.

A typical cloud consists of three main parts, namely; servers, internal LAN and storage. Clouds are usually built very near to core network nodes to benefit from the large bandwidth offered by such nodes to serve users. Fig. 2 shows how the different parts inside the cloud are connected and how the cloud is connected to the core network. If the cloud is serving users located at another

core node, traffic will flow through the optical switch and core router on its way towards the core network. On the other hand if the users are located on the same node, the traffic will flow through the optical switch on its path towards the aggregation router where it will be routed to local users. The core/edge network power consumption of the second scenario is limited to the optical switch and aggregation router.

III. Energy Efficient Content Delivery Cloud

A. Content Delivery Cloud MILP Model

Jointly optimizing content distribution for content providers (CPs) and traffic engineering for Internet service providers (ISPs) is studied in [20] from QoS perspective. The authors in [21] studied the same problem from energy point of view where ISP and CP cooperate to minimize energy. In [22], the authors compared conventional and decentralized server based content delivery networks (CDN), content centric networks (CCN), and centralized server based CDN using dynamic optical bypass where they took popularity of content into account. In their conventional CDN model, content is fully replicated to all datacenters regardless of content popularity. They showed that CCN is more energy efficient in delivering the most popular content while CDN with optical bypass is more energy efficient in delivering less popular content.

In this section we introduce the MILP model developed to minimize the power consumption of the cloud content delivery service over non-bypass IP/WDM networks. Given the client requests, the model responds by selecting the optimum number of clouds and their locations in the network as well as the capability of each cloud so that the total power consumption is minimized. The model also decides how to replicate content in the cloud according to its popularity so the minimum power possible is consumed in delivering content. A key difference between our content delivery model and the work done in the literature is the extensive study of the impact of content popularity among different locations where we compare content replication schemes.

Below we re-introduce our model developed in [14] for completeness. We assume the popularity of the different objects of the content follows a Zipf distribution, representative of the popularity distribution of several cloud content types such as YouTube and others [23] where the popularity of an object of rank i is given as follows:

$$P(i) = \varphi / i$$

where $P(i)$ is the relative popularity of the object of rank i and φ is

$$\varphi = \left(\sum_{i=1}^{N} \frac{1}{i} \right)^{-1}.$$

We divide the content in our model into equally sized popularity groups. A popularity group contains objects of similar popularity.

We define the following variables and parameters to represent the IP/WDM network:

Parameters:

N	Set of IP/WDM nodes.
Nm_i	Set of neighbors of node *i*.
$\|N\|$	Number of IP/WDM nodes.
Prp	Router port power consumption.
Pt	Transponder power consumption.
Pe	EDFA power consumption.
PO_i	Power consumption of optical switch installed at node $\boldsymbol{i} \in \boldsymbol{N}$.
Pmd	Multi/demultiplexer power consumption.
W	Number of wavelengths per fiber.
B	Wavelength bit rate.
S	Max span distance between EDFAs.
D_{mn}	Distance between node pair (*m, n*).
A_{mn}	Number of EDFAs between node pair (*m, n*).
PUE_n	IP/WDM network power usage effectiveness.
M	A large enough number.
Δt	Time granularity, which represents the evaluation period.

Variables (All are Nonnegative Real Numbers)

C_{ij}	Number of wavelengths in the virtual link (i, j).
L_{ij}^{sd}	Traffic flow between node pair (*s, d*) traversing virtual link (*i, j*).
W_{mn}^{ij}	Number of wavelength channels in the virtual link (*i, j*) traversing physical link (*m, n*).
W_{mn}	Total number of used wavelengths in the physical link (*m,n*).
F_{mn}	Total number of fibers on the physical link (*m,n*).
Q_i	Number of aggregation ports in router *i*.

Under the non-bypass approach, the total network power consumption is composed of [15], [18]:

1) The power consumption of router ports

$$\sum_{i \in N} Prp \cdot Q_i + \boldsymbol{Prp} \cdot \sum_{m \in N} \sum_{n \in N m_m} W_{mn}.$$

2) The power consumption of transponders

$$\sum_{m \in N} \sum_{n \in N m_m} Pt \cdot W_{mn}.$$

3) The power consumption of EDFAs

$$\sum_{m \in N} \sum_{n \in N m_m} Pe \cdot A_{mn} \cdot F_{mn}.$$

4) The power consumption of optical switches

$$\sum_{i \in N} PO_i.$$

5) The power consumption of multi/demultiplexers

$$\sum_{m \in N} \sum_{n \in N m_m} Pmd \cdot F_{mn}.$$

The content delivery cloud is represented by the following variables and parameters:

Parameters:

U_d	Set of users in node *d*.
PG	Set of popularity groups, $\{1 \ldots \|PG\|\}$.
$\|PG\|$	Number of popularity groups.

PUE_c Cloud power usage effectiveness.
S_PC Storage power consumption.
S_C Storage capacity of one storage rack in GB.
Red Storage and switching redundancy.
S_PPGB Storage power consumption per GB, $S\_PPGB = S\_PC/S\_C$.
S_Utl Storage utilization.
PGS_p Popularity group storage size, $PGS_p = (S\_C/|PG|) \cdot S\_Utl$.
CS_C Content server capacity.
CS_EPB Content server energy per bit.
Sw_PC Cloud switch power consumption.
Sw_C Cloud switch capacity.
Sw_EPB Cloud switch energy per bit, $Sw\_EPB = Sw\_PC/Sw\_C$.
R_PC Cloud router power consumption.
R_C Cloud router capacity.
R_EPB Cloud router energy per bit, $R\_EPB = R\_PC/R\_C$.
$Drate$ Average user download rate.
P_p Popularity of object p (Zipf distribution).
ND_d Node d total traffic demand, $ND_d = \sum_{i \in U_d} Drate$.
D_{Pd} Popularity group p traffic to node $d, D_{pd} = ND_d \cdot P_p$.

Variables (All are Nonnegative Real Numbers)

δ_{sdp} $\delta_{\mathrm{sdp}} = 1$ if popularity group p is placed in node s to serve users in node $d, \delta_{\mathrm{sdp}} = 0$ otherwise.
LP_{sdp} Traffic generated due to placing popularity group p in node $s$to serve users in node *d.*
L_{sd} Traffic from cloud $s$to users in node *d.*
Cup_s Cloud s upload capacity.
δ_{sp} $\delta_{sp} = 1$ if cloud s stores a copy of popularity group $p, \delta_{sp} = 0$ otherwise.
$Cloud_s$ $\mathrm{Cloud}_s = 1$ if a cloud is built in node $s, \mathrm{Cloud}_s = 0$ otherwise.
CN Number of clouds in the network.
CSN_s Number of content servers in cloud *s.*
SwN_s Number of switches in cloud *s.*
RN_s Number of routers in cloud *s.*
$StrC_s$ Cloud s storage capacity.

The cloud power consumption is composed of

1) The power consumption of content servers $(SrvPC\_CD)$:

$$\sum_{s \in N} Cup_s \cdot CS\_EPB.$$

2) The power consumption of switches and routers $(LANPC\_CD)$:

$$\sum_{s \in N} Cup_s \cdot (Sw\_EPB \cdot Red + R\_EPB).$$

3) The power consumption of storage $(StPC\_CD)$:

$$\sum_{s \in N} StrC_s \cdot S\_PPGB \cdot Red.$$

The model is defined as follows:

Objective: Minimize

$$\begin{aligned} PUE\_n \cdot \Bigg(& \sum_{i \in N} Prp \cdot Q_i + Prp \cdot \sum_{m \in N} \sum_{n \in N m_m} W_{mn} \\ & + \sum_{m \in N} \sum_{n \in N m_m} Pt \cdot W_{mn} \\ & + \sum_{m \in N} \sum_{n \in N m_m} Pe \cdot A_{mn} \cdot F_{mn} \\ & + \sum_{i \in N} PO_i + \sum_{m \in N} \sum_{n \in N m_m} Pmd \cdot F_{mn} \Bigg) \\ + PUE\_c \cdot \Bigg(& \sum_{s \in N} Cup_s \cdot CS\_EPB \\ & + \sum_{s \in N} Cup_s \cdot (Sw\_EPB \cdot Red + R\_EPB) \\ & + \sum_{s \in N} StrC_s \cdot S\_PPGB \cdot Red \Bigg). \end{aligned} \tag{1}$$

Equation (1) gives the model objective which is to minimize the IP/WDM network power consumption and the cloud power consumption

Subject to:

1) Flow conservation constraint in the IP layer

$$\sum_{j \in N: i \neq j} L_{ij}^{sd} - \sum_{j \in N: i \neq j} L_{ji}^{sd} = \begin{cases} L_{sd} & \text{if } i = s \\ -L_{sd} & \text{if } i = d \\ 0 & \text{otherwise} \end{cases}$$
$$\forall s, d, i \in N : s \neq d. \tag{2}$$

Constraint (2) is the flow conservation constraint for IP layer. It ensures that the total incoming traffic is equal to the total outgoing traffic for all nodes except for the source and destination nodes.

2) Virual link capacity constraint

$$\sum_{s \in N} \sum_{d \in N: s \neq d} L_{ij}^{sd} \leq C_{ij} \cdot B$$
$$\forall i, j \in N : i \neq j. \tag{3}$$

Constraint (3) ensures that the traffic traversing a virtual link does not exceed its capacity.

3) Flow conservation constraint in the optical layer

$$\sum_{n \in N m_m} W_{mn}^{ij} - \sum_{n \in N m_m} W_{nm}^{ij} = \begin{cases} C_{ij} & \text{if } m = i \\ -C_{ij} & \text{if } m = j \\ 0 & \text{otherwise} \end{cases}$$
$$\forall i, j, m \in N : i \neq j. \tag{4}$$

Constraint (4) represents the flow conservation for the optical layer. It ensures that the total number of outgoing wavelengths in a virtual link is equal to the total number of incoming wavelengths except for the source and destination nodes of the virtual link.

4) Physical link capacity constraints

$$\sum_{i\in N}\sum_{j\in N:i\neq j} W_{mn}^{ij} \leq W \cdot F_{mn} \quad \forall m \in N \ \forall n \in Nm_m. \tag{5}$$

$$\sum_{i\in N}\sum_{j\in N:i\neq j} W_{mn}^{ij} = W_{mn} \quad \forall m \in N \ \forall n \in Nm_m. \tag{6}$$

Constraints (5) and (6) represent the physical link capacity constraints. Constraint (5) ensures that the number of wavelength channels in virtual links traversing a physical link does not exceed the capacity of fibres in the physical link. Constraint (6) ensures that the number of wavelength channels in virtual links traversing a physical link is equal to the number of wavelengths in that physical link.

5) Number of aggregation ports constraint

$$Q_i = 1/B \cdot \sum_{d\in N:i\neq d} L_{id} \quad \forall i \in N. \tag{7}$$

Constraint (7) calculates the number of aggregation ports for each router. We use relaxation in our model due to the large number of variables. However, as we are interested in power consumption, relaxation has a limited impact on the final result as the difference will be within the power of less than one wavelength (router port and transponder) which is negligible compared to the total power consumption. The heuristics which produce comparable results provide independent verification for the model (and its relaxation), especially that the routing and placements in the heuristics follow approaches that are independent of the model.

6) IP/WDM network traffic

$$LP_{s\mathrm{dp}} = \delta_{s\mathrm{dp}} \cdot D_{Pd} \quad \forall s,d \in N \ \forall p \in PG \tag{8}$$

$$\sum_{s\in N} LP_{s\mathrm{dp}} = D_{Pd} \quad \forall d \in N \ \forall p \in PG. \tag{9}$$

$$L_{sd} = \sum_{p\in PG} LP_{s\mathrm{dp}} \quad \forall s,d \in N. \tag{10}$$

$$Cup_s = \sum_{d\in N} L_{sd} \quad \forall s \in N. \tag{11}$$

Constraint (8) calculates the traffic generated in the IP/WDM network due to requesting popularity group p that is placed in node s by users located in node d. Constraint (9) ensures that each popularity group request is served from a single cloud only. We have not included traffic bifurcation where a user may get parts of the content from different clouds. This is a useful extension to be considered in future. Constraint (10) calculates the traffic from the cloud in node s and users in node *d,* to be used in constraints (2) and (7). Constraint (11) calculates each cloud upload capacity based on total traffic sent from the cloud.

7) Popularity groups locations

$$\sum_{d\in N} \delta_{s\mathrm{dp}} \geq \delta_{sp} \quad \forall s \in N \ \forall p \in PG. \tag{12}$$

$$\sum_{d\in N} \delta_{s\mathrm{dp}} \leq M \cdot \delta_{sp} \quad \forall s \in N \ \forall p \in PG. \tag{13}$$

Constrains (12) and (13) ensure that popularity group p is replicated to cloud s if cloud s is serving requests for this popularity group, where M is a large enough unitless number to ensure that $\delta_{sp} = 1$ when $\sum_{\mathrm{d}\in\mathrm{N}} \delta_{s\mathrm{dp}}$ is greater than zero.

8) Cloud location and number of clouds

$$\sum_{p\in PG} \delta_{sp} \geq Cloud_s \quad \forall s \in N. \tag{14}$$

$$\sum_{p\in PG} \delta_{sp} \leq M \cdot Cloud_s \quad \forall s \in N. \tag{15}$$

$$CN = \sum_{s\in N} Cloud_s. \tag{16}$$

Constraints (14) and (15) build a cloud in location s if that location is chosen to store at least one popularity group or more, where M is a large enough unitless number to ensure that $\mathrm{Cloud}_s = 1$ when $\sum_{\mathrm{p}\in PG} \delta_{sp}$ is greater than zero.

