ABSTRACT

The energy consumption of the Internet is exploding due to the increase of the number of devices connected to it as well as the increase of traffic volume that is estimated in about 50% per year fuelled by mobile data and video applications; consequently, it is crucial to improve network efficiency to prevent the Internet from being throttled by an energy bottleneck. This PhD dissertation deals with the energy efficiency of telecommunication networks and specifically faces the problem of saving energy in the backbone section of the network. The main contributions concern the definition of algorithms and mechanisms that integrate and exploit the knowledge of the network topology, the carried traffic and the power behaviour of network devices to achieve energy efficiency in the network. Three novel contributions are presented which are related to three different saving strategies.

The first faced problem concerns the design of energy efficient virtual topologies in two-layers IP over WDM networks. The problem is formalized as a Mixed Integer Linear Programming (MILP) problem and a novel heuristic algorithm, named *Start-Single Hop and ReRoute*, is proposed and compared with the optimal solution and another heuristic solution proposed in the literature. Results obtained through extended simulations demonstrate that the proposed solution is able to reduce the network energy consumption with respect to other solutions and to perform close to the optimal solution.

Another key aspect faced in the dissertation is related to the possibility of saving energy in presence of variable traffic exploiting traffic engineering strategies that aim at aggregating the traffic on a subset of network elements in order to put in sleeping unused devices. Concerning this point, an Energy Aware Traffic Engineering mechanism, named DAISIES, is proposed. Besides achieving good level of energy efficiency, the proposed solution

overcome many practical issues rising from the adaptation of link switch-off mechanism, such as packet loss, out-of-order packet delivering and routing instabilities.

The last faced problem concerns the energy efficiency of the optical network layer. In this area a two different solutions are proposed. The first one is based on an iterative greedy algorithm that a posteriori tries to aggregate lightpaths on a subset of optical fibres links, re-optimizing the network when the traffic decreases. The second one, instead, directly works on lightpaths provisioning, exploiting a Power Aware Routing and Wavelength Assignment (PA-RWA) algorithm that tries to set up lightpaths in such a way that the total consumed energy is minimized. Specifically, the performance of the proposed PA-RWA algorithm is evaluated and compared with other solutions proposed in the literature showing better performance both in terms of energy efficiency and blocking probability.

In summary, this dissertation makes important contributions to the networking research community providing new methods and approaches for the definition of energy efficient networking techniques.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter I

INTRODUCTION

The widespread use of wideband technologies in customer access networks has made possible providing end users with high-capacity services. The growth of Internet users and high-bandwidth services have determined an increasing of the Internet traffic, which in turn has required the deployment of more transmission and switching equipments with more capacity. This has led to a rapid growth of the energy consumption of the network that has become the main operational expenditure of telcos.

Although, at today, the network is responsible of just 2% of the global electricity consumption [1], the expected traffic increase in coming years raises the issue of whether the energy consumption can be more constraining than the capacity of optical and electronic technologies for the growth of the Internet. Starting from the work in [2] the concept of Green Networking has been introduced in the telecommunication field indicating an umbrella of solutions aimed at reducing the energy required by the network to work. Several strategies can be undertaken to improve network energy efficiency; depending on the specific network section or network layer which we are focusing on, the specific approaches and solutions may be quite different with each other.

The present PhD thesis focuses on the backbone section of telecommunication networks which is expected to be the most energy-consuming section in coming years. The main contribution has concerned the definition of mechanisms, algorithms and procedures that integrate and exploit the knowledge of the network topology, the carried traffic and the power behaviour of network devices to achieve energy efficiency in the network.

I.1 Research Contribution

This dissertation presents three novel contributions in the area of energy efficiency in telecommunication backbone networks. Below, we briefly describe the contributions.

I.1.1 *Energy Efficient Design of Virtual Topology*

A key aspect of energy efficient backbone networks concerns the design of virtual topology in multi-layers networks; the problem is known as Energy Minimized Virtual Topology Design (EM-VTD). In IP over WDM networks the problem deals with designing the IP topology pursuing energy consumption minimization. The problem has been formalized as a Mixed Integer Linear Programming (MILP) problem and a novel heuristic algorithm has been proposed and compared with the optimal solution and another heuristic solution proposed in the literature. The main novelty of the proposed solution concerns the novel approach it follows and the use of a specific cost function inside the classical shortest path routing algorithm. The solution is described in IV.1.1.

I.1.2 *Energy Aware Traffic Engineering*

Such an energy saving strategy aims at adapting the energy consumption of the network to the actual offered load. The issue arises from the observation that the power consumption of network devices is quite independent of the actual supported load, making the power consumption of the network almost constant in time. Given that the offered traffic is quite variable in the network on a daily and weekly scale, the goal of Energy Aware Traffic Engineering (EA-TE) strategies is to make the power consumption of the network more load-proportional by aggregating traffic on a subset of network devices in order to put in low-power mode unused devices.

Although many solutions have been proposed in this field to efficiently determining the set of resources to keep active, a key point concerns the actual implementation of such solutions that in many cases do not consider the impact the switch-off of network elements has on the network. In this regard, the proposed solution provides a simple and efficient way for managing traffic in the network and switching-off idle devices, assuring no service interruption, preserving the good operation of layer 4 protocols and requiring minimal change to current routing and signalling protocols. The proposed solution, named DAISIES, is described in IV.1.2.

I.1.3 *Energy Efficient Optical Networks*

The last contribution concerns specific mechanisms to improve energy efficiency in transparent circuit-switched optical networks. Although the energy consumption of the optical network layer is much less than that of the IP layer, some studies point out that the grow rate of the former is expected to be higher than the latter's. For such a reason, in the present dissertation, energy saving strategies have been investigated also in the optical layer. Specifically, two different solutions have been proposed. The first solution is based on an iterative greedy algorithm that a posteriori tries to aggregate lightpaths on a subset of optical fibres links, re-optimizing the network when the traffic decreases. The second one, instead, directly works on lightpaths provisioning, exploiting a Power Aware Routing and Wavelength Assignment (PA-RWA) algorithm that tries to set up lightpaths in such a way that the total consumed energy is minimized.

The two proposed solutions are described in IV.2.1 and IV.2.2 respectively.

I.2 Organization

The rest of the thesis is organized as follows.

Chapter I provides a detailed analysis of the power consumption of Information and Communication Technology (ICT) and, in particular, of telecommunication networks by summarizing the most relevant studies in that field. Moreover, a power consumption model for the backbone network section is developed and some preliminary observations about the main goals that a generic energy saving strategy should pursue are carried out.

Chapter III introduces a first classification of energy saving strategies and focuses its attention on the energy saving opportunities in backbone networks. The energy minimization problem is mathematically formalized in III.2 considering the power consumption model provided in II.3. Then the problem is separately treated for the IP (III.3) and the optical (III.4) layers. As far as the IP layer is concerned, the two main strategies considered in this dissertation, namely EM-VTD and EA-TE, are respectively discussed in III.3.1 and III.3.2, and the most relevant solutions proposed in the literature are presented.

Chapter IV extensively describes the proposed solutions. Specifically, IP layer related solutions are presented in IV.1, whilst the ones related to the optical layer are described in IV.2.

Finally, Chapter V concludes the dissertation summarizing the main performed studies and the achieved results.

