

Energy-Aware Adaptive Network Resource Management

M. Charalambides, D. Tuncer, L. Mamatas, G. Pavlou

Department of Electronic & Electrical Engineering

University College London

London, UK

Abstract—Resource over-provisioning is common practice in network infrastructures. Coupled with energy unaware networking protocols, this can lead to periods of resource under-utilization and constant energy consumption irrespective of the traffic load conditions. Driven by the rising cost of energy – and therefore OPEX – and increasing environmental consciousness, research in power saving techniques has recently received significant attention. Unlike the majority of previous work in the area, which has focused on centralized offline approaches, in this paper we propose an online approach by which the capacity of the network can be adapted in a decentralized fashion. Based on the modular architectures of modern IP routers, adaptation is achieved by configuring individual line cards to enter sleep mode. Re-configuration is performed periodically by intelligent ingress nodes that coordinate their actions in order to control the traffic distribution in the network, according to the actual demand. We evaluate our approach using real network topologies and traffic traces. In the case of the GEANT network, the proposed approach can, on average, reduce the energy to power the required line cards by 46% for a maximum utilization below 65%, and by 18% under heavily loaded conditions.

Keywords- *green networking; decentralized resource management; bundled links; virtualized routing planes; online traffic engineering.*

I. INTRODUCTION

Energy awareness has been the subject of technological developments over the past decade, ranging from simple energy-saving techniques for battery-powered computer equipment to more sophisticated ones applying to data centers [4]. The increasing power consumption of modern networks, driven by bandwidth-hungry applications, in conjunction with the rising cost of energy, has led researchers to investigate methods by which the carbon footprint of network infrastructures can be reduced. Although some work has been carried out in this area, effective energy management solutions are still missing.

Existing approaches in the literature that address energy efficiency in network infrastructures mainly propose offline and centralized solutions which assume the availability of traffic demand, e.g. [8][9]. However, these approaches can have sub-optimal performance under changing or unpredicted traffic conditions and also have inherent scalability limitations. The most prominent method proposed by which energy consumption can be reduced is powering off links/routers [10], since the alternative method of rate adaptation [12] has not demonstrated significant gain. However, switching off entire links/routers can disconnect the network topology, which, in

addition to possible packet losses under traffic variations, is not suitable for online re-configurations given that the process of computing the routing configuration is not trivial.

This paper proposes a new online energy-aware resource management approach which can be seen as a middle ground between the reduced topology and rate adaptation solutions. By exploiting the fact that many links in core networks are bundles of multiple physical cables, the proposed approach adapts the link capacity at run time by switching off individual line cards (LCs). This is achieved by controlling the traffic splitting ratios at network ingress nodes in a decentralized fashion, which results in offloading some LCs that can subsequently enter sleep mode. Our choice of switching off LCs as the means to save energy was motivated by the real measurements reported in [10]. According to these, the overall power profile of a network is dominated by the energy consumption of the router LCs and the chassis, with an approximate ratio of 3:1 for a fully loaded Cisco GSR 12008.

To achieve the energy-saving objective we extend our previous work on adaptive resource management [1] and employ an intelligent in-network substrate, which was originally used for load balancing. The substrate is a logical structure formed between the ingress nodes of a network and encapsulates the necessary logic to realize self-management functionality. Substrate nodes coordinate re-configuration actions that control the traffic distribution in the network. Given that the optimal mapping of traffic demand on available resources is an NP-hard problem, and also the strict time constraints of an online re-configuration process, such as the one proposed here, we devised an efficient heuristic algorithm. This is executed by substrate nodes periodically (every 15 minutes) and produces a new configuration in the form of traffic splitting ratios. Our approach is based on the path diversity provided by multi-topology routing (MTR), with the traffic volume on paths between source-destination pairs being defined by the computed ratios. We evaluate our approach using real topologies and traffic traces, and compare its performance against that of a load balancing algorithm and plain MTR. The results indicate that substantial energy gain can be achieved without significantly compromising the balance of the network in terms of load.

The rest of the paper is organized as follows. Section II describes background work that is essential for understanding the presented ideas. Section III provides an overview of the proposed approach and Section IV details our adaptive traffic distribution algorithm. Section V presents the results of our evaluation and Section VI discusses related work. Finally, conclusions and future directions are provided in Section VII.

II. BACKGROUND

Resource management in fixed networks is typically performed by offline centralized systems, which optimize network performance over long timescales. However, their static nature can lead to sub-optimal configurations under unexpected traffic demand. For this reason, research in online approaches has investigated methods by which configurations can be dynamically adapted in short timescales, according to real-time information, in the context of IP [6] and MPLS [5] networks. This section provides an overview of the adaptive resource management scheme we have developed in previous work [1], which serves as the basis of the energy-aware approach we are proposing.

A. Multi-Topology Routing (MTR)

To achieve their objectives most resource management approaches employ routing protocols that can support path diversity. MTR [7] is such a protocol and can provide a set of multiple routes between source-destination (S-D) pairs in the network. It extends the OSPF and IS-IS routing protocols by enabling a virtualization of a single physical network topology into several independent virtual IP planes.

The configuration of the different virtual planes is part of an offline process which computes a set of desired IP virtual topologies given the physical network topology. The derived topologies are such that two objectives are satisfied: a) a set of non-completely overlapping paths between S-D pairs is provided, and, b) critical links are not introduced, i.e. given a link l that is traversed by some traffic from node S to node D, there always exists an alternative path that can be used for routing the traffic without traversing l .

Fig. 1 illustrates a simple example of how virtual topologies that satisfy the aforementioned requirements can be derived from a base physical topology. We consider the S-D pair 1-3 where traffic at source node 1 is forwarded towards destination node 3. In each of the alternative topologies T1, T2 and T3, some links are assigned a *maximum* weight (represented here with infinity) which prevents these links from being used for routing the traffic demand between node 1 and node 3. With these settings, three non-overlapping paths can be determined between node 1 and node 3: (1;4;2;3), (1;4;5;3) and (1;2;3) without creating a critical link. The configuration of the alternative topologies is represented at the network level by associating a vector of link weights to each link in the network, each component of the vector being related to one topology.