Constraint (16) calculates total number of clouds in the network.

9) Cloud Capability

$$CSN_s = Cup_s/CS\_C \quad \forall s \in N. \tag{17}$$

$$SwN_s = (Cup_s/Sw\_C) \cdot Red \quad \forall s \in N. \tag{18}$$

$$RN_s = Cup_s/R\_C \quad \forall s \in N. \tag{19}$$

$$StrC_s = \sum_{p\in PG} \delta_{sp} \cdot PGS_p. \quad \forall s \in N. \tag{20}$$

Constraints (17)–(19) calculate the number of content servers, switches and routers required at each cloud based on cloud upload traffic going through these elements. The integer value is obtained using the ceiling function. Note that the number of

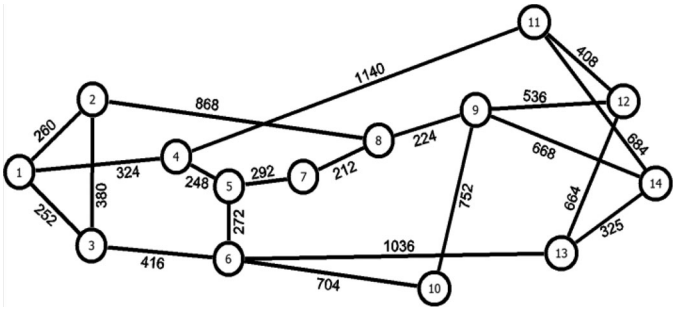

Fig. 3. The NSFNET network with link lengths in kilometer.

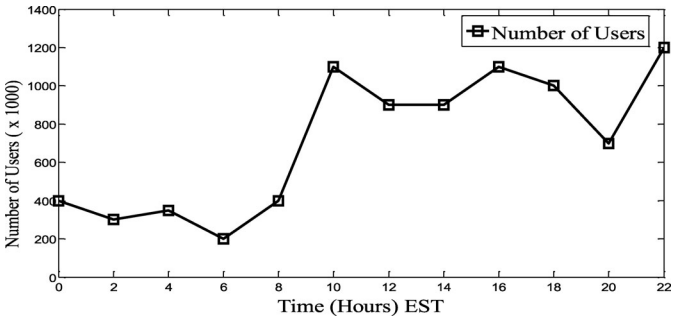

Fig. 4. Number of users versus time of the day.

switches (SwN_s) is calculated considering redundancy. Constraint (20) calculates the storage capacity needed in each cloud based on the number of replicated popularity groups.

B. The Content Delivery Cloud Model Results

The NSFNET network, depicted in Fig. 3, is considered as an example network to evaluate the power consumption of the cloud content delivery service over IP/WDM networks. It has 14 nodes and 21 bidirectional links [15]. In this paper the network is designed considering the peak traffic to determine the maximum resources needed. However we still run the model considering varying levels of traffic throughout the day to optimally replicate content as the traffic varies to obtain the minimum power consumption. One point in the operational cycle (the point that corresponds to the peak demand) is strict network design where the maximum resources needed are determined. At every other point the optimization is a form of adaptation where resources lower than the maximum are determined and used to meet the demand.

In our evaluation users are uniformly distributed among the NSFNET nodes and the total number of users in the network fluctuates throughout the day between 200 k and 1200 k as shown in Fig. 4. The values in Fig. 4 are estimated, based on the data in [18]. The number of users throughout the day considers the different time zones within the U.S. [18]. The reference time zone in Fig. 4 is EST. We use the model to evaluate the power consumption associated with the varying number of users at the different times of the day, with a 2 h granularity.

Table I gives the input parameters of the model. The router ports power consumption and number of wavelength per fiber is based on [15], note that the power consumption of the IP router port takes into account the power consumption of the different shared and dedicated modules in the IP router such as the switching matrix, power module and router processor. The eight-slot CRS-1 consumes about 8 kW and therefore the power consumption of each port is given as 1 kW.

TABLE I
INPUT DATA FOR THE MODELS

Power consumption of a router port (Prp)	1000 W [15]
Power consumption of transponder (Pt)	73 W [15]
Power consumption of an optical switch $(PO_i)\ \forall\ i \in N$	85 W [24]
Power consumption of EDFA (Pe)	8 W [25]
Power consumption of a Mux/Demux (Pmd)	16 W [26]
Number of wavelengths in a fiber (W)	16 [15]
Bit rate of each wavelength (B)	40 Gbps
Span distance between EDFAs (S)	80 km
Average client download rate $(Drate)$	5 Mbps [27]
Content server capacity (CS_C)	1.8 Gbps [28]
Content server energy per bit (CS_EPB)	211.1 W/Gbps [28]
Storage power consumption (S_PC)	4.9 kW [8]
Storage capacity (S_C)	75.6 TB [8]
Storage utilization (S_Utl)	50%
Storage and switching redundancy (Red)	2
Cloud switch power consumption (Sw_PC)	3.8 kW [8]
Cloud switch capacity (Sw_C)	320 Gbps [8]
Cloud router power consumption (R_PC)	5.1 kW [8]
Cloud router capacity (R_C)	660 Gbps [8]
Cloud power usage effectiveness (PUE_c)	2.5
IP/WDM power usage effectiveness (PUE_n)	1.5
Number of popularity groups (PGN)	50
Popularity group size $(\mathrm{PGS_p})$	0.756 TB
Time granularity (Δt)	2 Hours

The ITU grid defines 73 wavelengths at 100 GHz spacing or alternatively double this number approximately at 50 GHz channel spacing. In more recent studies [29] we have adopted lower router power per port, 440 W, based on Alcatel-Lucent designs. Also a larger (>16) number of wavelengths per fiber is possible at 100 or 50 GHz spacing, however with super channels and the introduction of 400 Gb/s and envisaged 1 Tb/s and possibly flexigrid developments, the number of channels may fall. Also note that a larger number of wavelengths per fiber will reduce the number of fibers and hence EDFAs, but the power consumption of the latter is small. To facilitate comparison with previous studies we have adopted the figures in [15], [18]. Note that data rate of router ports is the same as the wavelength rate (40 Gb/s).

The energy per bit (EPB) in the table is capacity based as we base it on the maximum power and maximum rate of the server. However note that the model always attempts to fully utilize servers, see constraint (17) and switches off unused servers to minimize power consumption. The combination of these two operational factors results in EPB based on capacity being a good representation. The 5 Mbps average download rate is based on the results of a survey conducted in the U.S. in 2011 [27].

A typical telecom office PUE is 1.5 [8]. Typical data centers PUE varies widely between 1.1 for large data centers able to

implement sophisticated water cooling [30] to small data centers with PUE as high as 3 [31]. We have adopted a PUE of 2.5 in this study for the small distributed clouds considered.

We divide the cloud content into 50 popularity groups which is a reasonable compromise between granularity and MILP model execution time. Note also that distributing the content into multiple locations does not increase the power consumption of the cloud, i.e., the power consumption of a single cloud with all the content is equal to the total power consumption of multiple distributed clouds storing the same content without replication. The reason is that each server has an idle power; however, it is either switched ON (and through packing is operated near maximum capacity) if needed or switched OFF. The cloud is made up of a large number of servers (200 for example, see Fig. 10). The cloud power consumption therefore increases with good granularity in steps equal to the power consumption of a single server. This leads to a staircase (with sloping stairs) profile of power versus load and with the very large number of steps, the profile is almost linear. This form of power management means that placing a given piece of content in different clouds amounts to the same power consumption approximately. We also assume similar type of equipment and PUE, hence we do not assume a fixed power component associated with placing a cloud in certain location. We also assume that underutilized storage units, switches, and routers) in the cloud are switched OFF or put in low power sleep mode. Note that savings are averaged over the 12 time points of the day (24 h). The metrics used are average savings over 24 h. These are the average network power saving, average cloud power saving and average total power saving.

Note that the variables specified above (with the exception of binary variables) take values that are dictated by the number of users, their data rates, content popularity and scenario considered.

We compare the following different delivery schemes:

1) *Single cloud:* Users are served by single cloud optimally located at node 6 as it yields the minimum average hop count. This scenario is obtained by setting the total number of clouds to 1, i.e., constraint (16) becomes

$$CN = \sum_{s \in N} Cloud_s = 1.$$

We consider two schemes in this model:

   a) *No power management (SNPM):* The cloud and the network are energy inefficient where different components are assumed to consume 80% of their maximum power consumption at idle state.
   b) *Power management (SPM):* The cloud and the network are energy efficient where underutilized components are powered off or put into deep sleep at off-peak periods.

2) *Max number of clouds with power management:* A cloud is located at each node in the network, i.e. the network contains 14 clouds. In this case the total number of clouds is set to 14, i.e., constraint (16) becomes

$$CN = \sum_{s \in N} Cloud_s = 14.$$

We consider three schemes in this model

   a) *Full replication (MFR):* Users at each node are served by a local cloud with a full copy of the content. This scheme is obtained by setting the number of popularity groups to 1, i.e., $|PG| = 1$.
   b) *No replication (MNR):* The content is distributed among all the 14 clouds without replication. This scheme is obtained by ensuring that the total number of replicas (δ_{sp}) does not exceed the original number of popularity groups $(|PG|)$:

$$\sum_{p \in PG} \sum_{s \in N} \delta_{sp} = |PG| .$$

   c) *Popularity based replication (MPR):* The model optimizes the number and locations of content replicas among all the 14 clouds based on content popularity.

3) *Optimal number of clouds with power management:* The number and location of clouds are optimized. We consider three schemes of this scenario:

   a) *Full replication (OFR):* Each cloud has a full copy of the content. This scheme is obtained by setting $|PG| = 1$.
   b) *No replication (ONR):* Content is distributed among the optimum clouds without replication. This scheme is obtained by setting

$$\sum_{p \in PG} \sum_{s \in N} \delta_{sp} = |PG| .$$

   c) *Popularity based replication (OPR):* The number and locations of content replicas are optimized based on content popularity.

We reduce the popularity group size to 20% of the size considered in the results of [14], i.e.,$\mathrm{PGS}_p = 0.756\mathrm{TB}$. A single download rate of 5 Mb/s is used for the two cases of file size. The two file sizes may for example correspond to two movies of different length and/or different resolution. At a constant user rate the time taken to download the larger file (before playing in this case) is longer. The smaller popularity group size represents a cloud scenario where music, for instance, is more popular than movies. Fig. 5(a) shows the total power consumption of the different schemes while Fig. 5(b) and (c) shows the network and cloud power consumptions, respectively.

The SNPM scheme, where all content is placed in a single cloud, results in the highest total power consumption as all the resources are switched on even at off-peak periods where idle power consumption contributes to 80% of the total power consumption. Therefore, the small variation in SNPM power consumption throughout the day is due to the 20% load induced power consumption. We use SNPM as our benchmark to calculate the power savings achieved by the other schemes. The power awareness of the SPM scheme saves 37% of both the network and cloud power consumption, compared to the SNPM scheme. The optimal location of the single cloud is selected based on network and cloud power consumption minimization using our MILP. Node 6 is the optimum location that minimizes power consumption based on MILP. It yields the minimum average hop count which is a good choice for power minimization given

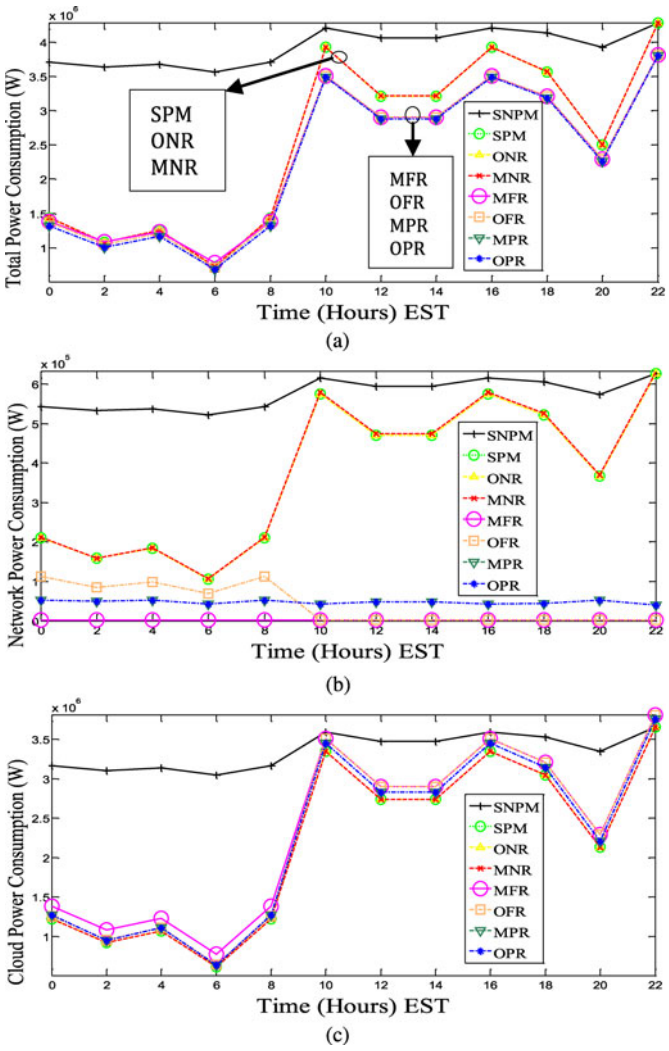

Fig. 5. (a) Total power consumption. (b) IP/WDM network power consumption. (c) Cloud power consumption.

that the traffic is uniformly distributed among the nodes. Note that this choice of the reference case is conservative. We could have chosen the reference case as a single cloud, but optimally placed to minimize delay or network CAPEX as is currently done, in which case the energy savings as a result of our work will be higher.