Chapter II
ENERGY CONSUMPTION IN TELECOMMUNICATION NETWORKS

II.1 Energy Consumption Trend in ICT

Information and Communication Technology (ICT) is widely used in our daily life and in most aspects of our society; it has always had an environment-friendly image and is often considered as a promising area for achieving energy conservation and reducing the environmental impact of our daily activities. Consider, for example, the various services that allow to reduce the human movement needed to perform oneself work, such as videoconference and telework, or also applications for e-commerce, e-booking, etc. Moreover, ICT is becoming very important to improve the energy efficiency in the energy production and distribution processes, as well as in improving energy-efficiency in all aspects of production and services. Ultimately, we can say that ICT has transformed our society providing practical means to reduce the human impact on nature.

However, an important issue arises from the ubiquitousness of ICT in daily life: the energy consumption of computers and network equipment is becoming a significant part of the global energy consumption [1], [3], [4]. As the coverage of ICT is rapidly spreading all over the world, the energy consumption of ICT is also increasing rapidly, since more equipment and components for networks and communications are being deployed annually. From the data of 2009, ICT consumes about 8% of the total electricity all over the world [5]. This has a significant impact on the environment in terms of CO2 emissions. It is estimated that today ICT accounts for about 2% of the worldwide carbon emissions and its carbon footprint will increase in next years more than other fields of applications. Moreover, as shown in

Figure II-1, most of its CO2 emissions are related to the operation of ICT devices rather than to the complete manufacturing process to produce ICT equipment.

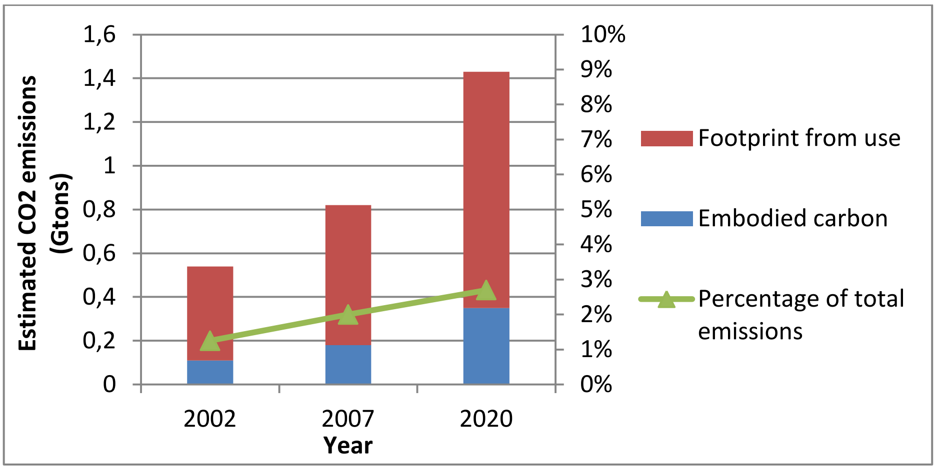


Figure II-1: Estimated CO2 emissions, source [10].

At today the main contributors of power consumption within the ICT are represented by consumer electronics like TVs, audio equipment, DVD players, gaming consoles, etc, that are responsible together of about a half of the total energy, whilst network equipment and data centers accounts for about 15% each. Nevertheless, the increase of the number of devices connected to the Internet as well as the increase of traffic volume, that is estimated in about 50% per year [6] fuelled by mobile data and video applications, is bringing an explosion of the energy consumption of the Internet. Figure II-2 shows forecasts of energy consumption of ICT equipment; it can be seen that, in coming years, the energy consumption of network equipment and data centers will increase more than consuming electronics' [1]. In fact, in the last years, broadband access technologies, such as Passive Optical Networks (PONs), ADSL2+, Long Term Evolution (LTE), etc., are rapidly spreading worldwide providing users with high-speed Internet connections and enabling the development of bandwidth-hungry services on the Web. This, jointly with the growth of costumer population, is leading to the deployment of an even larger number of devices in telcos and ISPs infrastructures, with sophisticated architectures able to perform increasingly complex operations in a scalable way.

For instance, high-end IP routers are even more based on multi-rack architectures with many line cards and switching matrixes, and provide more and more network functionalities. The capacity of such devices continues to grow, with an increase factor of 2.5 every 18 months [6]. At the same time, as shown in Figure II-3 and suggested by Dennard's scaling law [8], silicon technologies (e.g., CMOS) improve their energy efficiency with a lower rate with respect to routers' capacities and traffic volumes, by increasing of a factor 1.65 every 18 months (see Figure II-3).

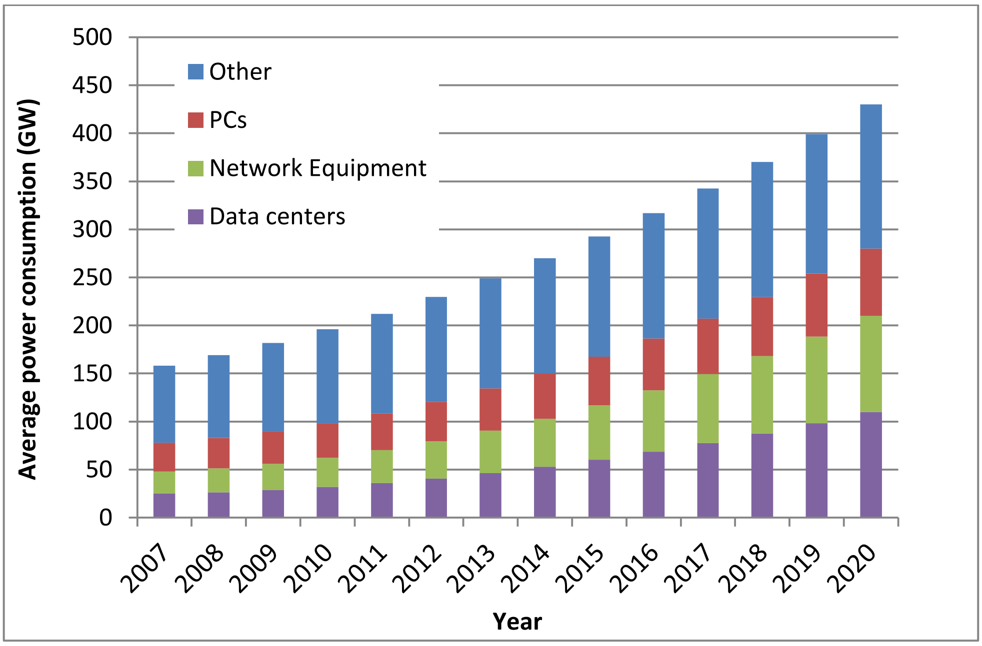


Figure II-2: Electricity consumption forecasts of ICT equipment during use, [1].

The sole introduction of low consumption silicon technologies cannot overcome such a problem and pursuing energy conservation in the future

Internet; it is crucial to improve network efficiency to prevent the Internet from being throttled by an energy bottleneck; specifically, the challenge is to achieve an increase of the technological and operational efficiency faster than the rate of traffic growth.