B. Decentralized and Adaptive Resource Management

In previous work we have designed and developed a new intra-domain resource management approach for IP networks, in which the traffic distribution is controlled in an adaptive and decentralized manner according to network conditions [1][2][3]. Based on path diversity provided by MTR, traffic flows between any S-D pair are balanced across several paths according to splitting ratios, which are (re)-computed by network ingress (source) nodes. Periodic re-configurations (e.g. every 15 minutes) are controlled by an intelligent in-network substrate. As depicted in Fig. 2, this is a logical management structure, embedded in ingress nodes, which facilitates the communication between participating nodes and the execution

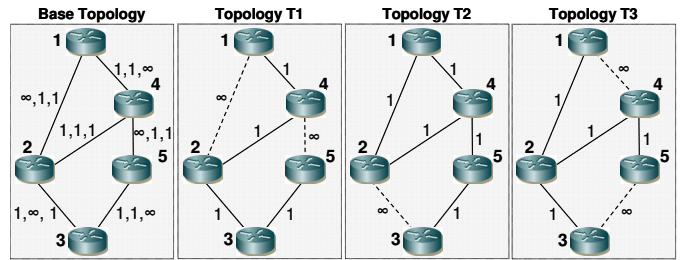


Figure 1. Building multiple topologies

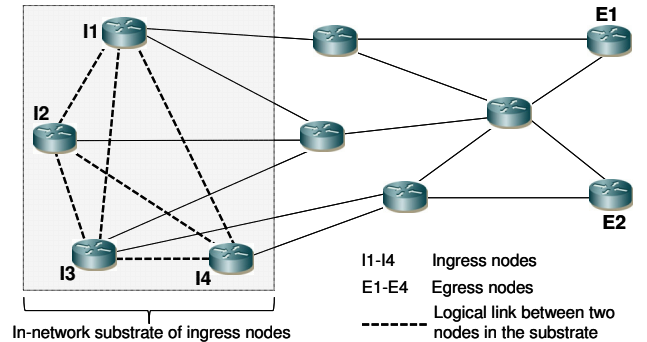


Figure 2. Full-mesh in-network substrate of ingress nodes

of a re-configuration algorithm. Although the substrate nodes shown here are arranged in a full-mesh, other structures can be used, each with its pros and cons. It should be noted that in this work we refer to a traffic flow as the volume of traffic between source and destination nodes.

During the adaptation process a sequence of re-configuration actions are decided in a coordinated manner between the source nodes in the substrate, with the objective of minimizing the link with the highest utilization l_{max} . At each iteration, they coordinate through the substrate to select one of them (called the Deciding Entity–DE) that will compute new splitting ratios over its locally originating traffic flows. Selecting a unique DE prevents inconsistencies between concurrent traffic splitting adjustments among multiple substrate nodes. This logic is represented by a static rule that allows the selection of the node associated with l_{max} .

While the DE is initially selected to perform re-configuration actions, it may not always be able to determine a configuration that satisfies the traffic engineering objective. In such cases, it sends a delegation request to the other substrate nodes (called Selected Entities–SEs), which compute new splitting ratios independently. Their results are communicated back to the DE, which then selects the configuration to apply (among successful ones), and notifies the relevant SE to enforce the configuration. For further details of the approach we refer the reader to [1].

III. ENERGY-AWARE DYNAMIC CAPACITY ADAPTATION

This section describes the proposed energy-aware resource management approach. Given that the approach is based on adapting the capacity of bundle links, we first illustrate link aggregation from an energy consumption viewpoint.

A. Link Aggregation

Aggregating multiple physical cables into a single logical link to connect IP routers has been standardized in [11] and is common practice in today’s core networks. The aggregation (or bundling) capability is a result of the modular architectures of modern routers, where multiple line cards (LCs) can be attached to the switching fabric. Under a single layer 3 logical address, bundled line cards can allow increased throughput beyond the capacity of a single connection, for example that of the fastest available link technology. Link aggregation can also be used as a flexible solution for network capacity upgrades whereby bundles are extended with new LCs instead of replacing existing links with higher-capacity ones.

Using a simple topology, Fig. 3 illustrates five different load distributions (separated by ‘|’) for an input traffic volume of 7.5 Gb/s, which is allocated by the source node R1 between three possible paths to the destination node R4. All bundled links consist of three LCs, each with a capacity of 2.5 Gb/s. In this work we consider all LCs in a bundle to have the same capacity, we view multiple ports on individual LCs as a single interface, and we assume that the allocation of traffic on a bundled link is carried out on a fill-first-available LC basis.

Table I summarizes the result of the traffic distribution examples in terms of the number of LCs used, with each unbundled link being associated with two LCs, (one at either end, e.g. egress of R1 and ingress of R4). The least number of LCs used is achieved in case (a), where all traffic is routed through the shortest path (R1-R4) of one hop. The LC number doubles to 12 when using the longest path (R1-R2-R4) of two hops in (b). Splitting the traffic with ratios of 0.75 and 0.25 between the longest paths R1-R2-R4 and R1-R3-R4, respectively in case (c), does not improve the cost compared to (b). However, equal splitting among the three paths in (d) reduces the LCs used by 2 since a third of the traffic is now routed over the shortest path. The last example (e) concerns the case where the load on a bundled link is not a multiple of the LC capacity. As a result, some LCs are not utilized to their full capacity and thus a higher number is required to accommodate the load.

Measuring the energy consumption of a network by the number of LCs used and based on the above, some observations can be made: (i) using shortest paths incurs the least cost, but can lead to congestion; (ii) optimized splitting ratios can reduce energy consumption with load balancing mitigating congestion; (iii) non-fully utilized LCs incur a high cost and should be avoided where possible. These observations were the key foundations in the design of the proposed algorithm.

B. Approach Overview

In contrast to the main body of work in energy efficiency of network infrastructures, which propose centralized offline solutions, we are proposing a flexible approach by which the configuration of a network can be adapted dynamically, in a decentralized fashion, to meet energy efficiency objectives.