We also investigate the other extreme scenario represented by the MFR scheme where content is fully replicated into all possible locations (14 nodes of the NSFNET). At the network side (see Fig. 5(b)), the MFR saves 99.5% of the network power consumption as all requests generated in a node are served locally. The other 0.5% of the power consumption is associated with the optical switch shown in Fig. 2. However, having a cloud with full content at each node, especially with large popularity group size, significantly increases the cloud power consumption due to storage. Therefore the MFR scheme is not efficient if storage power consumption continues to dominate the power consumption of datacenters, for instance, we have shown in [14] that MFR cloud power consumption even exceeds the SNMP cloud power consumption at peak periods of the day, such as at 10:00, 16:00 and 22:00.

Despite their different approaches in optimizing content delivery, the MNR, ONR and SPM schemes have similar total power consumption. In terms of the cloud power consumption, the three schemes produce the lowest power consumption as content is not replicated and as mentioned above distributing the content into multiple locations does not increase the power consumption. In the following we discuss the network power consumption of the different schemes.

The ONR scheme finds the optimal number of clouds required to serve all users from a single copy of the content, where different popularity groups are migrated to different clouds based on their popularity. However, our results show that the ONR scheme selects to serve users using a single cloud located at node 6, i.e., ONR imitates the SPM scheme. This is because distributing the content into multiple clouds without replicating it results in higher network power consumption. For instance, if the ONR scheme decides to build two clouds one at node 1 (far left end of NSFNET) and the other at node 12 (far right end of NSFNET), then users at nodes located at the left end of the network will have to cross all the way to the right end of the network to download some of their content from cloud at node 12 as we do not allow this content to be replicated into the cloud in node 1, and vice versa for users located at nodes at the right end of the network asking for content only available at node 1. However, if all popularity groups are kept in node 6 which is the node that yields the minimum average hop count to different nodes in the network (this will be shown later in Section III-C), then the power consumed in the network to download content will be minimized, resulting in 37% saving in both total and network power consumption compared to SNPM, similar to SPM.

The main insight of the ONR scheme is that energy efficient content delivery prioritizes single cloud solutions over distributed solutions if content is not allowed to be replicated, the content has similar popularity at every node, and we are free to choose the number of clouds. This raises the question of how should we replicate content in a scenario of more than one cloud. The extreme of such a scenario is represented by the MNR where each node in the network contains a cloud. To create a cloud in a certain location, an object should be migrated into that location. However, as discussed earlier migrating content into multiple clouds without replication results in high power consumption in the network. The MNR scheme selects to place content in agreement with the insights of the ONR scheme, where most of the content is located in node 6, while migrating only the least popular objects, associated with the lowest network power consumption, into other nodes. This placement strategy meets the MNR scheme constraints and at the same time minimizes the network power consumption needed to access content. The MNR network power consumption slightly deviates from the ONR network power consumption (36.5% rather than 37%) as the least popular contents are not served from the cloud in node 6. The MNR scheme, however maintains the total power saving achieved by the ONR scheme, i.e., 37% compared to the SNPM scheme.

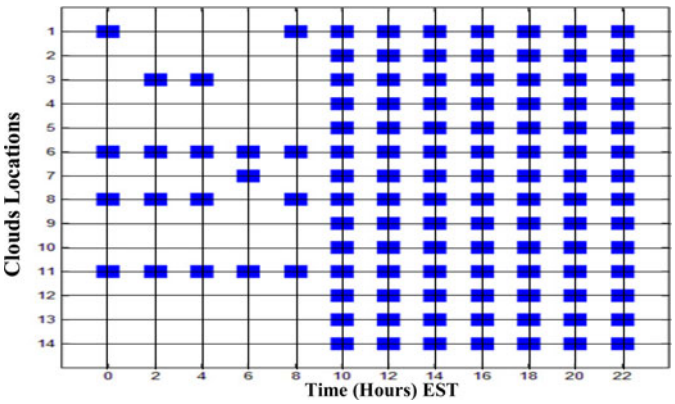

Fig. 6. OFR powered ON clouds at different times of the day.

ONR and MNR still follow the SPM scheme and centralize the content in one cloud at node 6, as observed in [14]. Therefore centralizing the content is the optimum solution if content is not allowed to be replicated regardless of content size.

As already observed, serving content locally by enforcing replication in all the 14 locations (MFR scheme) yields the highest cloud power consumption and the lowest network power consumption. In the OFR scheme we investigated the impact of removing the constraint on the number of clouds; so the model is free to choose the optimal number of clouds that have a full copy of the content. The OFR scheme manages to reduce the total number of clouds from 14 to 6 at off-peak periods of the day (between 00:00 and 08:00). These clouds are switched on and off according to the traffic variation as seen in Fig. 6. However, at peak periods (between 10:00 and 22:00) the OFR scheme converges to MPR scheme and builds clouds at all nodes as the power saved in the network is higher than the power lost in storage replication. OFR achieves significant power savings at the network side of 92% compared to the SNPM scheme, a saving higher than what is achieved by the SPM, ONR and MNR as the content is fully replicated to optimally located clouds. Eventually, the OFR scheme increases the total power saving compared to the SPM, ONR and MNR schemes, where the cloud in node 6 is mainly used, from 37% to 42.5% compared to SNPM, which is also slightly higher saving compared to MFR scheme due to deploying fewer clouds at off-peak periods.

From the discussion above it can be seen that the savings achieved by the OFR scheme are limited by the constraint on the content replication granularity where all content is replicated to a certain location if a cloud is created in that location. The question raised next is how much can the limits of power saving be pushed if the constraints on both the number of clouds and content replicated at each cloud are removed. This approach is implemented by the OPR scheme. The results in Fig. 5 indicate that the OPR scheme provides solutions with the lowest total power consumption.

The OPR, MPR and OFR schemes in this case have similar power consumption to the MFR scheme (43% total power savings compared to SNPM) which implies that the three schemes tend to replicate the majority of the content in all the 14 clouds

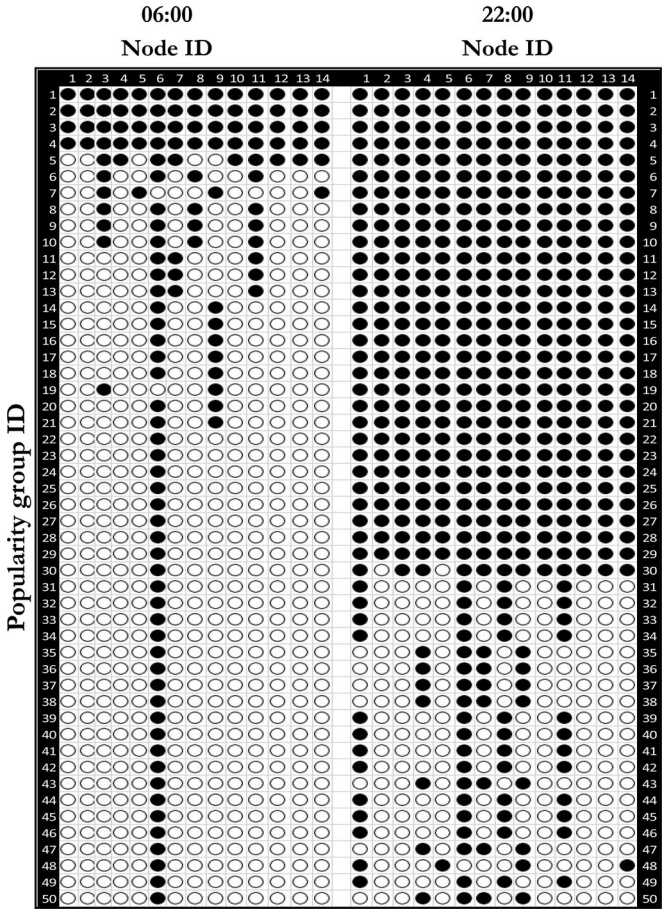

Fig. 7. Popularity groups placement under the OPR scheme at 06:00 and 22:00.

most of the time due to the small storage power cost as shown in Fig. 5(c).

Fig. 5(b) shows that the network power savings for the different models follow a similar trend to that in Fig. 5(a). At the network side, the OPR, MPR, OFR and MFR schemes save 92–99.5% of the network power consumption compared to the SNPM scheme while the SPM, ONR and MNR schemes save 37% of the network power consumption. Despite their similar average network power saving, OPR and OFR have different behaviour in saving power at the network side. OPR maintains the network power at the lowest level at all times by powering on the optimum number of clouds with the optimum content. However, OFR is less flexible as all the powered on clouds have to have a full copy of the content. Therefore OFR powers on only four clouds at low load and consumes higher network power while behaving as MFR and consuming power only in optical switches at high loads (10:00 to 22:00) by powering on 14 clouds (see Fig. 6), resulting in a daily average network power consumption similar to the OPR scheme.

Fig. 7 shows how the OPR scheme replicates the content at two times of the day, 06:00 (left half of Fig. 7) and 22:00 (right half of Fig. 7), which correspond to the lowest and highest load demands, respectively (as shown in Fig. 4). As the popularity of the content increases (popularity group 1 is the most popular),

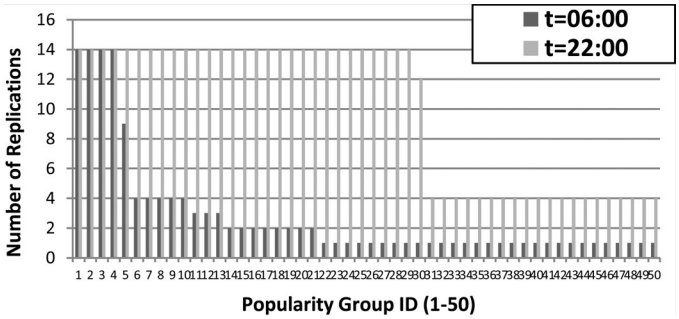

Fig. 8. Total number of replications per popularity group under the OPR scheme.

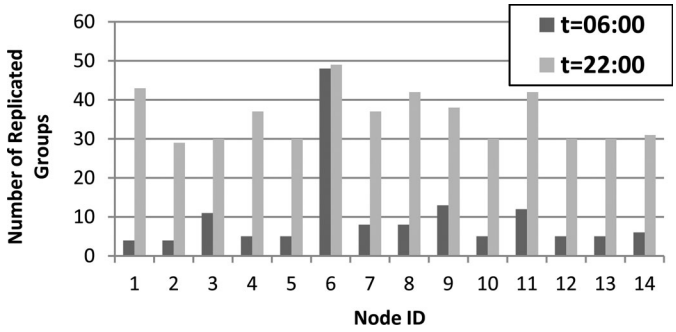

Fig. 9. Number of popularity groups replicated at each cloud under the OPR scheme.

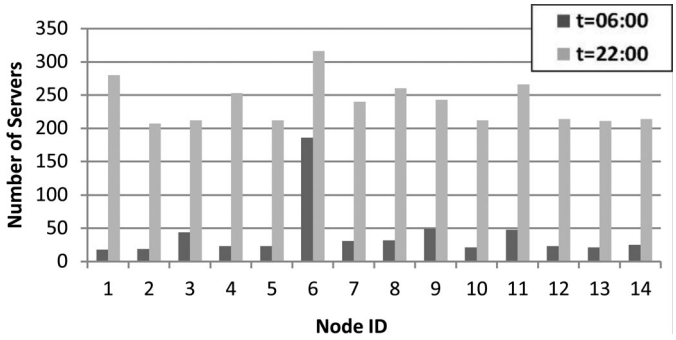

Fig. 10. Number of powered on servers in each cloud under the OPR scheme.

TABLE II
THE POWER SAVINGS GAINED BY THE DIFFERENT CONTENT DELIVERY CLOUDS SCENARIOS COMPARED TO THE SNPM SCENARIO

Scenario	Total Saving	Network Saving
OPR	43%	92%
MPR	43%	92%
OFR	42.5%	92%
SPM	37%	37%
ONR	37%	37%
MNR	37%	36.5%
MFR	42%	99.5%

it will be replicated (black circle) to more clouds as this will result in serving more requests locally and therefore reducing the network power consumption. At 06:00, most of the content is kept in node 6, while at 22:00, more popularity groups are replicated into more clouds as under higher loads the power consumption of the cloud is compensated by higher savings in the network side. Compared to the results of [14], Fig. 7 shows that OPR tends to replicate more content to more clouds due to lower power consumption of storage, especially that we fixed the average download rate at 5 Mbps. As content is equally popular among all users and users are uniformly distributed among nodes, OPR will replicate the most popular content into all the 14 nodes most of the day. Therefore OPR has similar power efficiency to the MPR where content is optimally replicated among all the 14 clouds based on its popularity.

Fig. 8 shows the total number of replicas per popularity group for the OPR scheme, obtained by summing the black circles for each popularity group in Fig. 7. At low demand, the number of replications follows a Zipf distribution. However, at high demand (at 22:00), the scheme does not follow a Zipf distribution in replicating content as the high demand and the low power cost of storage allow the majority of the popularity groups to be replicated into all the nodes.