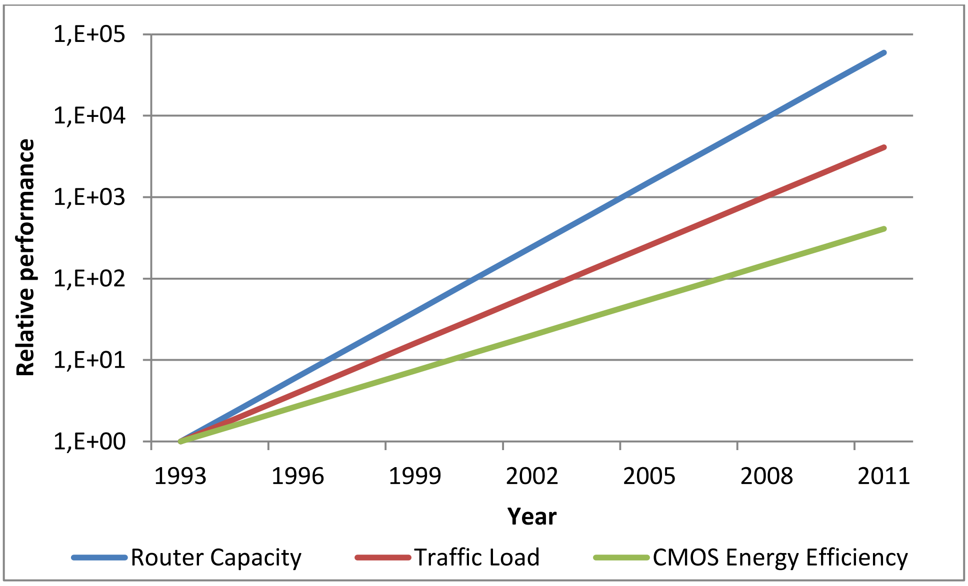


Figure II-3: Evolution from 1993 to 2010 of high-end router capacity (per rack), traffic volumes (Moore's law) and energy efficiency in silicon technologies; source [10].

Thus, as well as in other areas where energy efficiency is a concern, there are two main motivations to push ahead with the quest for "green" networking strategies:

1) the environmental one, which is related to the reduction of CO_2 emission;

2) the economic one, which is related to the reduction of costs sustained by the operators.

To this purpose, the research community has begun investigating new architectural solutions, algorithms and protocols, and developing innovative equipment with the aim of achieving a better ratio of performance to energy consumption. This has also driven major standardization bodies towards the definition of green solutions for data centers and network infrastructures [9].

As more energy efficient solutions are introduced in the networking field a considerable amount of energy can be saved leading to an important cost saving for telcos and ISPs. The paper in [10] reports a forecast of the saving in terms of energy and costs due to the application of green technologies in European telcos' infrastructure. Results presented in [10] and reported in Figure II-4 have been carried out by the European Commission DG INFSO [11].

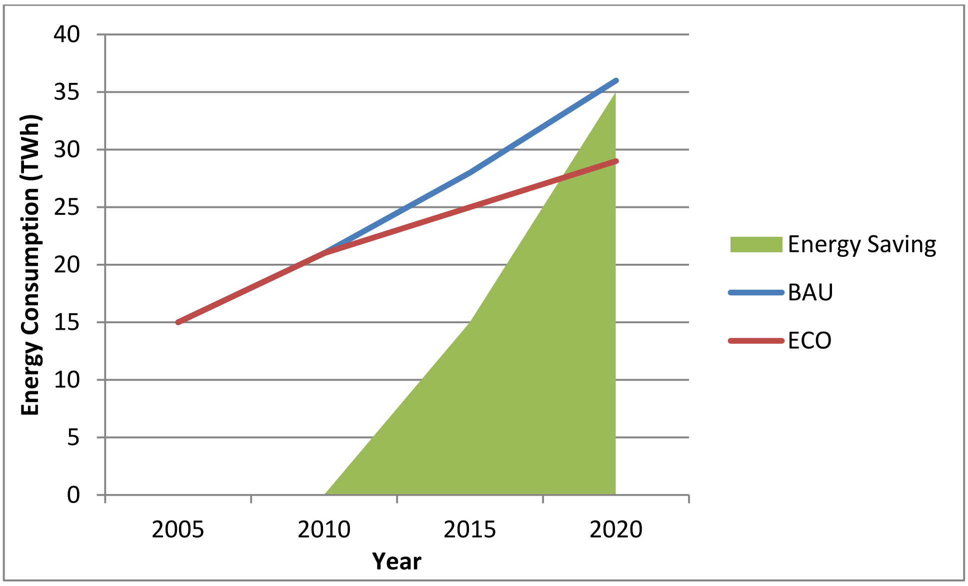


Figure II-4: Energy consumption estimation for the European telcos' network infrastructures in the "Business-As-Usual" (BAU) and in the Eco sustainable (ECO) scenarios, and cumulative energy savings between the two scenarios; source [10].

As it is evident, the expected saving is substantial both in terms of energy and costs. This is likely to lead to further investments for the definition of green networking technologies.

II.2 Energy Consumption of Different Network Sections

In order to develop appropriate green networking strategies, it has been necessary to carefully assess where the energy is most consumed in the network. In this regard, a great effort has been done by the research community in the last decade to identify the main contributors of power consumption and to develop proper consumption models in order to better address the green networking research.

Since telecommunication networks are typically composed of different sections, generally according to a hierarchical organization, it is particularly interesting evaluating the contribution of each section to the total energy consumption. In fact, each network section is generally based on a different technology, with a different structure, topology, communication protocols and transmission medium, and each of them requires a different strategy and approach. Concerning that aspect, a number of studies have been carried out with the objective of modelling a telecommunication network from the energy consumption perspective, mainly focusing on the whole network architecture rather than on the internal architecture of each device [12].

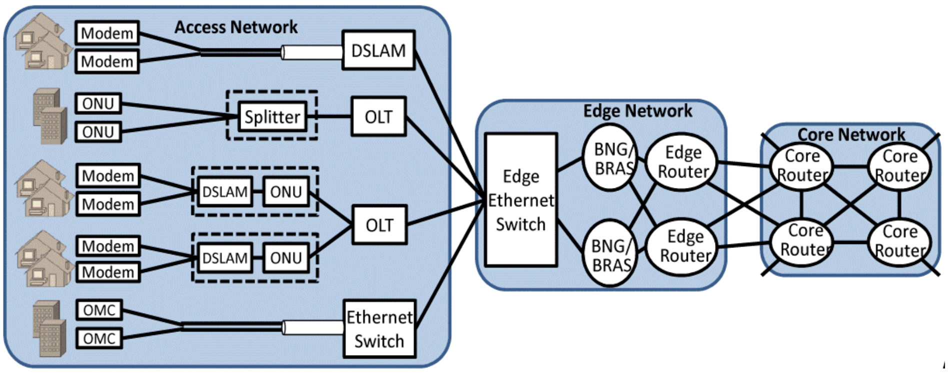


Figure II-5: Hierarchical architecture of a TLC network.