To reduce the energy required to sustain the operation of a network, our approach is based on the use of multiple LCs to implement the links between IP routers and the capability of

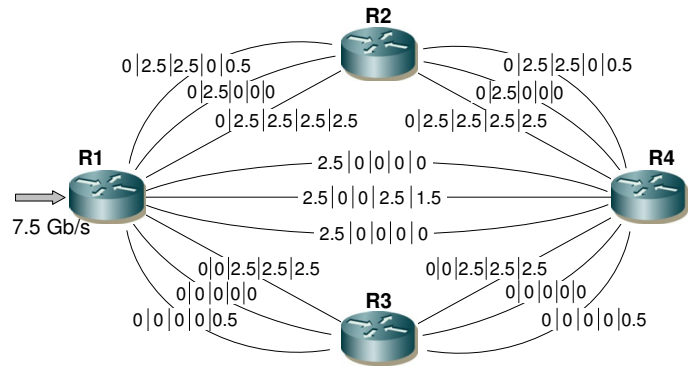


Figure 3. Examples of traffic distribution among multiple line cards

TABLE I. LINE CARD USAGE EXAMPLES

| Router | LCs Available | Line Cards Used | | | | |
|--------|---------------|-----------------|-----|-----|-----|-----|
| | | (a) | (b) | (c) | (d) | (e) |
| R1 | 9 | 3 | 3 | 3 | 3 | 5 |
| R2 | 6 | 0 | 6 | 4 | 2 | 4 |
| R3 | 6 | 0 | 0 | 2 | 2 | 4 |
| R4 | 9 | 3 | 3 | 3 | 3 | 5 |
| Total | 30 | 6 | 12 | 12 | 10 | 18 |

individual line cards to enter sleep mode. To achieve the latter, traffic is distributed in such a way that some of the LCs are not utilized, i.e. no traffic is transmitted through them. This is controlled by dynamically adapting the splitting ratios applied to incoming traffic at the ingress nodes, according to a re-configuration algorithm. The resulting configuration can allow some LCs to sleep, which thus adapts the link capacity. The main advantage of this approach compared to very recent work in [14] and [17], is that the physical topology does not get disconnected and new S-D paths do not need to be computed.

The re-configuration process is managed by the intelligent in-network substrate that we have successfully used for load balancing purposes in our previous work [1]. The substrate structure we use in this paper is one in which participating nodes are logically connected in a full-mesh. As described in Section II, the adaptation is performed periodically in short time scales (i.e. every 15 minutes) at which point substrate nodes coordinate their actions for the computation of new traffic splitting ratios in a decentralized manner. This time period was appropriate for our experiments so that traffic fluctuations could be followed without introducing frequent configuration changes. Adaptation is an iterative process with only one substrate node permitted to assume the role of the Deciding Entity (DE) at each iteration and subsequently execute the re-configuration algorithm. This selection is based on a predefined rule that allows the DE to associate itself with the bundled link from which traffic will be removed. Upon failure to determine an acceptable configuration, the DE sends a delegation request to the other substrate nodes. These execute the re-configuration algorithm concurrently and communicate their results back to the DE, which subsequently chooses which one should be applied.

Line cards can be *full* (LCF) if their load is equal to their capacity, *utilized* (LCU) if their load is not zero and less than

their capacity, and *non-utilized* (LCN) if they have zero load. With the objective to offload traffic from as many utilized line cards as possible, one of the key decisions in the proposed approach concerns the bundled link to consider for (a) removing traffic from, and (b) assigning that traffic to, at each iteration of the re-configuration process. This decision is based on a ranked list of all utilized LCs in the network according to their load (LCU_load_{BL})¹, which is common to all substrate nodes. Based on the load ($load_{BL}$) and the LC capacity (LC_cap_{BL}) of a bundled link (BL), LCU_load_{BL} can be determined as follows:

$$LCU_load_{BL} = LC_cap_{BL} \times \text{mod} \left(\frac{load_{BL}}{LC_cap_{BL}} \right)$$

The ranked list can be computed and provided to all substrate nodes by a distributed monitoring mechanism such as the one proposed in [18]. The impact of the monitoring system (e.g. monitoring accuracy or other capabilities) on the performance of our approach requires an independent study and is a subject of future work. Ranking is based on the load of utilized LCs, or on the collective load of BLs with a utilized LC, in increasing order. In our experiments, the two options produced similar results. The goal of the DE is to move the traffic load of the first LC in the list to another LC further down the list that can accommodate this load and thus potentially fill-up its remaining capacity. This involves the computation of new splitting ratios by the DE, which results to different volume of traffic sent over each of the virtualized MTR planes. The list is re-ranked at each iteration and, with the principle that traffic cannot be assigned to a LC from which traffic has been removed in a previous iteration, the process terminates when all the LCs from which traffic can be removed have been exhausted. To prevent the disconnection of S-D pairs during the re-configuration process, the approach does not allow a critical BL (one that is used by a S-D pair in every virtual topology) to be assigned with a zero splitting ratio, i.e. at least one LC is active (either full or utilized) in that BL. Our algorithm is detailed in the next section.

IV. ADAPTIVE TRAFFIC DISTRIBUTION ALGORITHM

Previous work in energy-aware networking identified that the optimal allocation of traffic to achieve green objectives is an NP-hard problem [9][21]. To satisfy the strict time constraints of an online re-configuration process, such as the one proposed, we devised an efficient heuristic algorithm. The main functionality of this algorithm involves determining the BLs to remove traffic from and re-assign that traffic to, as well as the amount of traffic to shift.

A. Definitions

Each bundled link is associated with a state S , which is expressed with two bits as follows:

- S_{00} – traffic can be both removed from and assigned to a BL.
- S_{01} – traffic can be removed from, but cannot be assigned to a BL.

¹ Subscript BL represents the id of the particular bundled link a LC is associated with.

- S_{10} – traffic cannot be removed from, but can be assigned to a BL.
- S_{11} – traffic can neither be removed from, nor can be assigned to a BL.

When generating a new configuration, in addition to the traffic load on the utilized LCs of BLs, the algorithm requires the amount of spare capacity ($spare_{BL}$) on utilized LCs for these links. This is determined as follows:

$$LCU_spare_{BL} = LC_cap_{BL} - LCU_load_{BL}$$

B. Traffic Re-assignment – Where?

At each iteration, the re-configuration entity (DE or delegated substrate node) executing the algorithm determines the volume of traffic that can be re-directed to offload a utilized LC. This is achieved by selecting the first element (LCU_fst_{BL}) from the ranked list of utilized LCs and determining all the traffic flows emanating from the re-configuration entity that are routed over the associated BL (BL_{LCU_fst}). Any flows for which BL_{LCU_fst} is a critical link are disregarded so that the physical topology can remain connected.