Fig. 9 shows the total number of popularity groups replicated at each built cloud, obtained by summing the black circles in Fig. 7 for each cloud; i.e., Fig. 9 reflects the relative cloud size built at each node. At low demand periods (06:00) most of the content is replicated to the cloud in node 6 only, while at peak period (time 22:00) fewer popularity groups are replicated to the cloud in node 6 as the high network power savings at peak period justify replicating the content to more clouds instead of having a single copy in a centralized cloud. Due to the low power cost of storage, content is replicated to more clouds at both high and low loads compared to the larger popularity group results in [14].

Fig. 10 shows the number of servers powered on at each cloud at $t = 06{:}00$ and $t = 22{:}00$. The number of servers powered on at each cloud at a certain time is proportional to the size of content replicated to the cloud given in Fig. 9. Fewer servers are powered on at node 6 at both times compared to the results of [14] as more requests are served from other clouds due to low storage cost. Note that the number of switches and routers at each cloud follows a similar trend to that of the number of servers.

From the results above it is observed that OPR is always the best scenario for content delivery. OPR converges to a single cloud scenario for content of larger size at low demand periods while it fully replicates contents at all network locations for content of smaller size at high demand periods. OPR can be realized either by replicating popular content to the optimized locations and continuously replacing it as the popularity changes throughout the day; or by replicating all the content to all clouds and only switching on the hard disks storing the content with the highest popularity, selected by the model at each location. The latter storage management approach saves power in the network side; however, it needs knowledge of content popularity to assign separate hard disks to different popularity groups.

Table II summarizes the power savings achieved by the different cloud Content delivery approaches considered compared to the SNPM model.

C. DEER-CD: Energy Efficient Content Delivery Heuristic for the Cloud

As OPR is the most energy efficient scheme among the different schemes investigated, in this section we build a heuristic (DEER-CD) to mimic the OPR model behaviour in real time. We need to build a mechanism to allow the cloud management to react to the changing network load by replicating the proper content to the optimum locations rather than deciding an average replication scheme that might not fit well with different load patterns throughout the day. The DEER-CD heuristic involves two phases of operation:

1) *Offline phase:* Each node in the network is assigned a weight based on the average number of hops between the node and the other nodes and the traffic generated by the node, i.e., the number of users in the node and their download rate. The weight of node *s,* NW_s, is given as

$$NW_s = \sum_{d \in N} U_d \cdot Drate \cdot H_{sd} \tag{21}$$

where U_d is the number of users in node d and H_{sd} is the minimum number of hops between node pair (*s, d*).

As the network power consumption is proportional to the amount of content transiting between nodes and the number of hops travelled by the content, nodes of lower weight are the optimum candidates to host a cloud. Equation (21) can be pre-calculated as it relies on information that rarely changes throughout the day such as the physical topology, average users population and download rate. We construct a sorted list of nodes from lowest to highest weight and use this list to make cloud placement decisions. In the absence of our sorted list approach, an exhaustive search is needed where a *PG* can be allocated to a single node leading to $|N|$ (number of nodes) combinations being evaluated for network energy efficiency. In addition the *PG* can be allocated to two nodes leading to the evaluation of $|N| \cdot (|N| - 1)$ combinations, and therefore in total this exhaustive search, which we totally avoid, requires $\sum_{i=1}^{|N|} \frac{|N|!}{(|N|-i)!}$ placement combinations to be assessed. Our sorted list reduces this search to the evaluation of $|N|$ combinations only. For NSFNET with $|N| = 14$ this is a complexity reduction by a factor of 1.6×10^{10}. Furthermore using a 2.4 GHz Intel Core i5 PC with 4 GB Memory, the heuristic took 1.5 min to evaluate the DEER-CD results.

In our scenario the users are uniformly distributed over all nodes and they all have the same average download rate of 5 Mbps. Therefore the nodes' ranking is mainly based on the average number of hops. The list of ordered nodes based on C_s from the lowest to the highest is as follows:

LIST $= \{6, 5, 4, 3, 7, 9, 13, 10, 11, 12, 14, 1, 8, 2\}$.

To place a given popularity group in one cloud, node 6 is the best choice as it has the minimum weight. If the model decides to have two replicas of the same popularity group, then they will be located at nodes $\{6, 5\}$. Higher numbers of replicas are located similarly by progressing down the list and replicating content in a larger ordered subset of the set above. We call each subset of the list a placement (J). Therefore, DEER-CD will only have 14 different placements for each popularity group to choose from,

Inputs:		LIST= {6, 5, 4, 3, 7, 9, 13, 10, 11, 12, 14, 1, 8, 2}, $PG = \{1..PGN\}$
Outputs:		Optimal Placement (J'), Total Power Consumption (TPC)
1.		**For** each popularity group $i \in PG$ **Do**
2.		**For** each placement $J \subseteq$ LIST **Do**
3.		**For** each node $d \in N$ **Do**
4.		**For** each cloud location candidate $s \in J$ **Do**
5.	(a)	**Add**$\{cost_{sd}\} = MinHop$ (s, d)
6.	(b)	**End For**
7.	(c)	**Get** s where: $\text{cost}_{sd} = Min\{cost_{sd}\}$
8.	(d)	$\text{L}_{iJsd} = D_{id}$
9.		**End For**
10.		$\text{NPC}_{iJ} = MultiHopHeuristic\{N, Nm_i, L_{iJsd}\}$
11.		$\text{CPC}_{iJ} = PUE\_c \cdot (SrvPC\_CD + LANPC\_CD + StPC\_CD)$
12.		$\text{TPC}_{iJ} = \text{NPC}_{iJ} + \text{CPC}_{iJ}$
13.		**End For**
14.		$\text{TPC}_i = \text{Min}\{\text{TPC}_{iJ}\}$
15		$J' = J$
16.		**End For**
17.		Calculate$\text{TPC} = \sum_{i \in PG} \text{TPC}_i$

Fig. 11. The DEER-CD heuristic pseudo-code.

which dramatically minimizes the number of iterations needed to decide the optimal placement for each popularity group.

2) *Online phase:* In this phase, the list generated from the offline phase is used to decide the placement of each popularity group. Fig. 11 shows the pseudo code of the heuristic.

For each popularity group, the heuristic calculates the total power consumption TPC_{iJ} associated with placing each popularity group $i \in PG$ in each placement, $J \subseteq$ LIST. The total power consumption is composed of network power consumption NPC_{iJ} and the cloud power consumption CPC_{iJ}, at each placement J. Each cloud location candidate in the placement, $s \in J$(loop(a)), is assigned to serve nodes according to the minimum hop count (loop (b)). (L_{iJsd}) is the traffic matrix generated by placing popularity group i in the set of nodes s that are specified by the associated placement J, d denotes the set of other nodes in the network where users are requesting files in popularity group i. We use multi hop non-bypass heuristic developed in [15] to route the traffic between nodes s and d and calculate the network power consumption that is induced due to cross traffic between the nodes associated with each placement (14 possible placements) for each popularity group (loop (c)). The total power consumption is calculated and the placement associated with the lowest TPC_i among the 14 possible placements is selected to replicate the popularity group (loop (d)).

The DEER-CD heuristic is able to build a core network that includes clouds and it is able to handle the resultant traffic to minimize power consumption. It is able to use the number of users in each node, the network hop counts, user data rates and content popularity distribution to specify the location of clouds and their capability in terms of servers, switches, storage and routers. It is able subsequently to route the resultant traffic using multi-hop non-bypass to minimize power consumption. As such

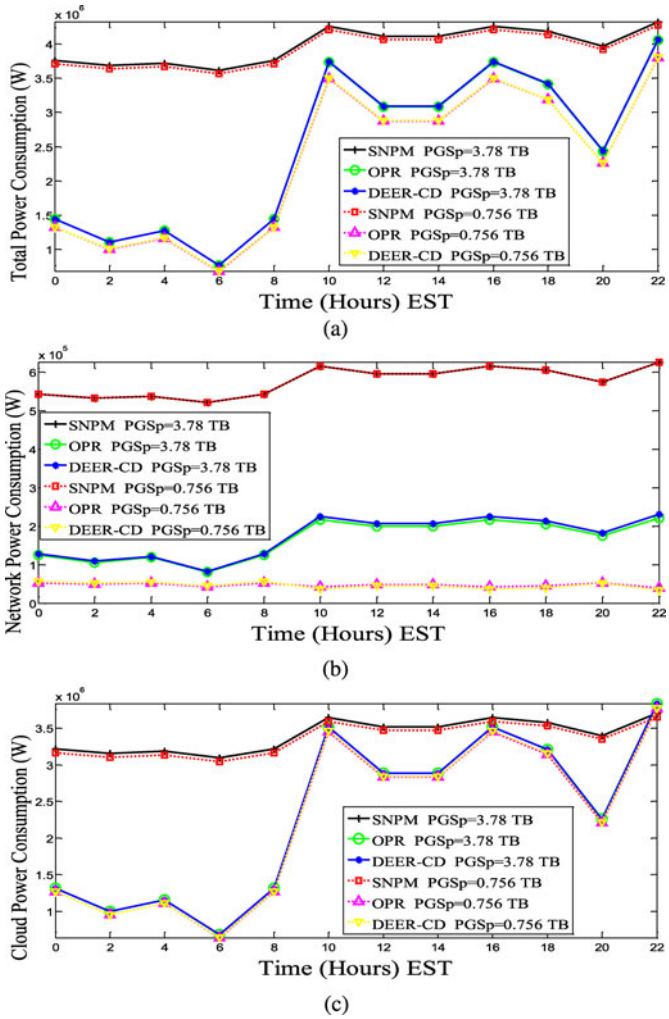

Fig. 12. (a) Total power consumption of the DEER-CD heuristic. (b) IP/WDM network power consumption of the DEER-CD heuristic. (c) Cloud power consumption of the DEER-CD heuristic.

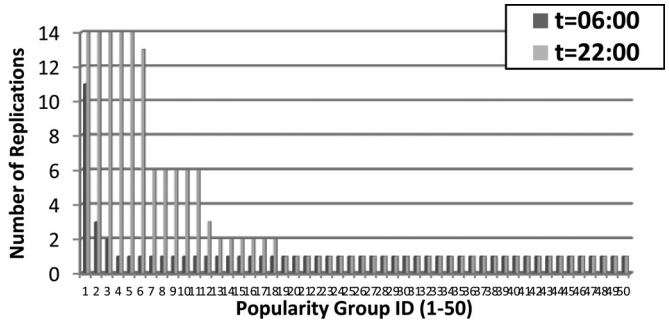

Fig. 13. Total number of replications per popularity group for DEER-CD ($\mathbf{PGS_p} = 3.78\mathbf{TB}$).

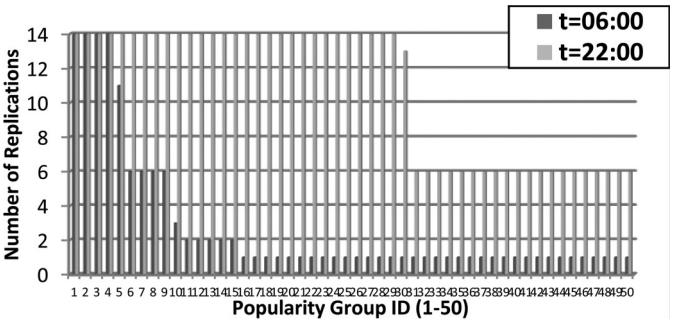

Fig. 14. Total number of replications per popularity group for DEER-CD($\mathbf{PGS_p} = 0.756\mathbf{TB}$).

TABLE III
OPTIMIZATION GAPS BETWEEN THE DEER-CD HEURISTIC AND MILP

Metric	Gap (%)			PG Size
	06:00	22:00	Avrg.	
Total Power	0.26%	0.10%	0.21%	0.756 TB
	0.22%	0.26%	0.25%	3.78 TB
Network Power	4.06%	-18.30%	-3.72%	0.756 TB
	2.06%	4.63%	3.32%	3.78 TB
Cloud Power	$1.56*10^{-5}$%	0.26%	0.15%	0.756 TB
	$1.46*10^{-5}$%	$-2.61*10^{-5}$%	0.01%	3.78 TB

it is able to carry out a complete network design and a complete cloud design which is different from dynamic operation over an existing cloud and an existing network.

Note that our heuristic in its current form does not consider the capacity constraints explicitly (such as the number of installed fibres and other network elements) as we consider designing the network for peak traffic (t = 22:00) first and running the heuristic at times where the traffic is less than the peak traffic (see Fig. 4). The heuristic can however be easily updated to work in a capacitated network.

Fig. 12(a) shows the total power consumption for the DEER-CD heuristic, while Fig. 12(b) and (c) show the IP/WDM network and cloud power consumptions, respectively, also under the DEER-CD heuristic. For the larger size popularity group scenario, the OPR model and the DEER-CD heuristic achieve comparable network power savings of 72% and 70%, respectively, compared to SNPM scheme. Also the cloud and total power savings achieved by the OPR model are maintained at 34% and 40%, respectively. This is due to the almost identical popularity groups' placement by the model and heuristic. A Similar observation is noticed for the smaller popularity group where the DEER-CD heuristic maintains the network, cloud and total power savings of 92%, 35% and 43%, respectively, achieved by the OPR model.