Figure II-5 shows typical network architecture composed of an access, a metro/edge and a core/backbone sections. The access network has the objective of connecting customers to the network; it typically has a tree topology where leaf nodes are the customer devices and the root node is the device in the central office, e.g. a DSLAM, an OLT or an Ethernet switch. Different technologies are available at today such as xDSL, PON or point-to-point optical access networks. The traffic coming from access sections is generally aggregated by the metro network by using layer 2 technologies such as Ethernet; they typically have a ring topology. Then the traffic reaches edge devices that generally are the first point where the IP protocol is managed in the network. Devices such as NAS and BRAS, representing the access points at the IP layer, are generally considered as belonging to the edge network. Finally, the aggregated traffic is routed in the core by IP routers that interconnect all the edge routers and provide a gateway towards the Internet. Links among IP core routers and between access and edge switches are generally realized by means of an underlying WDM optical transport network.

According to the study in [10], which results are shown in Figure II-6, at today about 70% of the total energy is consumed in the access section of the network due to the huge number of devices needed to connect all the customers distributed on the territory, whilst the transport and core networks account for 30% of the total. The typical power consumption of each device, instead, is much higher for core devices than for access ones due to the higher amount of treated traffic and the more complex processing that the formers must perform.

In this regard, an important aspect concerns how the energy required by each network section is expected to grow in coming years. Specifically, even if some countries are still developing their own network infrastructure and thus an increase of the number of Internet users is plausible in coming years, it is expected that the growth of the Internet traffic will be mainly due to the increase of the amount of traffic generated by each user. According to such a principle, authors of [12] show how the energy required by the network is expected to grow as the peak access rate (and consequently the average generated traffic) grows. The main assumption of the above-mentioned study can be summarized as follows:

i) the current equipment provisioning and used network design and operational practices are maintained along the whole period;

ii) according to technological evolution, an estimated average value of decrease of the energy per bit of networking equipment in the range of 15-20% per year is considered.

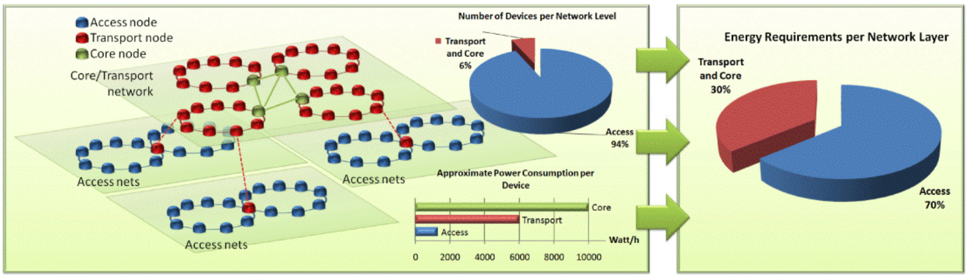


Figure II-6: Typical access, metro and core device density and energy requirements in today's typical networks deployed by telcos, and ensuing overall energy requirements of access and metro/core networks, source [10].

Results presented in Figure II-7 show that the energy consumption of access networks slightly depends on the amount of traffic generated by each customer ($A_I = A_p/M$) since it is mainly related to the number of access devices. On the contrary, the energy consumption of metro, edge and core devices linearly grows with A_I and is expected to exceed access ones' for A_I equal to about 10 Mbits/s; in particular, the energy consumed in the core grows more rapidly than the energy consumed by edge and transport devices. As wideband technologies are deployed in access networks, even more bandwidth hungry services will be developed on the Web and enjoyed via the public Internet, causing a growth of the internet traffic expected to be about 50% per year, as already outlined in II.1. On the basis of such results and considerations, the research community has been addressing a considerable effort towards the definition of energy saving mechanisms and strategies to reduce the energy footprint of the core network

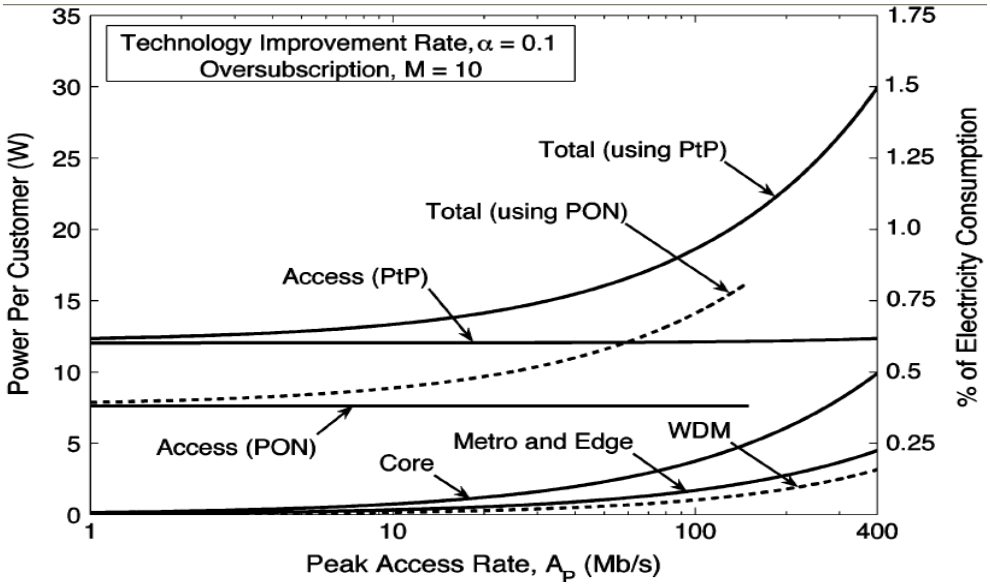


Figure II-7: Power consumption of the Internet with a PON or PtP network in the access network with no VDN and 10% router efficiency improvement per year. Includes power consumption of WDM links, core routers, metro and edge network, PON, ADSL and PtP access networks and total power consumption. Oversubscription M=10. Source [12].

II.3 Energy Consumption in the Backbone/Core Section

Core networks typically have the two layers architecture shown in Figure II-8. The higher layer is an IP network whilst the lower one is a Wavelength Division Multiplexing (WDM) optical layer. The former performs the switching operation in the electronic domain on a packet basis whilst the latter processes aggregated traffic flows directly in the optical domain, providing the IP layer with optical circuits, named lightpaths. Each lightpath constitutes a unidirectional link between two IP routers;

specifically, each lightpath connects a pair of physical line cards of the related routers. The set of all interconnected line cards constitutes the IP layer topology that is generally referred to as Virtual Topology. Since a lightpath can be potentially established between any two routers, the IP topology can be designed regardless of the WDM layer topology. The latter is generally referred to as Physical Topology.

The transmission capacity ΔC provided by each lightpath, and thus the amount of carried traffic, is equal to the line card bit-rate. Generally, each IP link can be composed of several aggregated line cards composing a single logical entity (virtual link). Thus, the capacity of a virtual link *i,j* in the IP topology can be modelled as:

$$C_{i,j} = x_{i,j} * \Delta C \qquad \text{(II.1)}$$

Wherein $x_{i,j}$ is the number of line cards deployed on the link *i,j*.