The next step is to determine, for each of these flows, the virtual topologies from which traffic can be removed and the ones to which this traffic can be re-assigned. This is signified by the two state bits of BL_{LCU_fst} as defined in the previous section, which are updated at each iteration. To achieve our energy reduction objective we do not allow the offloading of full LCs or the traffic assignment to non-utilized LCs. As such, BLs having all their LCs either full or non-utilized are set to state S_{11} . BLs with a utilized LC are set to state S_{00} . The flows over a virtual topology using BL_{LCU_fst} are marked for removal if all the BLs along their paths have a first state bit of 0. Conversely, the alternative topologies (where BL_{LCU_fst} is not used) over which traffic can be re-directed are the ones that have all the BLs with a second state bit of 0 along the paths of the considered flows. This results to a list of flows that can be removed from LCU_fst_{BL} and assigned to other utilized LCs.

C. Traffic Re-assignment – How Much?

Once the flows that can potentially be shifted have been determined, they are removed iteratively from the relevant topologies. This process terminates if the list of flows is exhausted or if the entire load on LCU_fst_{BL} has been removed. It may not always be possible to move the entire volume of a flow since the available capacity on alternative virtual topology paths acts as a constraint. For this reason, the proposed algorithm computes the maximum volume that can be removed from each flow in three steps:

- 1) For each topology where LCU_fst_{BL} is involved, the LC with the least load along the associated path is determined. The minimum of these values over the relevant topologies is subsequently computed, $LCU_minLoad_{BL}$.
- 2) For each alternative topology where LCU_fst_{BL} is not involved, the utilized LC with the least spare capacity along the associated path is determined. The minimum of these values over the relevant topologies is subsequently computed, $LCU_minSpare_{BL}$.

3) The maximum volume of traffic to remove from LCU_fst_{BL} is $LCU_minSpare_{BL}$ if:

$$LCU_minLoad_{BL} > LCU_minSpare_{BL}$$

and $LCU_minLoad_{BL}$, otherwise. In the case where LCU_fst_{BL} is involved in multiple virtual topologies, traffic is removed equally between them.

Having determined the volume of traffic to remove from each topology, the algorithm ranks the utilized LCs with the least spare capacity, computed in step 2) above, in decreasing order. Starting from the top of this list, the removed traffic is assigned to the relevant topology, filling up LCUs, until the entire traffic volume has been allocated. The resulting total traffic volume assigned to each virtual topology allows the algorithm to compute new splitting ratios.

At each iteration of the re-configuration process, the ranked list of utilized LCs is updated and new BL states are set. State S_{11} is assigned to a BL whose LCs become either full or non-utilized and state S_{01} to a BL if LCU_fst_{BL} was a constituent LC, but its entire load was not successfully removed in the previous iteration. The latter state is communicated to all nodes of the intelligent substrate so that traffic is not assigned to the associated LCU during the next iteration. Alternate actions of traffic removal and assignment on the same LC (flip-flops), which can lead to configuration instabilities, can thus be avoided.

V. EVALUATION

To evaluate the performance of our energy-aware approach we used real topologies and traffic traces from two PoP-level networks, Abilene [26] (12 nodes) and GEANT [27] (23 nodes). This section presents the results from experiments, which were carried out on a high-end laptop, in terms of LC gain and maximum utilization.

A. Experimentation Setup

For all experiments we have used 672 traffic matrices (TMs) covering a period of 7 days so that a wide range of traffic conditions are taken into account. We use 4 virtual topologies for Abilene and 5 for GEANT, which were generated according to the guidelines described in Section II.A. With the bundled LCs having the same capacity, Table II summarizes the configuration of the BLs in the two physical topologies. The 30 unidirectional BLs of Abilene are implemented with 120 LCs and the 74 BLs of GEANT with 222 LCs.

We compare the performance of our approach (abbreviated NRG on plots and tables) against: (i) basic MTR (MTR_basic) with static splitting ratio (does not adaptively change) equal to the inverse of the capacity of the bottleneck link in each virtual topology, (ii) the adaptive resource management scheme (DACORM) described in [1] with changing splitting ratios that achieve a load balancing objective. Adaptation is performed every 15 minutes in both Abilene and GEANT topologies with a maximum of 50 algorithm iterations. At the start of each re-configuration cycle we initially use the splitting ratios generated in the previous cycle. Utilized LCs are ranked according to their load in increasing order; ranking based on the load of BLs has demonstrated similar results. For each of

TABLE II. ABILENE AND GEANT BLS CONFIGURATION

| | Abilene | | GEANT | | |
|--------------------|---------|--------|--------|--------|--------|
| | Type 1 | Type 2 | Type 1 | Type 2 | Type 3 |
| BL capacity (Gb/s) | 10 | 2 | 10 | 2 | 1 |
| BL size | 4 | 4 | 4 | 2 | 2 |
| LC capacity (Gb/s) | 2.5 | 0.5 | 2.5 | 1 | 0.5 |
| Number of BLs | 28 | 2 | 37 | 17 | 20 |
| Number of LCs | 112 | 8 | 148 | 34 | 40 |

the schemes we compare against, we measure the gain (G) in terms of number of LCs used to accommodate the traffic and the maximum utilization in the network (max-u), i.e. typically used as a measure of the load balancing level.

In addition to evaluating the proposed approach under the load imposed by the available TMs, we investigate its performance under heavily loaded traffic conditions. Instead of scaling all the TMs, which increases the size of every flow, we introduce background traffic that fully consumes one LC on every BL. This selectively increases the size of only some flows. We structure the results presented in this section in the presence and absence of background traffic.