Figs. 13 and 14 show the number of replications per popularity group for the two popularity group sizes. Similar to the OPR model, the heuristic displays a Zipf like behaviour in replicating popularity groups for the larger popularity group size (see Fig. 13). Similar distributions to those of the model in Fig. 8 are exhibited by the heuristic for the smaller size where at low demand the content placement follows a Zipf distribution while at higher demand it follows a simple binary distribution. Such behaviour can be harnessed to simplify the content placement heuristic in networks characterized by long periods of peak traffic. Comparing the distribution in Fig. 14 at (t = 22:00) to Fig. 8 shows that the heuristic creates more replications for popularity groups of lower popularity which results in higher cloud power consumption and lower network power consumption compared to the OPR model as shown in Table III which reports the

optimization gaps between the DEER-CD heuristic and the OPR model for both larger and smaller popularity group sizes at low and high traffic.

IV. Energy Efficient Storage as a Service

StaaS can be viewed as a special case of the content delivery service where only the owner or a very limited number of authorized users have the right to access the stored content. Dropbox, Google Drive, Skydrive, iCloud, and Box are examples of cloud based storage. In energy efficient StaaS, all content is stored in one or more central locations and dynamically migrated to locations in proximity of its owners to minimize the network power consumption. The content can be migrated, content migration, however, consumes power at the IP/WDM network as well as in the servers and internal LAN of the clouds. Therefore, StaaS should achieve a trade-off between serving content owners directly from the central cloud/clouds and building clouds near to content owners. Upon registration for StaaS, users are granted a certain size of free storage (Quota). DropBox [32], for instance, grants its users 2 GB. Different users might have different levels of utilization of their StaaS quota as well as different files access frequency. Large file access frequency has two meanings: either one user highly accesses it or many authorized users have low/moderate access frequency to the same file.

A. StaaS MILP Model

We extend the model in Section III-A to capture the distinct features of StaaS. As only the owner or a very limited number of authorized users have the right to access the stored content, the concepts of popularity and replication do not apply to StaaS.

In addition to the parameters and variable defined in Section III-B, we define the following:

Parameters:

$Freq$	Average file download frequency per hour.
$Dsize$	Average file size in Gb.
$Rate$	Average user rate per second, where $Rate = 2 \cdot Freq \cdot Dsize/3600$.
ND_d	Node d total traffic demand, $ND_d = \sum_{i \in U_d} Rate$.
$Quota$	Users storage quota in Gb.
SU_i	Storage Quota utilization of user *i*.
α	1 or 2.

Variables (All are Nonnegative Real Numbers)

IT_{cd}	Traffic between the central cloud and cloud in node d due to content migration.
CU_{sd}	Traffic from the cloud in node s to users in node *d*.
π_{sd}	$\pi_{sd} = 1$ if cloud s serves users in node d, $\pi_{sd} = 0$ otherwise.
L_{sd}	Total traffic between node pair (s, d).

Note that the average user rate, Rate, substitutes Drate in Section III-A and is calculated by dividing total amount of data sent and received which (are measured in Gb, and equal to$2 \cdot \mathrm{Freq} \cdot \mathrm{Dsize}$) over one hour by number of seconds in one hour as users need to download the full file before editing and need to upload the full file after finishing (or intermediately). Waiting is not desirable and the file access frequency dictates the data rate. The factor of 2 is introduced to represent the fact that users usually re-upload their files back to the cloud after downloading and processing them.

The objective of the model in Section III-A applies to the StaaS model, as well as constraints (1)–(7), (17)–(19). The following additional constraints are introduced:

1) Clouds to users traffic

$$\sum_{s \in N} CU_{sd} = ND_d \quad \forall d \in N \tag{22}$$

$$M \cdot CU_{sd} \geq \pi_{sd} \quad \forall s, d \in N \tag{23}$$

$$CU_{sd} \leq M \cdot \pi_{sd} \quad \forall s, d \in N \tag{24}$$

$$CU_{cd} = ND_d \cdot \left(1 - \sum_{b \in N : b \neq c} Cloud_b\right). \quad \forall d \in N : d \neq c \tag{25}$$

Constraint (22) ensures that the traffic demand of all users in each node is satisfied. Constraints (23) and (24) decide whether a cloud serves users in node d or not, where M is a large enough number, with units of 1/Gb/s and Gb/s in (23) and (24), respectively, to ensure that $\pi_{sd} = 1$ when CU_{sd} is greater than zero. Constraint (25) sets the traffic between the central cloud and users in other nodes to 0 if those users have a nearby cloud to download their content from.

2) Clouds locations

$$M. \sum_{d \in N} CU_{sd} \geq Cloud_s \quad \forall s \in N \tag{26}$$

$$\sum_{d \in N} CU_{sd} \leq M \cdot Cloud_s. \quad \forall s \in N \tag{27}$$

Constraints (26) and (27) build a cloud in location s if that location is selected to serve the requests of users of at least one node d, where M is a large enough number, with units of 1/Gb/s and Gb/s in (26) and (27), respectively, to ensure that $\mathrm{Cloud}_s = 1$ when $\sum_{d \in N} \mathrm{CU}_{sd}$ is greater than zero

3) Clouds storage capacity

$$StrC_s = \sum_{d \in N} \sum_{i \in U_d} \pi_{sd} \cdot Quota \cdot SU_i \cdot Red. \quad \forall s \in N \tag{28}$$

Constraint (28) calculates the cloud storage capacity based on the number of users served by the cloud, their storage quota and utilization, taking redundancy into account.

4) Inter clouds traffic:

$$IT_{cd} = \sum_{i \in U_d} Cloud_d \cdot Quota \cdot SU_i / (3600 \cdot \Delta t).$$
$$\forall d \in N : d \neq c \tag{29}$$

Constraint (29) calculates the content migration traffic between the central cloud and local clouds. The factor of Δt in the denominator scales the power consumption down to be consistent with our evaluation period of Δt hours.

5) Total traffic between nodes and clouds upload capacity

$$L_{sd} = CU_{sd} + IT_{sd}$$
$$\forall s, d \in N \tag{30}$$
$$Cup_s = \sum_{d \in N} (L_{sd} + \alpha \cdot IT_{sd}).$$
$$\forall s \in N \tag{31}$$

Constraint (30) calculates the total traffic between node pair (*s, d*) as the summation of the inter cloud traffic and clouds to users traffic. Constraint (30) substitutes constraint (10) in calculating network traffic. Constraint (31) calculates clouds upload capacity which includes the clouds to users traffic and the clouds inter traffic when $s = c$. The factor α is set to 2 if we are interested in total power consumption, as the receiving local cloud will consume similar power in its servers and internal LAN to the central sending cloud during migration. However, it is set to 1 if we are only interested in the actual clouds upload capacity and capability calculated by constraints (17)–(19). In this section we set $\alpha = 2$ as we are interested in power consumption.

B. StaaS Model Results

The NSFNET network, the users' distribution and input parameters discussed in Section III-C are also considered to evaluate the StaaS model. Node 6 is optimally selected based on the insights of Section III to host one central cloud.

We analyze 1200 k users uniformly distributed among network nodes which correspond to time 22:00 in Fig. 4. Power consumption calculation is averaged over the range of access frequencies considered (10 to 130 downloads per hour).

We analyse three different schemes to implement StaaS:

1) *Single cloud:* Users are served by the central cloud only.
2) *Optimal clouds:* Users at each node are served either from the central cloud or from a local cloud by migrating content from the central cloud.
3) *Max clouds:* Users at each node are served by a local cloud.

We evaluate the different schemes considering two file sizes of 22.5 MB and 45 MB and a user storage quota of 2 GB.

Note that the file sizes reflect content of high resolution images or videos. Files of smaller sizes will result in low network traffic that will not justify replicating content into local clouds.

Users' storage utilization SU_i is uniformly distributed between 0.1 and 1.

Fig. 15(a) shows the total power consumption versus the content access frequency while Fig. 15(b) and (c) decompose it into the IP/WDM network power consumption and cloud power consumption, respectively. At lower access frequencies, the optimal clouds scheme selects to serve all users from the central cloud. At higher access frequencies, however, the impact of the file size becomes more relevant. For the larger file size (45 MB) scenario local clouds are built whenever the access frequency is equal to or higher than 50 downloads per hour. On the other hand for the smaller file size of 22.5 MB, users are served from the central cloud up to access frequencies as high as 90 downloads per hour for this smaller file size (22.5 MB) scenario. This is because a larger file size results in higher traffic and consequently larger reduction in traffic between the central cloud and users when serving the requests locally. Therefore it will compensate for the power consumption of content migration. To eliminate the impact of content replication on storage power consumption, this scheme requires switching off the central clouds storage that stores the migrated content after migration. Apple iCloud is an example of a cloud implementation that typically hosts large files representing images and videos. Migration to local clouds is energy efficient in this case as opposed to clouds that typically host small text documents.

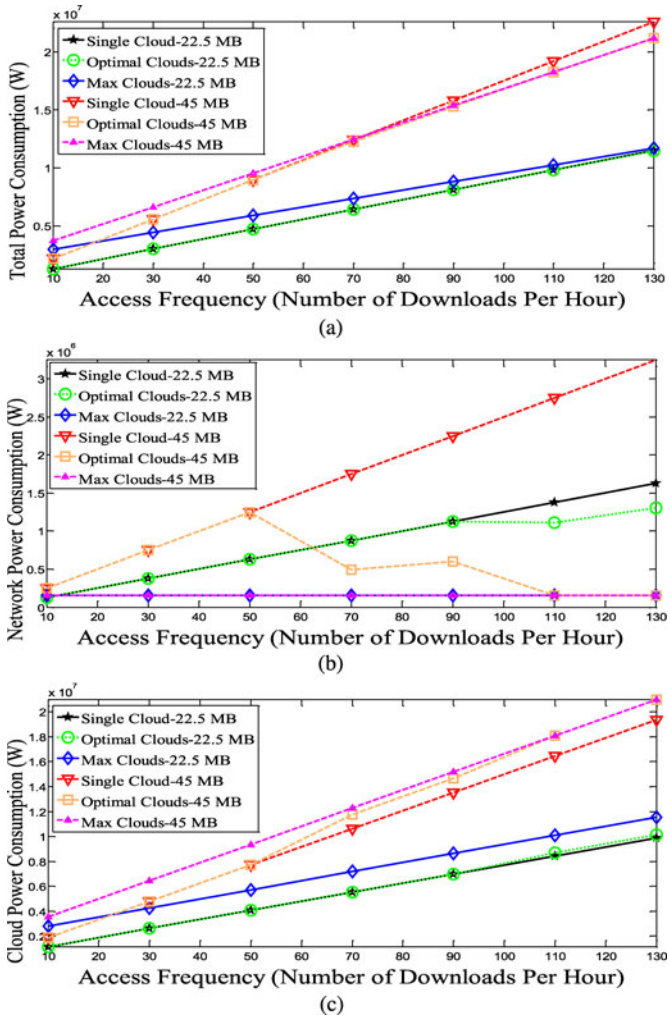

Fig. 15. (a) Total power consumption of StaaS. (b) IP/WDM network power consumption of StaaS. (c) Cloud power consumption of StaaS.

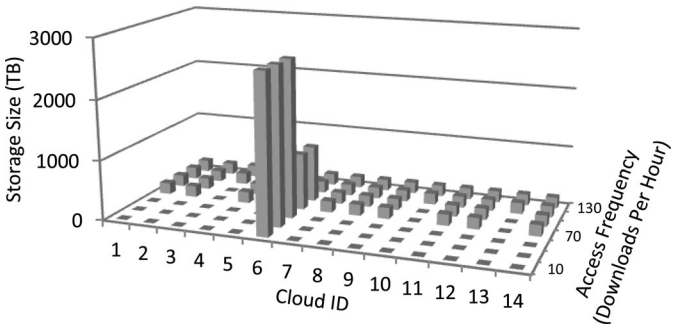

Fig. 16. Power-on clouds storage size versus content access frequency for the StaaS optimal scheme (45 MB file size).

On the network side, the Max Clouds scheme has the lowest network power consumption, saving 67% and 83% of the network power consumption compared to the Single Cloud scheme considering the 22.5 MB and the 45 MB average file size, respectively. This is because all users are served locally and the IP/WDM network power consumption is only due to content migration from the central cloud. However, this saving is at the cost of high power consumption inside the clouds to migrate content, resulting in an increase of 28% and 19% in the cloud power consumption for the 22.5 and 45 MB cases, respectively, as shown in Fig. 15(c).

For the 22.5 MB file size scenario, limited total and network power savings are obtained by the optimal clouds scheme compared to the single cloud scheme (0.1% total and 5% network power saving) as shown in Fig. 15(b). On the other hand, for the 45 MB file size scenario, more total and network power savings are obtained by the Optimal Cloud scheme compared to the Single Cloud scheme (2% total and 48% network power saving) as shown in Fig. 15(b).

Fig. 16 shows the variation in the number and size of clouds with the content access frequency for the 45 MB file size. As the content access frequency increases, migrating content from the central cloud to local clouds becomes more energy efficient compared to delivering content directly from the central cloud and therefore more clouds are needed to serve users locally. Note that the storage size in Fig. 16 represents the powered-on storage. This results in decreased storage size of the central cloud with higher content access frequency as shown in Fig. 16. Fig. 16 also shows that other clouds have almost similar storage size, at high access rates, due to the uniform distribution of users in the network.