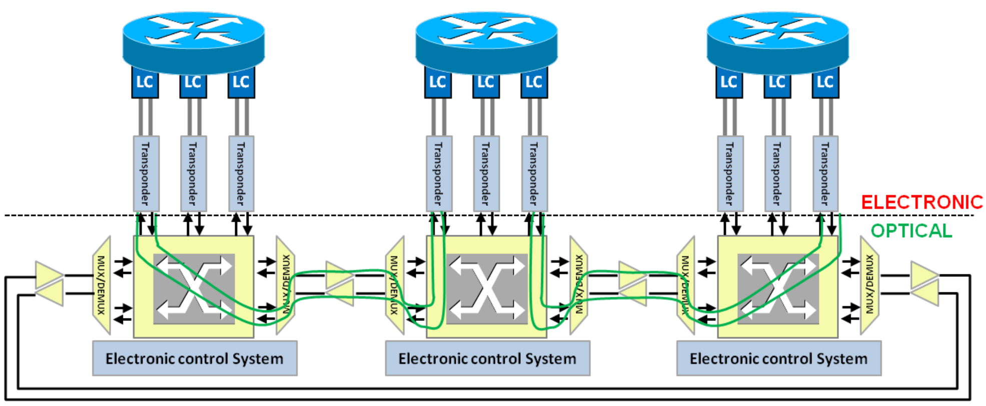


Figure II-8: Typical two-layers IP over WDM architecture of core/backbone networks.

Concerning the energy consumption of the core network it is necessary to distinguish between the contributions of the two layers. In fact, due to the different technology which they are based to, their energy consumption results quite different. Specifically, the IP network is a packet-switched network thus IP routers have to process traffic on a packet basis; at today such an operation can be performed only in the electronic domain making IP routers very energy hungry devices. On the contrary, the WDM optical network treats aggregated traffic flows and performs the routing operation on a wavelength basis directly in the optical domain. For such a reason, OXCs generally require much less energy than their IP counterparts. A careful evaluation of the two components of energy consumption is presented in [13], where a transparent optical network is considered; results are shown in Figure II-9.

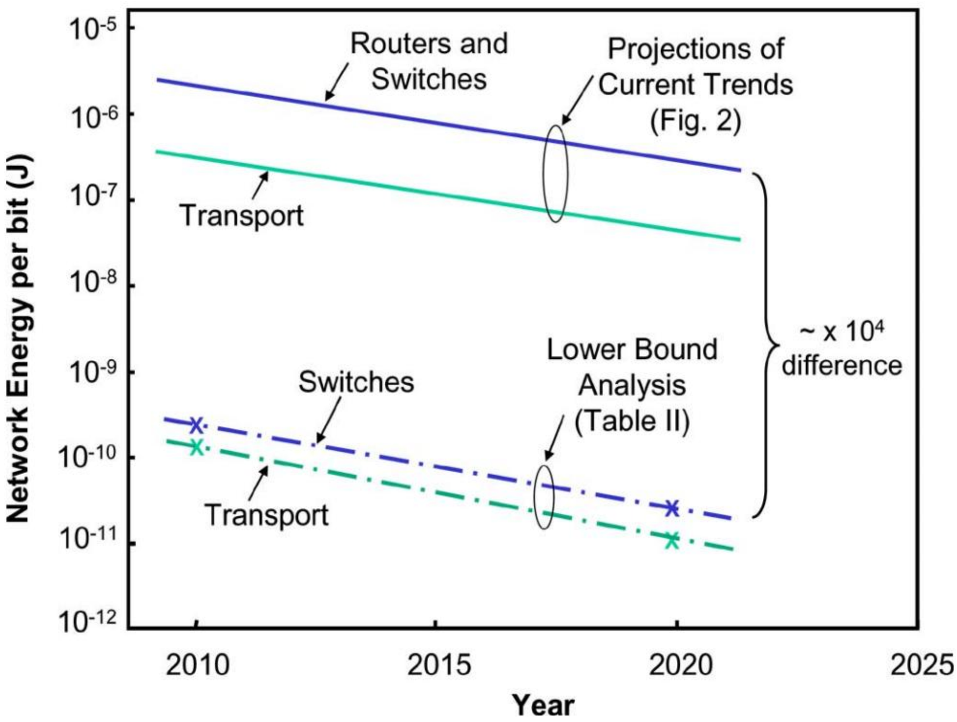


Figure II-9: End-to-end switch and transport network energy consumption. Source [13].

It can be seen that the energy required by the IP layer is expected to be about ten times higher than the energy required by the optical network layer.

This is mainly due to the lower energy consumption of optical devices with respect to electronic devices. In the next subsections a detailed power consumption model is presented for both the IP and the optical networks.

II.3.1 *Energy Consumption Model of the IP Network Layer*

With reference to the two-layer architecture shown in Figure II-8 the IP network is composed by IP routers interconnected by lightpaths established in the optical network. Thus its energy consumption is due to the energy consumed by each router which it is composed of.

A simple model of the architecture of an IP router is shown in Figure II-10.

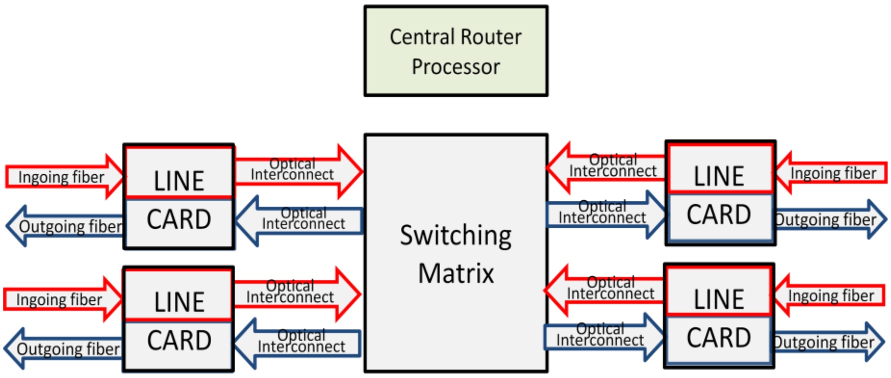


Figure II-10: Block diagram of high-end electronic IP router.

A router is mainly composed of:

- a set of line cards (LCs);
- one or more switching matrixes;
- a central processor.

The Line cards and the switching matrix are involved in data-plane operation.

Line cards represent the physical interfaces of the router; they are the end points of a link of the IP topology. The traffic entering in an ingoing interface that lies on a given line card is physically transferred on the proper outgoing interface, i.e. the proper line card, through the switching matrix whose role is to interconnect all the line card of the same router in order to allow traffic to be forwarded on the appropriate link. It can be seen that line cards are connected to the switching matrix by means of optical interconnections

The central router processor is responsible of all the operation related to the control plane. In particular the routing engine lies on this centralized entity. It must be outlined that the routing engine has the objective of constructing the routing table that is used by IP routers to perform the forwarding functionality, i.e. deciding the outgoing interface which a packet must be forwarded to. Thus it involves a routing protocol e.g. OSPF and a routing algorithm e.g. Dijkstra algorithm.

Each line card performs a number of operations not limited to the simple physical transmission. In particular, as shown in Figure II-11, the forwarding functionality is generally implemented directly in line cards. Each line card receives the routing table by the central router processor and makes decisions for ingoing packets. Once the outgoing interface has been decided, data are passed to the switching fabric interface module which interacts with the SW matrix control module and transfers data to the switching matrix.