B. Definitions

The following variables, which are related to the BLs in the network, are used when evaluating the performance of the proposed approach: the bundle size (BL_sz_{BL}), i.e. the number of LCs that implement a BL, the number of full (LCF_count_{BL}), utilized (LCU_count_{BL}), and non-utilized (LCN_count_{BL}) LCs. These are calculated as follows:

$$BL_sz_{BL} = \frac{BL_cap_{BL}}{LC_cap_{BL}}$$

$$LCF_count_{BL} = \text{int} \left(\frac{load_{BL}}{LC_cap_{BL}} \right)$$

$$LCU_count_{BL} = \begin{cases} 0 & \text{if } \text{mod} \left(\frac{load_{BL}}{LC_cap_{BL}} \right) = 0 \\ 1 & \text{otherwise} \end{cases}$$

$$LCN_count_{BL} = BL_sz_{BL} - (LCF_count_{BL} + LCU_count_{BL})$$

The gain G associated with an energy-aware re-configuration (NRG) compared to a non-NRG one, e.g. basic MTR or DACORM, is defined as follows:

$$G = 1 - \frac{(LCF_tot + LCU_tot)_{NRG}}{(LCF_tot + LCU_tot)_{non_NRG}}$$

where the total number of full (LCF_tot) and utilized (LCU_tot) LCs in the network are computed by:

$$LCF_tot = \sum_{BL \in \lambda} LCF_count_{BL}$$

$$LCU_tot = \sum_{BL \in \lambda} LCU_count_{BL}$$

with λ being the set of BLs in the network.

C. Algorithm Performance – No Background Traffic

The first set of experiments concerns the case where background traffic is not present. Table III summarizes the performance of our approach (NRG), in terms of active LCs, compared to MTR_basic and DACORM for the Abilene and GEANT topologies. A much higher average gain can be observed in the case of GEANT compared to Abilene. Our approach performs better on average compared to MTR_basic (22%) rather than DACORM (18%) in the case of Abilene, but the performance is very similar in the case of GEANT. There were no instances where the gain is negative.

Fig. 4 plots the distribution of the gain for the Abilene network. Compared to MTR_basic, our approach achieves a gain between 19% and 24% for more than 85% of the TMs, with the most frequent gain of 21% applying to around 40% of the TMs. The comparison with DACORM shows that a gain between 16% and 19% is achieved for more than 85% of the TMs, with the most frequent gain of 16% applying to around 60% of the TMs. In both comparisons the gain achieved by our approach is always non-negative, with the worst case being the configuration for only one TM, for which there was no gain. Compared to DACORM, the gain is slightly smaller on average than that of MTR_basic, but comparable in terms of order of magnitude. The results for the GEANT topology in Fig. 5 show that our approach achieves a gain between 44% and 48% compared to both MTR_basic and DACORM for more than 85% of the TMs, with the most frequent gain of 46% applying to around 53% of the TMs. As in the case of Abilene, the gain is always non-negative.

Figs. 6 and 7 plot the evolution of the maximum utilization (max-u) under the three different schemes in the Abilene and

TABLE III. % GAIN WITHOUT BACKGROUND TRAFFIC

| Gain (%) | Abilene | | GEANT | |
|----------|-----------------|--------------|-----------------|--------------|
| | NRG V MTR_basic | NRG V DACORM | NRG V MTR_basic | NRG V DACORM |
| Minimum | 6.06 | 0 | 17.02 | 17.02 |
| Maximum | 32.43 | 34.21 | 51.06 | 51.06 |
| Average | 21.85 | 18.03 | 46.10 | 46.08 |

GEANT networks, respectively. For Abilene, the maximum utilization achieved by our approach is significantly lower than that of MTR_basic, but higher than that of DACORM by 25% on average. Flat-lining for NRG occurs because the BL with the highest utilization remains the same over a period of time. In the case of GEANT, max-u for NRG follows the same trend but is higher than that of the two other schemes, with an average difference of 47% compared to DACORM – almost double the deviation compared to Abilene.

D. Algorithm Performance – Background Traffic

Table IV summarizes the performance of our approach, in the presence of background traffic, compared to MTR_basic and DACORM for the Abilene and GEANT topologies. As in the absence of background traffic, a higher average gain can be observed in the case of GEANT compared to Abilene. Our approach performs better on average compared to MTR_basic (11%) rather than DACORM (8%) in the case of Abilene, but the performance is very similar in the case of GEANT. There were no instances where the gain is negative.

Fig. 8 plots the distribution of the gain for the Abilene network. Compared to MTR_basic, our approach achieves a gain between 9% and 12% for more than 90% of the TMs, with