V. VM Placement Optimization

Machine virtualization provides an economical solution to efficiently utilize the physical resources, opening the door for energy efficient dynamic infrastructure management as highlighted by many research efforts in this field. The authors in [33] studied the balance between server energy consumption and network energy consumption to present an energy aware joint VM placement inside datacenters. The authors in [34] proposed the use of multiple copies of active VMs to reduce the resource requirement for each copy of the VM by distributing the incoming requests among them to increase the energy efficiency of the consolidation and VM placement algorithm. They considered heterogeneous servers in the system and used a two dimensional model that considers both computational and memory bandwidth constraints. The authors in [35] proposed a MILP formulation that virtualizes the backbone topology and places the VMs in several cloud hosting datacenters interconnected over an optical network with the objective of minimizing power consumption.

In this section we optimize the placement of VMs in IP/WDM networks to minimize the total energy consumption. We consider different VM distribution schemes. In our analysis, a VM is defined as a logical entity created in response to a service request by one or more users sharing that VM. A user request is defined by two dimensions: (i) the CPU utilization (normalised workload) of the VM and (ii) the traffic demand between the VM and its user. In this section we use the terms CPU utilization and normalised workload interchangeably.

A. Cloud VM Placement MILP Model

We develop a MILP model to optimize the number, and location of clouds and optimize the placement of VMs within the clouds as demands vary throughout the day to minimize the network and clouds power consumption. The model considers three VM placement schemes:

1) *VM replication:* More than one copy of each VM is allowed in the network.
2) *VM migration:* Only one copy of each VM is allowed in the network. We assume that the internal LAN capacity inside datacenters is always sufficient to support VM migration.
3) *VM slicing:* The incoming requests are distributed among different copies of the same VM to serve a smaller number of users as proposed in [34]. We call each copy a slice as it has less CPU requirements. As VMs with small CPU share might threaten the SLA, we enforce a limit on the minimum size of the VM CPU utilization. Unlike [34] where CPU and memory bandwidth are considered, we consider the CPU and traffic dimensions of the problem where each slice is placed in a different cloud rather than in different server inside the same cloud.

In addition to the variables in Section III-A, we define the following variables and parameters:

Parameters:

VM	Set of virtual machines.
U_v	Set of users requesting VM v.
NVM	Total number of virtual machines.
$P_{\max}$	Maximum power consumption of a server.
$W_{\max}$	Maximum normalised workload of a server.
∇	Server energy per bit, $\nabla = P_{\max}/W_{\max}$.
W_v	Total normalised workload of VMv.
D_{dv}	Traffic demand from VM v to node $d, D_{dv} = \sum_{i \in U_d : i \in U_v} Drate.$
$MinW$	Minimum allowed normalised workload per VM.

Variables (All are Nonnegative Real Numbers)

$L_{s\mathrm{dv}}$	Traffic demand from VM v in cloud s to node *d.*
δ_{sv}	$\delta_{sv} = 1$ if cloud s hosts a copy of VM *v,* otherwise $\delta_{sv} = 1$.

CW_s Total normalised workload of Cloud *s*.
W_{sv} Normalised workload of the slice of VM v in node *s*.
PSN_s Number of processing servers in cloud *s*.

The power consumption of the cloud considering the machine virtualization scenario is composed of

1) The power consumption of servers $(SrvPC\_VM)$

$$\sum_{s\in N} \nabla \cdot CW_s.$$

2) The power consumption of switches and routers $(LANPC\_VM)$

$$\sum_{s\in N} Cup_s \cdot (Sw\_EPB \cdot Red + R\_EPB).$$

Note that we do not include the storage power consumption in our models. Although the server power consumption is a function of the idle power, maximum power and CPU utilization [36], for large number of servers, taking only $\nabla = P_{\max}/W_{\max}$ as the server EBP to calculate its power consumption yields very close approximation as the difference will be only in last powered on server. Note that a cloud is composed of a large number of servers and through "packing," each server in our case is either as close to fully utilized as possible or is off. In such a case "idle power plus linear increase in power with load" is equivalent to "linear increase in power with load" as both servers are either operated near the peak or are off and the peak powers are identical. For the overall cloud either a single server (or more generally a very small minority of servers) may be partially loaded. Therefore for a cloud made up of a large number of highly used servers (unused servers are turned off), the power consumption increases in proportion to load approximately. Note that if servers are not fully packed, this approximation becomes less accurate. This warrants further investigation, however our approach is followed in the literature [28]. As for the content delivery model, we also assume here that other storage and network elements have similar power management as servers.

The model is defined as follows:

Objective: Minimize

$$\begin{aligned} PUE\_n \cdot \Bigg(& \sum_{i\in N} Prp \cdot Q_i + Prp \cdot \sum_{m\in N} \sum_{n\in N m_m} W_{mn} \\ & + \sum_{m\in N} \sum_{n\in N m_m} Pt \cdot W_{mn} \\ & + \sum_{m\in N} \sum_{n\in N m_m} Pe \cdot A_{mn} \cdot F_{mn} \\ & + \sum_{i\in N} PO_i + \sum_{m\in N} \sum_{n\in N m_m} Pmd \cdot F_{mn} \Bigg) \\ & + PUE\_c \cdot \Bigg(\sum_{s\in N} \nabla \cdot CW_s + \sum_{s\in N} Cup_s \cdot (Sw\_EPB \cdot Red \\ & + R\_EPB) \Bigg). \end{aligned} \tag{32}$$

Subject to

1) VMs demand

$$\sum_{s\in N} L_{sdv} = D_{dv} \quad \forall d \in N \ \ \forall v \in VM. \tag{33}$$

Constraints (33) ensures that the requests of users in all nodes are satisfied by the VMs placed in the network.

2) VMs locations

$$M \cdot \sum_{d\in N} L_{sdv} \geq \delta_{sv} \quad \forall s \in N \ \ \forall v \in VM \tag{34}$$

$$\sum_{d\in N} L_{sdv} \leq M \cdot \delta_{sv} \quad \forall s \in N \ \ \forall v \in VM. \tag{35}$$

Constraints (34) and (35) replicate VM v to cloud s if cloud s is selected to serve requests for v where M is a large enough number, with units of 1/Gb/s and Gb/s in (34) and (35), respectively, to ensure that $\delta_{sv} = 1$ when $\sum_{d\in N} L_{sdv}$ is greater than zero.

3) Clouds locations

$$\sum_{v\in \mathrm{VM}} \delta_{sv} \geq Cloud_s \quad \forall s \in N \tag{36}$$

$$\sum_{v\in \mathrm{VM}} \delta_{sv} \leq M \cdot Cloud_s. \quad \forall s \in N \tag{37}$$

Constraints (36) and (37) build a cloud in location s if the location is selected to host one or more VMs where M is a large enough unitless number to ensure that $\mathrm{Cloud}_s = 1$ when $\sum_{v\in \mathrm{VM}} \delta_{sv}$ is greater than zero.

4) Total cloud normalised workload for replication and migration schemes

$$CW_s = \sum_{v\in VM} \delta_{sv} \cdot W_v. \quad \forall s \in N \tag{38}$$

Constraint (38) calculates the total normalised workload of each cloud by summing its individual VMs normalised workloads.

5) VM migration constraint

$$\sum_{s\in N} \delta_{sv} = 1 \quad \forall v \in VM. \tag{39}$$

Constraint (39) is used to model the VM migration scenario where only one copy of each VM is allowed.

6) VM Slicing constraints

$$\sum_{s \in N} W_{sv} = W_v$$
$$\forall v \in VM \tag{40}$$
$$W_{sv} \geq MinW \cdot \delta_{sv}$$
$$\forall s \in N \ \forall v \in VM \tag{41}$$
$$W_{sv} \leq M \cdot \delta_{sv}$$
$$\forall s \in N \ \forall v \in VM \tag{42}$$
$$CW_s = \sum_{v \in VM} W_{sv}$$
$$\forall s \in N. \tag{43}$$

Constraints (40)–(43) are used to model the VM slicing scenario. Constraint (40) ensures that the total normalised workload of all slices is equal to the original VM normalised workload before slicing. Constraints (41) and (42) ensure that the locations of the slices of a VM are consistent with those selected in constraints (34) and (35) and also they ensure that the slices normalised workload does not drop below the minimum allowed normalised workload per slice where M is a large enough number, with units of %, to ensure that $\delta_{sv} = 1$ when W_{sv} is greater than zero. Constraint (43) calculates the work load of each cloud by summing the load of the slices of the different VMs hosted by the cloud.

7) Single cloud scheme constraint

$$\sum_{s \in N} Cloud_s = 1. \tag{44}$$

Constraint (44) is used to model the single cloud scheme which is used as our benchmark to evaluate the power savings achieved by the different VM distribution schemes.

8) Number of processing servers

$$PSN_s = CW_s / W_{\max}$$
$$\forall s \in N. \tag{45}$$

Constraint (45) calculates the number of processing servers needed in each cloud. The integer value is obtained using the ceiling function.

B. Cloud VM Model Results

The VM placement schemes are evaluated considering the NSFNET network and the users distribution discussed in Section III-B. In addition to the input parameters in Table I, the VM model considers the parameters in Table IV. To reflect various users' processing requirements and represent different types of VMs, we uniformally assign normalised workloads to VMs from the set given in Table IV.

Fig. 17 shows the server CPU and network bandwidth utilization of 10 VMs at two times of the day (06:00 and 22:00). Note that network utilization is calculated by assuming a 10 Gb/s servers' interface speed and users traffic rates are kept at 5 Mbps as in Section III.

In a practical cloud implementation the CPU normalised workload needed is estimated by the cloud provider based on users' requirements. Instead of starting with users requests for VMs and assigning them to VMs, we simplify the generation of CPU normalised workload by considering a set of 1000 VMs (limit of what MILP can handle) of different types and assign each VM a uniformly distributed normalised workload between 10% and 100% of the total CPU capacity. We then randomly and uniformly assign each VM to serve a number of users. This approach is less complex to analyze in terms of number of variables and it captures the same picture. This can be understood by noting that a cloud provider will assign the incoming requests to a given VM according to its specialization up to a certain maximum normalised workload. This results in a distribution of VM normalised workloads and an assignment of users (from different nodes) to a VM which is what our approach also achieves.

Fig. 18(a) shows that the VM replication and migration schemes reduce to a single cloud scheme at low demand periods where all users are served from a cloud built in node 6 where all the 1000 VMs reside. For the migration scheme, node 6 is always (low traffic and high traffic) the optimum location for all VMs as it yields the minimum average hop count to all the network nodes given that users are uniformally distributed among nodes and requests for a VM are also uniformally distributed among users. For migration and at higher demands, and if the users connected to each VM have a Geo location clustering tendency, then there will be a benefit at high demand in migrating VMs nearer to such clusters. On the other hand for a uniform distribution of users there is fundamentally no benefit in migrating VMs. Our model concurs with this reasoning

TABLE IV
INPUT DATA FOR THE VM MODELS

Number of virtual machines (NVM)	1000
Server maximum power consumption (P_{max})	300 W
Server maximum normalised workload (W_{max})	100%
VM v total normalised workload (W_v)	RAND{10,20,30,40,50,60,70,80,90,100}%
Minimum allowed normalised workload per VM ($MinW$)	5%

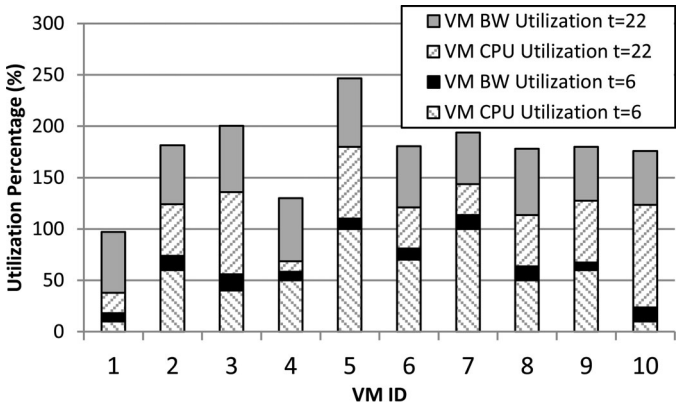

Fig. 17. Sample VMs CPU and network utilization.

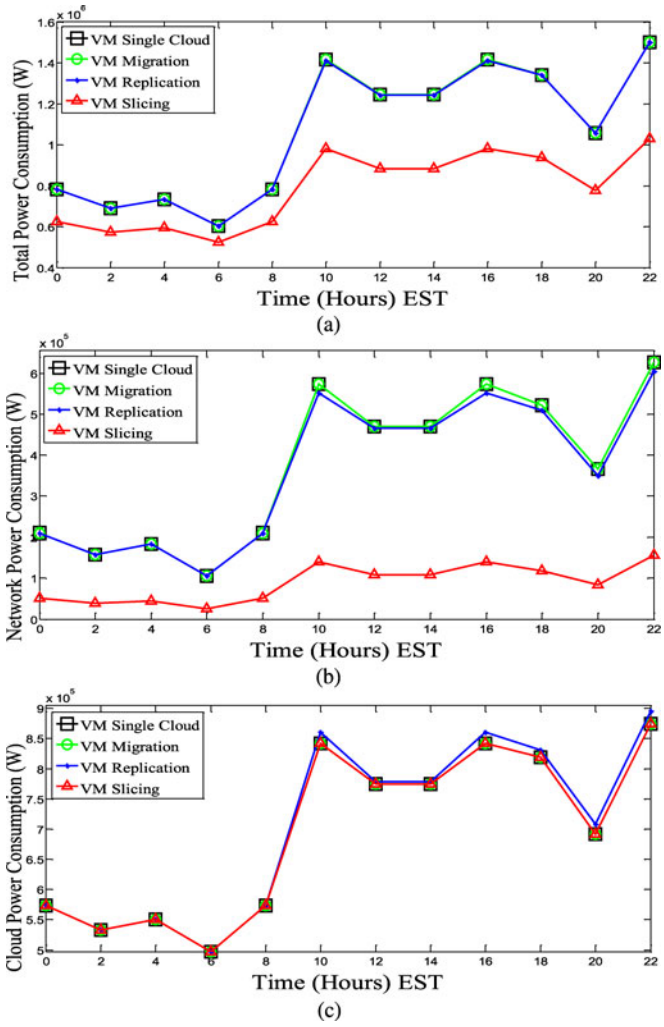

Fig. 18. (a) Total power consumption of the different VMs scenarios. (b) IP/WDM network power consumption of the different VMs scenarios. (c) Cloud power consumption of the different VMs scenarios.

and keeps the VMs at node 6 even at high demand under the migration scheme.