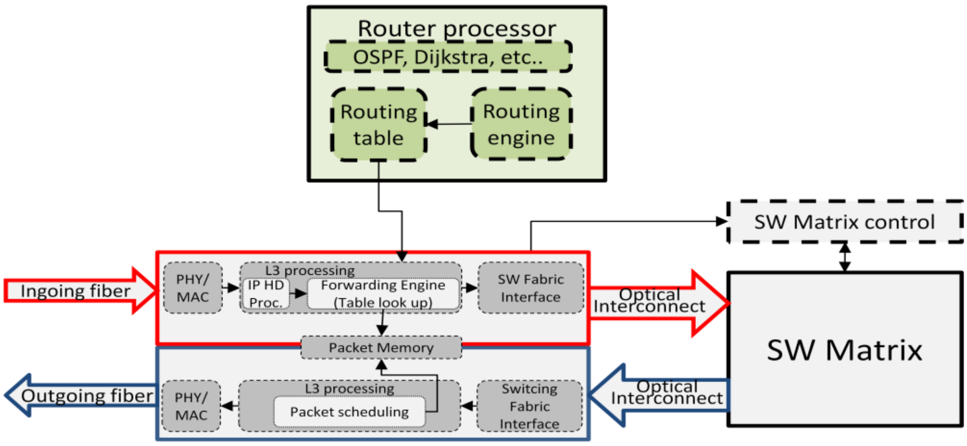


Figure II-11: Block diagram of high-end electronic IP router, showing main functionalities.

The papers in [14], [15], [16] present a careful analysis of the power consumption of the different hardware devices lying on an IP router and of the different functionalities implemented in each device. The results of these studies are summarized in Figure II-12 and Figure II-13. According to the study in [15], about 35% of the total energy is due to power supply inefficiencies and to fans and blowers needed to cool the router. This component of consumption can be seen as energy overhead and its total value is a fixed fraction of the total; this means that by reducing the total power absorbed by the router of a factor α such a component will be reduced of the same factor. More in detail:

$$P_{tot} = P_{CP\&DP} + P_{Cool\&PS} = P_{CP\&DP}*(1+H_{Cool\&PS}). \qquad (II.2)$$

$$H_{Cool\&PS} = P_{Cool\&PS} / P_{CP\&DP} \qquad (II.3)$$

In (II.2) P_{tot} is the total power consumption composed of two terms: i) the power related to control and data plane operations ($P_{CP\&DP}$); ii) the power related to power supply inefficiencies and cooling ($P_{Cool\&PS}$). The ratio $H_{Cool\&PS}$ between the two components is here assumed to be fixed.

Concerning the component $P_{CP\&DP}$ it can be divided as shown in Figure II-12.

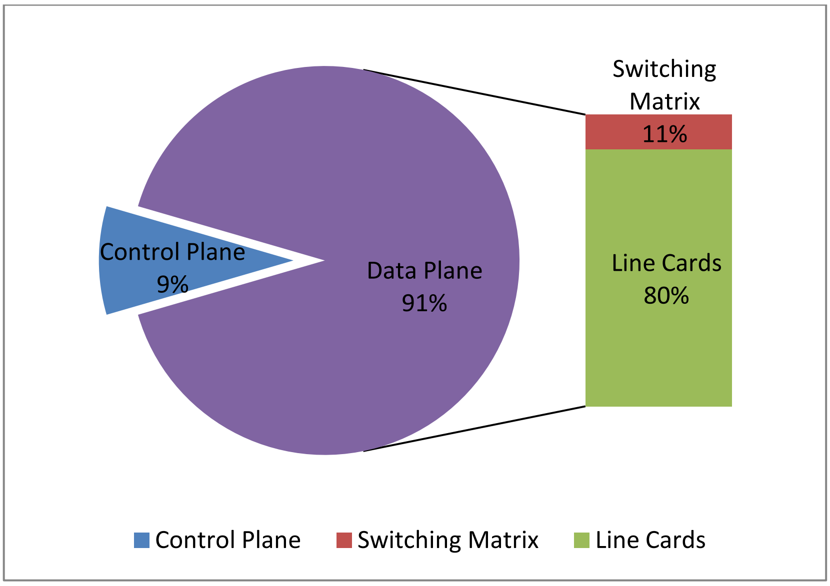


Figure II-12: Fractional energy consumed by different components of an IP router.

It can be seen that LCs are responsible of about 80% of the total power consumption whilst the remaining 20% is equally distributed between the switching matrixes and the control plane. Specifically, the study in [14] provides a detailed analysis of the different contributions of power consumption related to the different functionalities implemented in a LC. Results in Figure II-13 show that the most energy consuming functionality is the forwarding engine/packet processing functionality which consumes about 60% of the total power consumed by a LC. This is mainly due to the complexity related to the operations performed by such functionalities such as the longest prefix matching operation performed by the forwarding engine functionality of an IP router. It can also be seen that physical and data link functionalities account for just 13% of the total; this means that improving the power efficiency of modulation formats cannot be enough to improve the energy efficiency of the whole IP network.

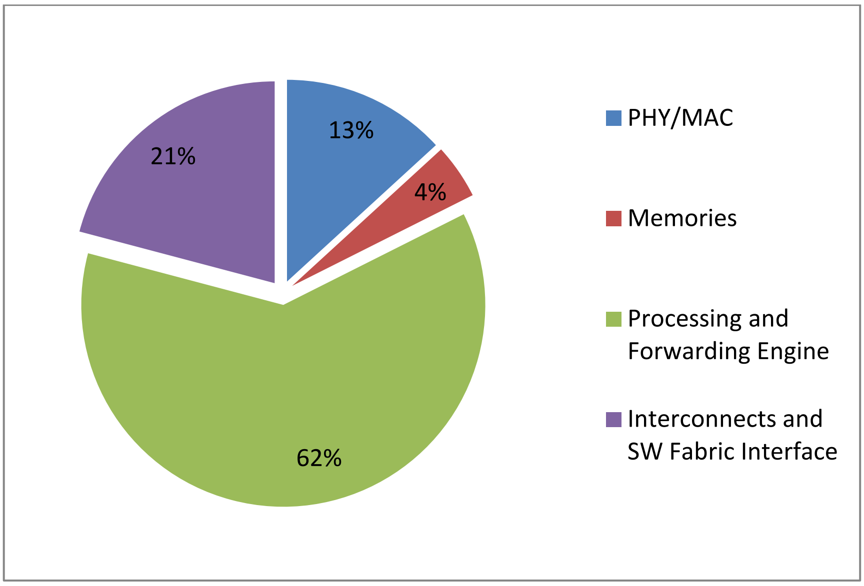


Figure II-13: Fractional energy consumed by different functionalities in Line Cards.

A key point concerns the relationship between the power consumed by each device and the load that it supports. Clearly, a full proportionality between power consumption and load would be desirable and would allow to automatically save energy during low traffic load periods; however, since devices are active also when idle, their power consumption results non null even when no operation is performed. A simple model of the power consumption of a generic device can be given as follows.

$$PC(x) = a_0 + f(x);$$
$$0 \leq x \leq 1. \quad \text{(II.4)}$$

In (II.4) the term $PC(x)$ is the power consumption of a generic device that depends on the normalized load x, whilst the term a_0 represents the power consumption when the device is idle. Thus, a very important aspect concerns the evaluation of the ratio:

$$\gamma = a_0/PC(1). \tag{II.5}$$

In this regard, Authors of [17] carried out an experimental evaluation of the power consumption of an IP router in different load conditions. The main results of the above cited study, reported in Figure II-14, can be summarized as follows.