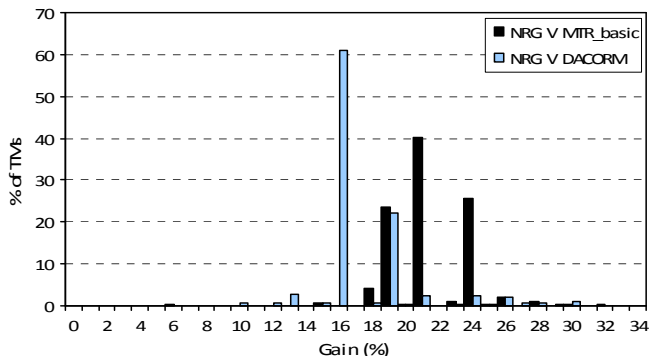


Figure 4. Gain distribution for Abilene

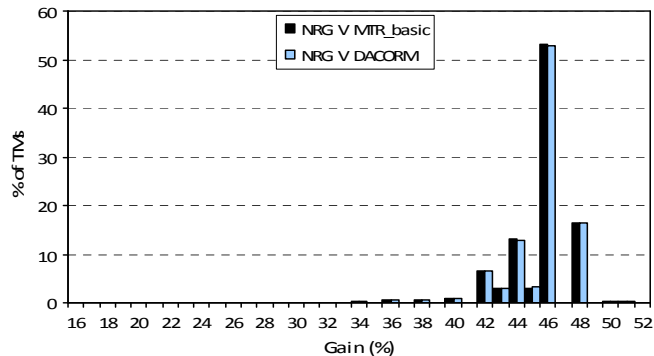


Figure 5. Gain distribution for GEANT

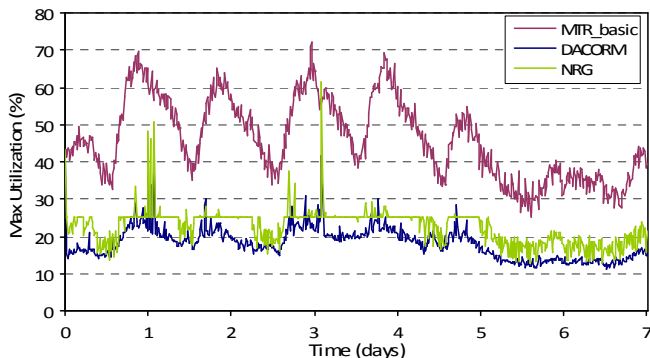


Figure 6. Evolution of maximum utilization in Abilene

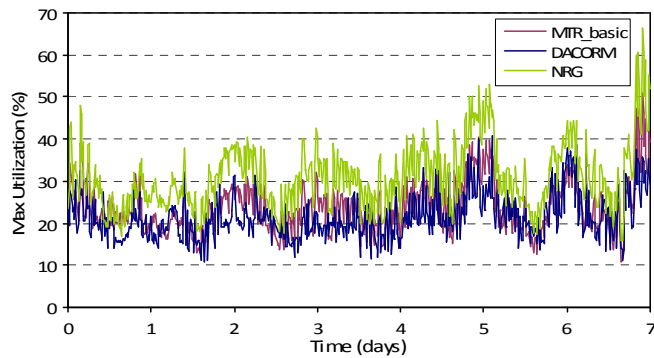


Figure 7. Evolution of maximum utilization in GEANT

TABLE IV. % GAIN WITH BACKGROUND TRAFFIC

| Gain (%) | Abilene | | GEANT | |
|----------|--------------------|-----------------|--------------------|-----------------|
| | NRG V MTR_basic | NRG V DACORM | NRG V MTR_basic | NRG V DACORM |
| Minimum | 3.17 | 0 | 6.61 | 6.61 |
| Maximum | 17.91 | 19.12 | 20.33 | 20.32 |
| Average | 11.33 | 8.41 | 17.92 | 17.91 |

the most frequent gain of 11% applying to around 40% of the TMs. The comparison with DACORM shows that a gain between 6% and 9% is achieved for more than 90% of the TMs, with the most frequent gain of 8% applying to around 84% of the TMs. Compared to DACORM, the gain is slightly smaller on average than that of MTR_basic, but comparable in terms of order of magnitude. The results for the GEANT topology in Fig. 9 show that our approach achieves a gain between 16% and 19% compared to both MTR_basic and DACORM for more than 90% of the TMs, with the most frequent gain of 18% applying to around 52% of the TMs. As in the case of Abilene, the gain is always non-negative.

Figs. 10 and 11 plot the evolution of max-u under the three different schemes in the Abilene and GEANT networks, respectively. For Abilene, the maximum utilization achieved by our approach is significantly lower than that of MTR_basic, but higher than that of DACORM by 7% on average. In the case of GEANT, max-u for NRG is higher than that of the two other schemes, but not as much as in the case of Abilene, with an average deviation of 7% compared to DACORM.

E. Analysis and Discussion

The gain distribution plots illustrate strong concentration around the average in all cases indicating that the performance of the proposed approach is consistent with most TMs. The higher gain observed in the case of GEANT compared to Abilene, in both the presence and absence of background traffic, is attributed to the variation in the GEANT BL capacities. Offloading LCs of smaller capacity is a simpler task since less traffic needs to be shifted. As such, the capacity homogeneity of the Abilene BLs (only 6.6% are 2 Gb/s links) acts as a limiting factor. Furthermore, it is evident that the average gain in the presence of background traffic, for both GEANT and Abilene networks, is lower. This is because the proportion of utilized LCs (that can be switched off) in the total number of active LCs (sum of utilized and full) is always lower under loaded traffic conditions.

As expected, the lowest max-u in all experiments is obtained by DACORM, which is a load balancing approach with the objective of minimizing the maximum utilization in the network. The difference in max-u obtained by NRG for Abilene and GEANT is generally insignificant demonstrating that energy conservation may come with a degree of load balancing. We plan to further investigate the cases where this is not true (e.g. Fig. 7). As such, methods that attempt to tune the two objectives better (i.e. load balancing and energy efficiency) will be part of our future work.

Given that the proposed algorithm is executed by ingress nodes, its time complexity is theoretically defined by the