Under the replication scheme on the other hand the decision to make extra copies of a VM is driven by the tradeoff between the network power saved as a result of having extra copies of the VM that reduce network journeys, versus the increase in power as a result of the extra VMs. Therefore we found out that the model chooses to replicate lightly loaded VMs (i.e., VMs with load nearer to the 10%, away from 100%) as these consume less power. The replication scheme under uniformly distributed users obtains a limited network saving of 2% (see Fig. 18(b)) compared to the single cloud scheme. This is achieved by replicating a limited number of VMs (see Fig. 18(c)) with an overall network and VM combined power saving that is near zero, but positive.

The VM slicing scheme is the most energy efficient scheme as slicing does not increase the cloud power consumption (constraint (40)), allowing the VMs slices to be distributed over the network, yielding 25% and 76% total and network power saving, respectively compared to the single cloud scheme as shown in Fig. 18(a) and (b).

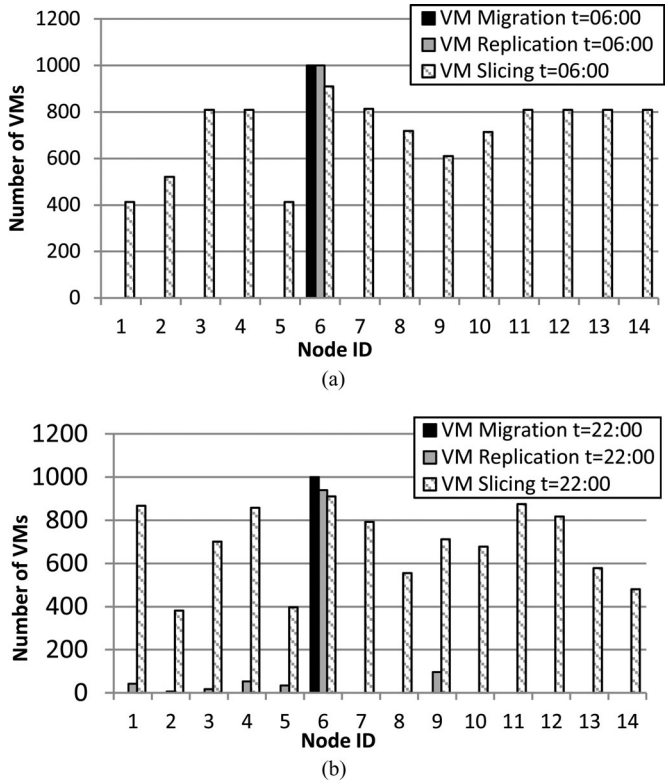

Fig. 19. (a) VMs distribution scheme at 06:00. (b) VMs distribution scheme at 22:00.

Fig. 19 shows the VMs placement over the network nodes for the three schemes at low and high demands. As discussed above the migration scheme places all VMs at node 6 all the time. At high loads the replication scheme, creates more than one copy of a limited number of VMs. The slicing scheme slices the machines so that each node has about 400–900 VMs at the different times of the day.

Note that the limited impact of the migration and replication schemes is due to the geographical uniformity of the traffic between VMs and nodes. Therefore we also evaluate a scenario with an extreme non-uniform distribution of requests for VMs where only users in a certain neighbourhood request a VM. In this scenario, only users in three neighbouring nodes (one physical hop between each) connect to a particular VM. For each of the 1000 VMs, we generate a uniform normalized workload between 10% and 100% and also generate for the same VM a random number uniformly distributed between 1 and N representing a network node, we then look up the single hop neighbourhood list of that node and select any other two neighbouring nodes. For example at 22:00, when there are 1.2 M users, we choose to allocate in this example to each VM an equal deterministic share, i.e., 1200 users, and so 400 users will be located in each of the three neighbouring nodes. This means that all users requesting a given VM are highly geographically localised.

Table V shows the power savings achieved by the different schemes compared to the single cloud scheme at two times of the day, 06:00 and 22:00. The power savings of the migration and replication schemes increase compared to the uniform traffic

TABLE V
POWER SAVING OF THE DIFFERENT SCHEMES COMPARED TO SINGLE CLOUD UNDER GEOGRAPHICALLY NON-UNIFORM TRAFFIC

Scheme	Power Saving			
	06:00		22:00	
	Total	Network	Total	Network
VM Migration	8%	48%	20%	49%
VM Replication	8%	48%	21%	59%
VM Slicing	16%	96%	40%	97%

Input: LIST= {6, 5, 4, 3, 7, 9, 13, 10, 11, 12, 14, 1, 8, 2}, $VM = \{1..NVM\}$

Output: Optimal Placement(J'), Total Power Consumption (TPC)

1. **For** each Virtual Machine $v \in VM$ **Do**
2. **For** each Placement $J \subseteq$ LIST **Do**
3. **For** each node $d \in N$ **Do**
4. **For** each location candidate $s \in J$ **Do**
5. (a) **Add**$\{\text{cost}_{sd}\} = MinHop$ (s, d)
6. (b) $CW_{vjs} = W_v$
7. (c) **End For**
8. (d) **Get** s where: $\text{cost}_{sd} = Min\{cost_{sd}\}$
9. $\text{L}_{vJsd} = \text{D}_{vd}$
10. **End For**
11. $\text{NPC}_{vJ} = MultiHopHeuristic\{N, Nm_i, L_{vJsd}\}$
12. $\text{CPC}_{vJ} = PUE\_c \cdot (SrvPC\_VM + LANPC\_VM)$
13. $\text{TPC}_{vJ} = \text{NPC}_{vJ} + \text{CPC}_{vJ}$
14. **End For**
15. $\text{TPC}_v = \text{Min}\{\text{TPC}_{vJ}\}$,
16. $J' = J$
17. **End For**
18. Calculate $\text{TPC} = \sum_{v \in VM} \text{TPC}_v$

Fig. 20. DEER-VM heuristic pseudo-code.

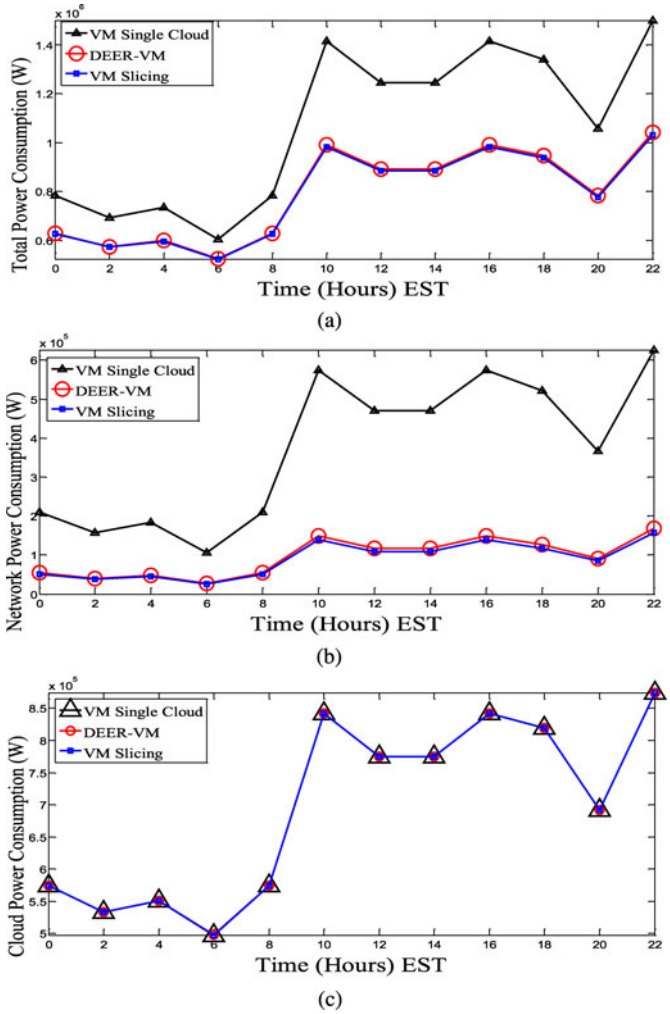

Fig. 21. (a) DEER-VM total power consumption. (b) DEER-VM IP/WDM network power consumption. (c) DEER-VM cloud power consumption.

scenario as the popularity of a VM in a certain neighbourhood justifies migrating or replicating it to that neighbourhood. The non-uniformity of traffic also allows the slicing scheme to save more power. The performance under the uniform and extreme non-uniform traffic distributions gives the lower and the upper bounds on the power savings that can be achieved by the different VM placement schemes.

C. DEER-VM: Energy Efficient VM Distribution for the Cloud

The results of the VM models showed that VM slicing is the most energy efficient VM placement scheme. In this section we develop a heuristic to perform VM-Slicing in real time using the same concept used in the DEER-CD heuristic.

The pseudo code of the DEER-VM heuristic is shown in Fig. 20. Nodes are ordered from the lowest to the highest based on their weight and the heuristic searches the 14 possible placements deduced from the LIST for each VM. The minimum hop count (loop (a)) is used to assign each node, to the nearest node hosting a VM which yields (L_{vJsd}), the traffic generated between nodes (s, d) due to placing VM v according to placement J (loop (b)). Knowing(L_{vJsd}), the heuristic calculates the total power consumption, TPC_{vJ}(loop (c)). VMs are located at the placement associated with the lowest power consumption among the 14 possible placements (loop (d)). The same process is repeated for other VMs till all VMs are placed.

In the absence of our sorted list approach, an exhaustive VM search is needed. In a similar fashion and as outlined for the DEER-CD, for NSFNET with $N = 14$ this is a complexity reduction by a factor of 1.6×10^{10}. Furthermore using a 2.4 GHz Intel Core i5 PC with 4 GB Memory, the heuristic took 35 min to evaluate the DEER-VM results.

Fig. 21(a) reveals that while the VM slicing approach has saved 25% of the total power compared to the single cloud scenario, the DEER-VM heuristic achieved 24%. This slightly lower saving is due to the multi hop non-bypass heuristic which is less efficient than the MILP model in routing traffic and therefore its network power consumption has increased by 6.4% as seen in Fig. 21(b).

As slicing the VM does not increase the power consumption of the machine, the results of model and heuristic considering VM slicing maintain the cloud power consumption of the Single Cloud scenario as shown in Fig. 21(c).

Table VI reports the optimization gaps between the DEER-VM heuristic and MILP.

TABLE VI
OPTIMIZATION GAPS BETWEEN THE DEER-VM HEURISTIC AND MILP

Metric	Gap (%)		
	06:00	22:00	Average
Total Power	0.36%	1%	0.73%
Network Power	7%	6.77%	6.47%
Cloud Power	0%	0%	0%

The 24% saving of DEER-VM is due to the IP/WDM network side rather than the cloud side as the CPU normalised workload is the same for both scenarios and the cloud servers' power consumption is normalised workload dependant. Most of the savings come from VMs which generate a lot of traffic in the network compared to their CPU utilization. Slicing such VMs reshapes the traffic pattern in the network so that users get their service from nearby locations, thereby saving network power consumption. However, if most VMs have high CPU utilization and low network traffic then slicing VMs will not save power and a Single Cloud scenario will be a better solution, especially taking VMs management cost into account. Therefore, VM slicing can play a major role in saving power consumption in VM based content delivery such as IPTV, and video on demand (VoD). The current version of the DEER-VM heuristic is designed to work under a uniform distribution of users as the sorted list is produced only once, however in the case of non-uniformly (geographically) distributed users, the placement of one VM will affect the placement selection of the next one. As such a progressively shorter list has to be sorted at each step. This is the subject of on-going work that we hope to report on in future. Note that the placement decision of one VM (and one *PG* for the content delivery model) is independent of the others; therefore, the heuristics can be distributed among different servers to reduce the computation time as different instances will work in parallel.

Note that our MILP model and heuristics are free to place content, storage or VMs in the locations that minimize power consumption and select the optimum route between the user and the content, stored documents or VMs such that power is minimized. The model also allows a user to be served by multiple VMs.