1) The power consumption of an IP router is mainly due to the number of LCs installed on it.

2) The power consumption of an IP router slightly depends on the amount of treated traffic.

3) The power consumption of each LC slightly depends on the typology of treated traffic; specifically, for the same amount of treated traffic the power consumption mildly increases in case of many small packets with respect to the case of few big packets.

Concerning the last point, it can be observed that, even if the most power is related to the forwarding functionality, the number of times that such an operation is actually performed does not have a considerable impact on the consumption.

Results reported in this subsection give a general indication about the main strategy that should be followed to improve the energy efficiency of the IP network layer: minimizing the total number of LCs deployed in the network by properly designing the IP layer topology and routing the traffic.

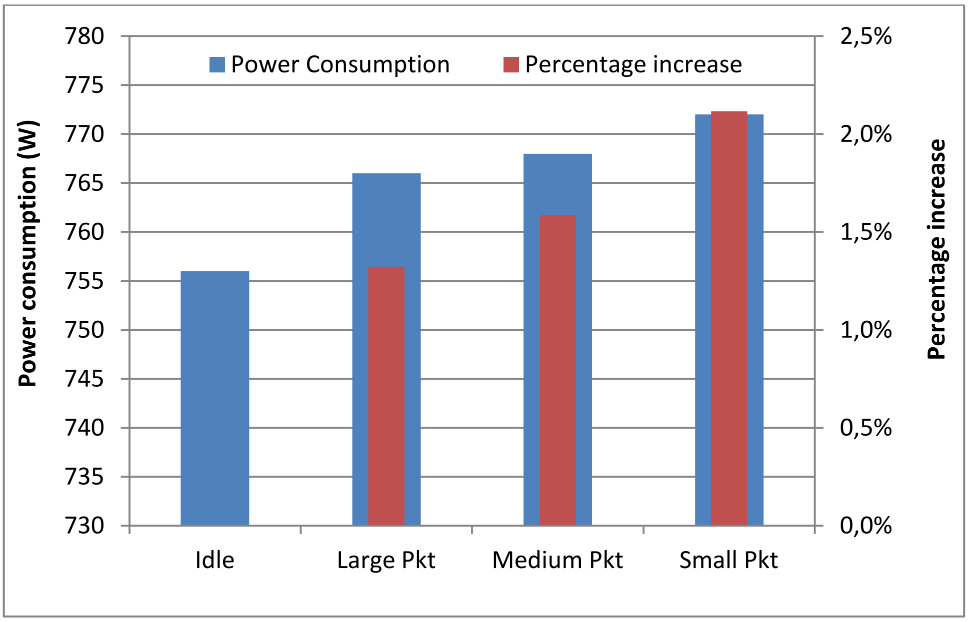


Figure II-14: Power consumption of an IP router for constant bit rate UDP traffic with different packet sizes at about 2.5 Gbits/s, along with idle baseline. Source [17]

II.3.2 *Energy Consumption Model of the WDM Optical Layer*

In the IP over WDM network scenario described in II.3 the optical network layer works as a circuit-switched transport network that provides the upper layer with connections representing the links of the virtual topology. The topology of the optical layer is often referred to as physical topology since links are realized on a real transmission medium. Specifically, a link is an optical fibre carrying a number of wavelength channels, each of which is used to accommodate a lightpath; the wavelength represents the minimum bandwidth granularity. It is very common to refer to such networks as wavelength routed networks since the task that each node has to perform is to switch the traffic carried by a wavelength on a given input port to the proper output port. It may be distinguished between transparent and opaque optical networks. The latter performs opticalelectrical-optical (O/E/O) conversion at each intermediate node via transponders deployed on OXC line interfaces, whilst the former does not. The advantage of an opaque network is to provide fully wavelength conversion allowing a better utilization of the huge capacity provided by optical fibres; nevertheless, it requires a considerable number of transponders leading to a considerable energy waste. If H is the number of hops in the physical topology, $2 \cdot H$ transponders are needed for each bidirectional connection (a pair of lightpaths) which the IP layer is provided with. In a transparent optical network the O/E/O conversion is performed only on client interfaces where client signals are inserted in or extracted from the optical network. Thus, the number of needed transponders is equal to the number of client interfaces. Transponders on client interfaces are necessary to convert gray signals transmitted by client devices on appropriate wavelengths and to provide the channel coding and the modulation suitable for the long-haul WDM transmission. Alternatively, by providing IP routers with tuneable laser interfaces, transponders are not needed at all and the power efficiency of the whole network can be further improved.

The general architecture of a transparent optical network is depicted in Figure II-15. It is composed of:

- an electronic control system (ECS);
- a 3D Micro Electromechanical System (MEMS) based optical switching matrix, that might be of the type proposed in [18];
- a set of transponders on client interfaces;
- a pair of passive optical MUX/DEMUX for each line interface composed of a couple of input/output fibres.

The total power consumption of a node i can be modelled as follows.

$$P_i^{Node}\,(N_i, I_i, W) = P^{ECS} + I_i{*}P^{Trans} + (N_i{*}W + I_i){*}\mathrm{P}^{\mathrm{MEMS}}. \qquad \text{(II.6)}$$

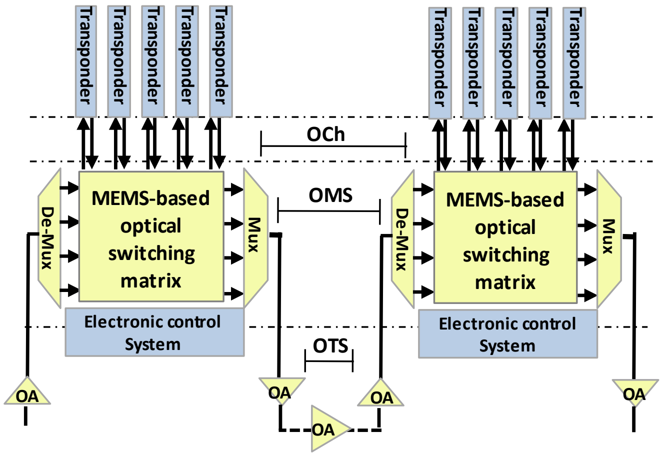


Figure II-15: Architecture and components of transparent circuit-switched optical networks.

In (II.6), P_i^{Node} is the total power consumption of the node *i*; *Ni* is the number of ingoing/outgoing fibres pair; *Ii* is the number of client interfaces and *W* is the number of wavelengths of each fibre. P^{ECS} is the power absorbed by the ECS; P^{MEMS} is the power absorbed by a single element of the node switching matrix; finally, P^{Trans} is the power absorbed by a transponder associated to each client interface. It has to be outlined that the power consumption of 3D MEMS-based switching matrixes is due to electronic circuitries that are responsible for the correct space positioning of the mirrors; since two mirrors, and an electronic circuitry, are placed for each input/output port pair, the power consumption linearly grows with them.