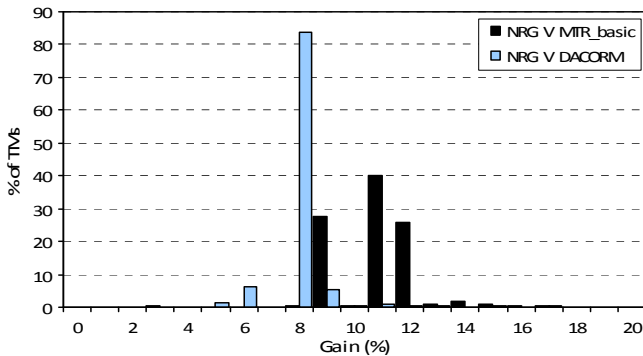


Figure 8. Gain distribution for Abilene

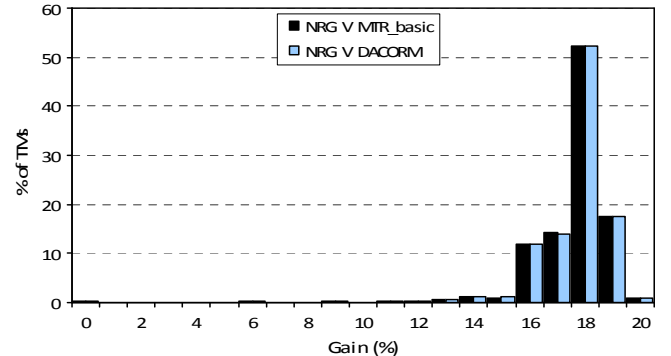


Figure 9. Gain distribution for GEANT

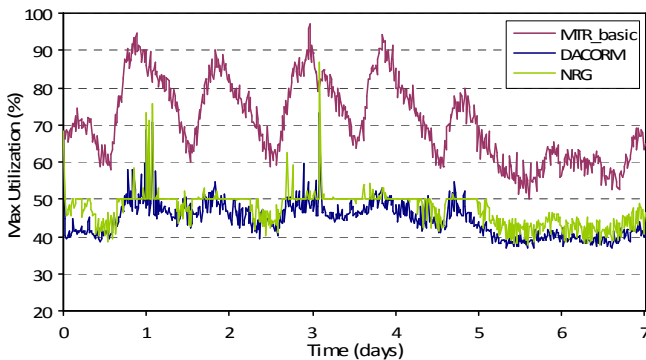


Figure 10. Evolution of maximum utilization in Abilene

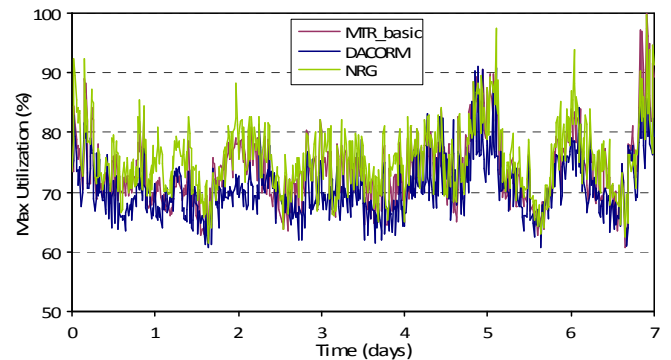


Figure 11. Evolution of maximum utilization in GEANT

number of local traffic flows. In a PoP-level topology with N nodes, the total number of local flows is $N(N-1)$ and the complexity is thus in the order of $O(N^2)$. In practice, however, the complexity is directly driven by the actual number of flows that need to be considered such that all the traffic can be removed from a utilized LC. In the best case, a solution can be determined by considering a single local flow. The average execution time of our algorithm is 15.23ms for Abilene and 15.43ms for GEANT, without significant deviations from these values. Despite the fact that GEANT has almost four times as many local flows as Abilene (506 compared to 132), the execution time for the two topologies is similar. This result shows that the time complexity does not increase quadratically in practice since a similar number of flows are considered for determining a solution in both Abilene and GEANT. The complexity of the overall adaptation process is influenced by the number of re-configuration iterations and the frequency of delegation since the latter incurs communication overhead. For further details we refer the reader to our previous work in [1].

VI. RELATED WORK

Resource management in fixed networks is recently adopting the objective of energy efficiency, attempting to associate network conditions with energy consumption better. The common strategy of resource over-provisioning with constant energy consumption will be gradually replaced by new adaptive and deployable approaches, balancing well performance with energy-efficiency.

Proposed solutions are based on various assumptions for the network environment. They are considering networks with bundled links, links with single LCs or even going deeper at the level of the LC ports. The bundled link could be a fiber and its set of WDM channels as in [13], or other types of fixed physical links. The topology could be dynamic, in the sense that it is associated with reduced connectivity graphs due to strategies that turn off LCs and corresponding links, e.g. [14] [15][16]. Other approaches assume fixed topologies but dynamic traffic conditions, at short or longer time scales, e.g. [21][22]. In [13], the authors define theoretical upper-bounds in the energy-saving potential for different types of dynamicity in the environment, including topology and routing strategies. Such assumptions could be realistic but are associated with reasonable complexity.

The assumptions for the network environment call for mechanisms with a certain level of network environment awareness. Most of the proposals use integrated monitoring facilities (e.g. [8][14]) that may be restricted to the studied context or not detailed at a sufficient level. We argue that a monitoring or information management system could be decoupled from the resource management solution. It is a complicated aspect that should be investigated independently. Example platforms that could be used are [18][19][20].

A schema that is aware of the network environment conditions should be able to take energy-saving actions as responses to these conditions. The type of actions range from pruning the network topology (e.g. exploiting algebraic connectivity [15]), routing changes (e.g. of shortest path tree [14]), to traffic engineering approaches (e.g. setting traffic split ratio among multiple paths [21][22]). The traffic engineering

approaches are usually using MPLS tunnels or virtual network links (e.g. [22][23]). Furthermore, the majority of the proposed solutions take actions in an offline or centralized manner (e.g. [9][21]). Very few of them have distributed or online considerations, e.g. [15], and others targeting these issues are still at an early stage [22]. Other works adopt approaches belonging in the middle ground between the offline and online solutions; for example using topology switching [8].

By performing capacity adaptation instead of topology adaptation, our approach prevents the disconnection of S-D pairs during the re-configuration process. In conjunction with MTR, this allows the computation of alternative paths in advance, thus reducing complexity at the routing level. Furthermore, we consider dynamic traffic conditions, at short or long scales. This allows our approach to adapt to the dynamicity of fixed infrastructures, following changes in user surf habits and available network applications. Of course, this level of complexity could be tackled due to the previous reasonable assumptions. Monitoring of dynamic conditions is a complicated subject that requires an independent infrastructure. In our case, we use the information management overlay proposed in [18]. The decentralized and online nature of our approach avoids the scalability problems encountered by centralized proposals, which constitute the majority of the work in this area. Last but not least, an important aspect of our approach is deployability, since the logic to realize the energy objective requires modifications only in the ingress nodes of a network, but not in the core ones.

VII. CONCLUSIONS AND FUTURE WORK

This paper describes a new resource management approach for IP networks that addresses the very important issue of energy efficiency. The objective is achieved by controlling the traffic distribution through an intelligent substrate at the edges of the network. In contrast to the majority of previous work that suggests switching off entire links, the proposed approach is more flexible since source-destination pairs always remain connected and thus new routing configuration is not required. Being decentralized and online, it overcomes many of the limitations of existing solutions in the literature. Furthermore, the efficiency of the heuristic employed by the approach makes it suitable for an online re-configuration process compared to the rather ‘heavy’ algorithms previously proposed.

Although the evaluation results demonstrate substantial energy gain, the performance can be further enhanced by complementing our approach with an offline one (e.g. [8]) that computes MTR topologies for different intervals of the day (e.g. reduced topologies for off-peak times). In future extensions of this work we plan to develop and evaluate additional heuristic algorithms and investigate possible energy gains by controlling individual ports on router line cards. We also plan to investigate the use of a hybrid structure to realize the intelligent substrate so that the management overhead can be further minimized.

ACKNOWLEDGMENT

This work was partly funded by Flamingo, a Network of Excellence project (ICT-318488) supported by the European Commission under its Seventh Framework Programme.