VI. CONCLUSION

This paper has introduced a framework for energy efficient cloud computing services over non-bypass IP/WDM core networks. We have analysed three cloud services, namely; content delivery, Storage as a Service (StaaS) and virtual machines based applications. A mixed integer linear programming (MILP) optimization was developed for this purpose to study network related factors including the number and location of clouds in the network and the impact of demand, popularity and access frequency on the clouds placement, and cloud capability factors including the number of servers, switches and routers and amount of storage required at each cloud. We have studied different replication schemes and analysed the impact of content storage size. Optimizing the cloud content delivery reveals that replicating content into multiple clouds based on content popularity (OPR scheme) is the optimum scheme to place content in core networks where at low traffic most of the content is kept in node 6 of the NSFNET network, while at high traffic more popularity groups are replicated into more clouds as under higher loads the power consumption of the cloud is compensated by higher savings on the network side. OPR resulted in 92% and 43% network and total power savings respectively, for content of small size such as music files. MPR is shown to have similar performance as OPR, however, under a scenario with a non-uniform user distribution or with fixed (idle) power component for placing a cloud in certain location, OPR is expected to outperform MPR as it might not be necessary to build 14 clouds, which we will investigate in future work. For real time implementation, we have developed an energy efficient content delivery heuristic, DEER-CD, based on the model insights. Comparable power savings are achieved by the heuristic. The results of the StaaS scenario show that at lower access frequencies, the optimal clouds scheme selects to serve all users from the central cloud while at higher access frequencies content is migrated to serve users locally, resulting in saving 48% and 2% of the network and total power compared to serving contents from a single central cloud for an average file size of 45MB. Limited total power savings are obtained for smaller file sizes. Optimizing the placement of VMs shows that VM Slicing is the best approach compared to migration or replication schemes. However this is under the assumption that the minimum normalized workload per slice is 5%. For VMs with larger minimum normalised workload per slice, slicing might approach VM migration in power saving as it will be difficult to have more local slices. VMs slicing saves 76% and 25% of the network and total power, respectively, compared to a single virtualized cloud scenario. Comparable power savings are obtained by placing VMs using a heuristic (DEER-VM) developed to mimic the model behaviour in real time.

REFERENCES

[1] N. Bessis, E. Asimakopoulou, T. French, P. Norrington, and F. Xhafa, "The big picture, from grids and clouds to crowds: a data collective computational intelligence case proposal for managing disasters," in *Proc. Fifth IEEE Int. Conf. P2P, Parallel, Grid, Cloud Internet Comput.*, Fukuoka, Japan, Nov. 4–6 2010.

[2] (2012). Cisco Global Cloud Networking Survey 2012. Cisco Systems. [Online]. Available: http://www.cisco.com/en/US/solutions/ns1015/2012_Cisco_Global_Cloud_Networking_Survey_Results.pdf

[3] T. Hansen. (2012). The future of knowledge work. Intel Corp, White Paper. Available: http://blogs.intel.com/intellabs/files/2012/11/Intel-White-Paper-The-Future-of-Knowledge-Work4.pdf

[4] B. P. Rimal, E. Choi, and I. Lumb, "A taxonomy and survey of cloud computing systems," in *Proc. Fifth Int. Joint Conf. INC, IMS IDC*, 2009, pp. 44–51.

[5] F. Rocha, S. Abreu, and M. Correia, "The final frontier: confidentiality and privacy in the cloud," *IEEE Comput.*, vol. 44, no. 9, pp. 44–50, Sep. 2011.

[6] (2013). Security and high availability in cloud computing environments. IBM Global Technology Services. [Online]. Available: http://www-935.ibm.com/services/za/gts/cloud/Security_and_high_availability_in_cloud_computing_environments.pdf.

[7] A. Beloglazov, R. Buyya, Y. Lee, and A. Zomaya, "A taxonomy and survey of energy-efficient data centers and cloud computing systems," *Advances Comput.*, vol. 82, pp. 47–111, 2011.

[8] J. Baliga, R. W. A. Ayre, K. Hinton, and R. S. Tucker, "Green cloud computing: Balancing energy in processing, storage, and transport," *Proc. IEEE*, vol. 99, no. 1, pp. 149–167, 2011.
[9] M. F. Habib, M. Tornatore, M. D. Leenheer, F. Dikbiyik, and B. Mukherjee, "Design of disaster-resilient optical datacenter networks," *IEEE /OSA J. Lightw. Tech.*, vol. 30, no. 16, pp. 2563–2573, Aug. 2012.
[10] B. Kantarci and H. T. Mouftah, "Designing an energy-efficient cloud network [Invited]," *IEEE /OSA J. Opt. Commun. Netw.*, vol. 4, no. 11, pp. B101–113, Oct. 2012.
[11] J. Buysse, C. Cavdar, M. De Leenheer, B. Dhoedt, and C. Develder, "Improving energy efficiency in optical cloud networks by exploiting anycast routing," in *Proc. Asia Commun. Photon. Conf. Exhib.*, Nov. 2011, pp. 1–6.
[12] J. Buysse, K. Georgakilas, A. Tzanakaki, M. De Leenheer, B. Dhoedt, and C. Develder, "Energy-efficient resource-provisioning algorithms for optical clouds," *IEEE /OSA J. Opt. Commun. Netw.*, vol. 5, no. 3, p. 226, Feb. 2013.
[13] A. H. Mohsenian-Rad and A. Leon-Garcia, "Energy-information transmission tradeoff in green cloud computing," in *Proc. IEEE Global Commun. Conf.*, 2010, pp. 1–6.
[14] A. Lawey, T. E. H. El-Gorashi, and J. M. H. Elmirghani, "Energy efficient cloud content delivery in core networks," in *Proc. Cloud Comput. Syst., Netw., Appl. Workshop*, 2013, pp. 1–7.
[15] G. Shen and R. Tucker, "Energy-minimized design for IP over WDM networks," *IEEE /OSA J. Opt. Commun. Netw.*, vol. 1, no. 1, pp. 176–186, Jun. 2009.
[16] J. Berthold, A. A. M. Saleh, L. Blair, and J. M. Simmons, "Optical networking: Past, present, and future," *IEEE /OSA J. Lightw. Tech.*, vol. 26, no. 9, pp. 1104–1118, May 2008.
[17] X. Dong, T. El-Gorashi, and J. M. H. Elmirghani, "IP over WDM networks employing renewable energy sources," *IEEE /OSA J. Lightw. Tech.*, vol. 29, no. 1, pp. 3–14, Jan. 2011.
[18] X. Dong, T. El-Gorashi, and J. M. H. Elmirghani, "Green IP Over WDM networks with data centers," *IEEE /OSA J. Lightw. Tech.*, vol. 29, no. 12, pp. 1861–1880, Jun. 2011.
[19] X. Dong, T. El-Gorashi, and J. Elmirghani, "On the energy efficiency of physical topology design for IP over WDM networks," *IEEE /OSA J. Lightw. Tech.*, vol. 30, no. 12, pp. 1931–1942, Jun. 2012.
[20] W. Jiang, R. Zhang-Shen, J. Rexford, and M. Chiang, "Cooperative content distribution and traffic engineering in an ISP network," in *Proc. SIGMETRICS*, 2009, pp. 239–250.
[21] L. Chiaraviglio and I. Matta, "An energy-aware distributed approach for content and network management," in *Proc. IEEE Comput. Commun. Workshops*, Apr. 2011, pp. 337–342.
[22] K. Guan, G. Atkinson, D. Kilper, and E. Gulsen, "On the energy efficiency of content delivery architectures," in *Proc. IEEE ICC Green Commun. Workshop*, Kyoto, Japan, Jun. 2011, pp. 1–6.
[23] M. Soraya, M. Zamani, and A. Abhari, "Modeling of multimedia files on the web 2.0," in *Proc. Canadian Conf. Elect. Comput. Eng.*, May 2008, pp. 1387–1392.
[24] (2013). Data sheet of glimmer glass intelligent optical system 500. Glimmerglass. [Online]. Available: http://www.glimmerglass.com/products/intelligent-optical-systems/
[25] (2014). Cisco ONS15501 erbium doped fiber amplifier data sheet. Cisco Systems. [Online]. Available: http://www.cisco.com/warp/public/cc/pd/olpl/metro/on15500/on15501/prodlit/ons15_ds.pdf
[26] (2014). Data sheet of Cisco ONS 15454 100-GHz 4-CH multi/ demultiplexer. Cisco Systems. [Online]. Available: http://www.cisco.com/en/US/prod/collateral/optical/ps5724/ps2006/product_data_sheet09186a00801a5572.pdf
[27] (2014). Global internet speed study. Pando Networks. [Online]. Available: http://www.pandonetworks.com/company/news/pando-networks-releases-global-internet-speed-study
[28] V. Valancius, N. Laoutaris, L. Massoulie, C. Diot, and P. Rodriguez, "Greening the internet with nano data centers," in *Proc. CoNEXT*, 2009, pp. 37–48.
[29] GreenTouch, "Green Meter Research Study: Reducing the Net Energy Consumption in Communications Networks by up to 90% by 2020," A GreenTouch White Paper, Version 1, Jun. 26, 2013, available: http//www.greentouch.org/uploads/documents/GreenTouch_Green_Meter_Research_Study_26_June_2013.pdf. Last access date 14 Nov. 2013.
[30] (2013). Efficiency: How we do it. Google. [Online]. Available: http://www.google.co.uk/about/datacenters/efficiency/internal/
[31] P. Mathew, S. Greenberg, D. Sartor, J. Bruschi, and L. Chu. (2010). Self-benchmarking guide for data center infrastructure: Metrics, benchmarks, actions. [Online]. Available: http://hightech.lbl.gov/benchmarking-guides/docs/datacenter-benchmarking-guide.pdf
[32] DropBox. (2013, Nov. 10). [Online]. Available: https://www.dropbox.com/business/pricing
[33] D. Huang, D. Yang, and H. Zhang, "Energy-aware virtual machine placement in data centers," in *Proc. IEEE Glob. Commun. Conf.*, Dec. 2012, pp. 3243–3249.
[34] H. Goudarzi and M. Pedram, "Energy-efficient virtual machine replication and placement in a cloud computing system," in *Proc. IEEE Fifth Int. Conf. Cloud Comput.*, Jun. 2012, pp. 750–757.
[35] B. Kantarci, L. Foschini, A. Corradi, and H. T. Mouftah, "Inter-and- intra data center VM-placement for energy-efficient large-scale cloud systems," in *Proc. IEEE GLOBECOM Workshop Manage. Security Technol. Cloud Comput.*, 2012, pp. 1–6.
[36] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future Generation Comput. Syst.*, vol. 28, no. 5, pp. 755–768, May 2012.

Ahmed Q. Lawey received the B.S. degree (first-class Hons.) in computer engineering from the University of Nahrain, Iraq, in 2002, the M.Sc. degree (with distinction) in computer engineering from University of Nahrain, Iraq, in 2005. He is currently working toward the Ph.D. degree at the School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K. From 2005 to 2010, he was a core network engineer in ZTE Corporation for Telecommunication, Iraq. His current research interests include energy optimization of IT networks, energy aware content distribution in the Internet, and energy efficient routing protocols in optical networks.

Taisir E.H. El-Gorashi received the B.S. degree (first-class Hons.) in electrical and electronic engineering from the University of Khartoum, Sudan, in 2004, the M.Sc. degree (with distinction) in photonic and communication systems from the University of Wales, Swansea, U.K., in 2005, and the Ph.D. degree in optical networking from the University of Leeds, Leeds, U.K., in 2010. She is currently a Postdoctoral Research Fellow at the School of Electronic and Electrical Engineering, University of Leeds. Her research interests include next-generation optical network architectures and green Information and Communication Technology.

Jaafar M. H. Elmirghani received the B.Sc. degree (first-class Hons.) in electrical and electronic engineering from the University of Khartoum, Sudan, in 1989 and the Ph.D. degree in the synchronization of optical systems and optical receiver design from the University of Huddersfield, U.K., in 1994. He is the Director of the Institute of Integrated Information Systems at the School of Electronic and Electrical Engineering, University of Leeds, U.K, and is an IEEE Distinguished Lecturer. He joined Leeds in 2007. Prior to that (2000–2007) he was a Chair in optical communications at the University of Wales Swansea, he founded, developed and directed the Institute of Advanced Telecommunications and the Technium Digital (TD), a technology incubator/spin-off hub. He has provided outstanding leadership in a number of large research projects at the IAT and TD. He has co-authored *Photonic switching Technology: Systems and Networks*, (Wiley) and has published more than 350 papers. He has research interests in optical systems and networks and signal processing. He was the Chairman of IEEE Comsoc Transmission Access and Optical Systems technical committee and was the Chairman of IEEE Comsoc Signal Processing and Communications Electronics technical committee, and an Editor of IEEE Communications Magazine. He was a Founding Chair of the Advanced Signal Processing for Communication Symposium which started at IEEE GLOBECOM'99 and has continued since at every ICC and GLOBECOM. He was also founding Chair of the first IEEE ICC/GLOBECOM optical symposium at GLOBECOM'00, the Future Photonic Network Technologies, Architectures and Protocols Symposium. He chaired this symposium, which continues to date under different names. He received the IEEE Communications Society Hal Sobol Award, the IEEE Comsoc Chapter Achievement Award for excellence in chapter activities (both in 2005), the University of Wales Swansea Outstanding Research Achievement Award, 2006, the IEEE Communications Society Signal Processing and Communication Electronics outstanding service award, 2009 and a best paper award at IEEE ICC'2013. He is currently an Editor of *IET Optoelectronics* and of *Journal of Optical Communications*, a Co-Chair of the GreenTouchWired, Core and Access Networks Working Group, an Adviser to the Commonwealth Scholarship Commission, a Member of the Royal Society International Joint Projects Panel and Member of the Engineering and Physical Sciences Research Council (EPSRC) College. He has been awarded in excess of £20 million in grants to date from EPSRC, the EU and industry and has held prestigious fellowships funded by the Royal Society and by BT. He is a Fellow of the IET and a Fellow of the Institute of Physics.