REFERENCES

- [1] D. Tuncer, M. Charalambides, G.Pavlou, N. Wang, "DACORM: A coordinated, decentralized and adaptive network resource management scheme," Proceedings of 13th IEEE/IFIP Network Operations and Management Symposium (NOMS), Hawaii, USA, April 2012.
- [2] M. Charalambides, G. Pavlou, P. Flegkas, N. Wang, D. Tuncer, "Managing the future Internet through intelligent in-network substrates," IEEE Network, Special Issue: Managing an Autonomic Future Internet, Vol. 25, No. 6, Nov/Dec 2011.
- [3] D. Tuncer, M. Charalambides, G.Pavlou, N. Wang, "Towards decentralized and adaptive network resource management," Proceedings of 7th IEEE/IFIP Conference on Network and Service Management (mini-CNSM), Paris, France, October 2011.
- [4] L. Liu, H. Wang, X. Liu, X. Jin, W.B. He, Q.B. Wang, Y. Chen, "GreenCloud: a new architecture for green data center," Proceedings of 6th IEEE/ACM International Conference on Autonomic Computing and Communications (ICAC-INDST), Barcelona, Spain, June 2009.
- [5] S.Kandula, D.Katabi, B.Davie, A.Charny, "Walkingthetightrope: responsive yet stable traffic engineering," Proceedings of ACM SIGCOMM conference (SIGCOMM), USA, 2005.
- [6] S. Fischer, N. Kammenhuber, A. Feldmann, "Replex: dynamic traffic engineering based on wardrop routing policies," Proceedings of 2nd ACM CoNEXT conference (CoNEXT), Lisbon, Portugal, December 2006.
- [7] P. Psenak et al., "Multi-Topology (MT) Routing in OSPF," IETF RFC 4915, June 2007.
- [8] F. Francois, N.Wang, K. Moessner, S. Georgoulas, "Optimization for time-driven link sleeping reconfigurations in ISP networks," Proceedings of 13th IEEE/IFIP Network Operations and Management Symposium (NOMS), Hawaii, USA, April 2012.
- [9] W. Fisher, M. Suchara, J. Rexford, "Greening backbone networks: reducing energy consumption by shutting off cables in bundled links," Proceedings of 1st ACM SIGCOMM workshop on Green Networking (Green Networking), New Delhi, India, August 2010.
- [10] J. Chabarek, J. Sommers, P. Barford, C. Egan, D. Tsang, S. Wright, "Power awareness in network design and routing," Proceedings of 27th IEEE Conference on Computer Communications (INFOCOM), Arizona, USA, April 2008.
- [11] IEEE Standard 802.1AX: Link Aggregation, IEEE Computer Society, November 2008.
- [12] S. Nedeveschi, L. Popa, G. Iannaccone, S. Ratnasamy, D. Wetherall, "Reducing network energy consumption via sleeping and rate-adaptation," Proceedings of 5th USENIX Symposium on Networked Systems Design and Implementation (NSDI), California, USA, April 2008.
- [13] F. Idzikowski, S. Orłowski, C. Raack, H. Woesner, A. Wolisz, "Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios," Proceedings of 14th Conference on Optical Network Design and Modeling (ONDM), Kyoto, Japan, February 2010.
- [14] A. Coiro, F. Iervini, M. Listanti, "Distributed and adaptive interface switch off for Internet energy saving," Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN), Hawaii, USA, July 2011.
- [15] F. Cuomo, A. Cianfrani, M. Polverini, D. Mangione. "Network pruning for energy saving in the Internet," Elsevier Computer Networks, Vol. 56, No. 10, July 2012.
- [16] F. Giroire, D. Mazauric, J. Moulrierac, B. Onfroy, "Minimizing routing energy consumption: from theoretical to practical results," Proceedings of IEEE/ACM International Conference on Green Computing and Communications (GreenCom), Hangzhou, China, December 2010.
- [17] R. Bolla, R. Bruschi, A. Cianfrani, M. Listanti, "Enabling backbone networks to sleep," IEEE Network, Vol. 25, No. 2, March/April 2011.
- [18] L. Mamatas, S. Clayman, M. Charalambides, A. Galis, G. Pavlou, "Towards an information management overlay for the future Internet," Proceedings of 12th IEEE/IFIP Network Operations and Management Symposium (NOMS), Osaka, Japan, April 2010.
- [19] H. Asgari, P. Trimintzios, G. Pavlou, R. Egan, " Scalable monitoring support for resource management and service assurance," IEEE Network, Vol. 18, No. 6, pp. 6-18, November 2004.
- [20] R. G. Clegg, S. Clayman, G. Pavlou, L. Mamatas, A. Galis, "On the selection of management/monitoring nodes in highly dynamic networks," IEEE Transactions on Computers, Vol. PP, No. 99, March 2012.
- [21] M. Zhang, C. Yi, B. Liu, B. Zhang, "GreenTE: Power-aware traffic engineering," Proceedings of 18th IEEE International Conference on Network Protocols (ICNP), Kyoto, Japan, October 2010.
- [22] G. Athanasiou, K. Tsagkaris, P. Vlacheas, P. Demestichas, "Introducing energy-awareness in traffic engineering for future networks," Proceedings of the 7th International Conference on Network and Services Management (mini-CNSM), Paris, France, October 2011.
- [23] A. Kvalbein, O. Lysne, "How can Multi-Topology Routing be used for intradomain traffic engineering?," Proceedings of ACM SIGCOMM Workshop on Internet Network Management (INM), Kyoto, Japan, August 2007.
- [24] J.R. Loyola, J. Serrat, M. Charalambides, P. Flegkas, G. Pavlou, "A methodological approach towards the refinement problem in policy-based management systems," IEEE Communications, Vol. 44, No. 10, pp. 60-68, October 2006.
- [25] A. Bandara, E. Lupu, A. Russo, N. Dulay, M. Sloman, P. Flegkas, M. Charalambides, G. Pavlou, "Policy refinement for IP differentiated services quality of service management," IEEE Transactions on Network and Service Management (TNSM), Vol. 3, No. 2, pp. 2-13, second quarter 2006.
- [26] The Abilene network:
<http://www.cs.utexas.edu/~y Zhang/research/AbileneTM/>
- [27] The GEANT network: <http://www.dante.net/server/show/nav.007009007>