Each optical link is composed of a number of optical fibre spans; pre and post Erbium Doped Fibre Amplifiers (EDFAs) are deployed at the two ends of each fibre and a number of In-Line Amplifiers (ILAs) are placed along it. So, the power consumption $P_{i,j}^{Fibre}$ associated to each fibre of the link i,j can be written as:

$$P_{i,j}^{Fibre}(L_{i,j}) = \lfloor L_{i,j} / d^{amp} \rfloor * P^{ILA} + (P^{Pre} + P^{Post}). \quad \text{(II.7)}$$

In (II.7) $L_{i,j}$ is the physical distance between nodes i and j and d^{amp} the distance between two ILAs; P^{ILA}, P^{Pre} and P^{Post} are the powers absorbed by the ILA, pre and post amplifiers, respectively.

A preliminary study performed during the present PhD thesis has concerned the evaluation of the power consumption of the different devices composing the described network architecture. Results presented in [54] and shown in Figure II-16 have been carried out considering a network extending on An area of 4000x4000 Km^2 and assuming classical power-unaware routing and wavelength assignment algorithms.

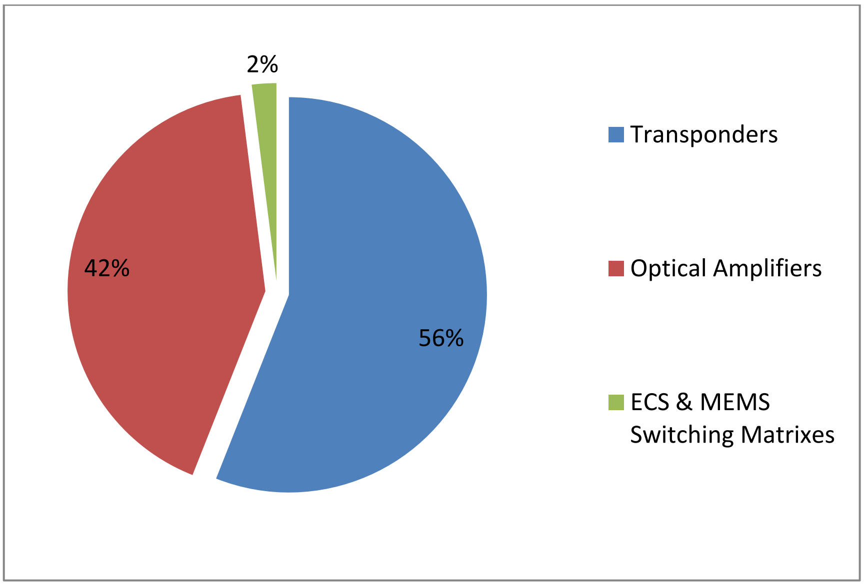


Figure II-16: Power consumption of different network devices using Least Congested Path | First Fit RWA algorithm.

As it is evident, most of the power is consumed by EDFAs and transponders, about 42% and 56% respectively, with optical switching matrixes and electronic control systems consuming together less than 2% of power. However, since the number of transponder just depends on the number of line cards (corresponding to the number of requested lightpaths) deployed at the IP layer, the only saving opportunity concerns the possibility of reducing the number of needed optical amplifiers.

Chapter III

ENERGY EFFICIENCY IN TLC NETWORKS: PROBLEM STATEMENT AND STATE OF THE ART

III.1 Energy Saving Strategies

The process of “greening” of the Internet, as firstly outlined by the pioneeristic work on such topic [2], passes through the individuation of the main wastes experienced in a TLC network. Concerning this, it is possible to identify two main opportunities for improving the energy efficiency of the network. First of all, current networks are generally designed with the target of minimizing the capital expenditure (CAPEX) without taking into account their energy consumption that represents instead the major contributor of operational expenditure (OPEX). By pursuing energy efficiency during the network design process, the power consumption of the network can be drastically reduced.

In this regard, the Energy-Minimized Network Design (EM-ND) is one of the main strategies able to achieve a considerable energy saving in multilayer networks; such a problem will be investigated in III.3.1. It takes in input a traffic matrix, representing the traffic demand among IP layer nodes, the power consumption model of each network device and the physical topology, and provides the IP layer topology (the virtual topology), the routing of the traffic matrix on the IP topology and the routing of the virtual topology, represented by the set of lightpaths, on the physical topology exploiting proper traffic grooming strategy aiming at minimizing the total power consumed by the network.

However, it is worth remembering that, at present, networks are dimensioned statically, based on a peak traffic load plus a reserve; this peak capacity is directly connected with the network peak energy consumption. Moreover the traffic presents significant periodic fluctuations over day/month time basis, e.g., during night hours, as shown in Figure III-1 where the traffic trend over an Italian ISP backbone link [19] is reported. It can be seen that the average link load is less than 50% of the peak load, which is experienced only during some days and for short periods.

Unfortunately this traffic drop is not followed by a corresponding decrease in the amount of energy consumed by the network. This is mainly due to the mild proportionality between load and power consumption of network devices, as outlined in II.3.1. Thus, the overall objective is to achieve a near load proportional power profile of the network in order to save a considerable amount of energy during periods of low traffic load. We refer to such a class of strategies as Energy Minimized Network Operation (EM-NO).

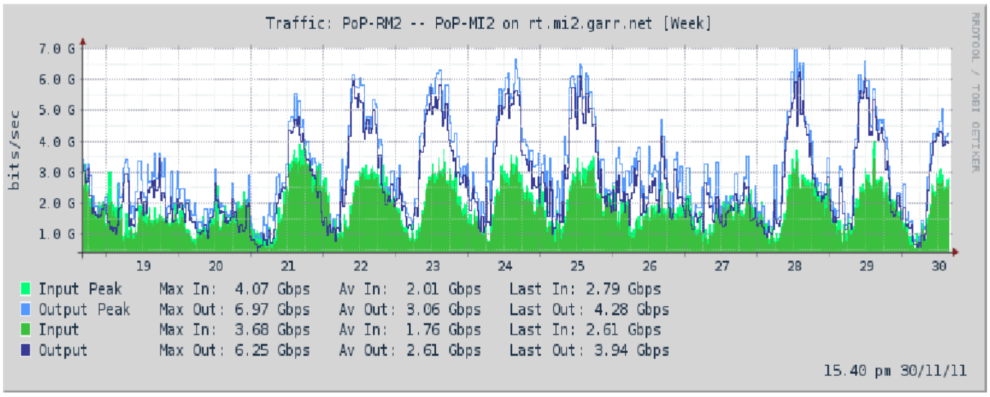


Figure III-1: Traffic statistics measured over the backbone link of the Italian GARR network.

Concerning the last point, among the possible approaches that can be used are the following.

- **Dynamic adaptation:** at an individual device level, i.e. switch or router, it is possible to put in low power mode some of the subcomponents such as line cards when they are idle or to clock the hardware at lower rates. This will lead to a more proportional power profile of each single sub-component, and consequently of the whole network device, that will come transparently with respect to other network devices.

- **Traffic and topology control:** power saving mechanisms based on traffic and topology control are currently founded on the extension of traffic engineering and routing criteria to use [20]. In this area, the basic idea is to adapt the network capacity in terms of links and nodes to the actual traffic volumes. The main objective of is to move traffic flows among network nodes in order to find the minimum number of network resources (i.e., links and nodes), guaranteeing the best trade-off between end-to-end network performance and overall power consumption.

The first approach is suitable for electronic devices operating on a packet basis and can exploit different techniques such as Dynamic Voltage Scaling and Dynamic Frequency Scaling, or low power idle modes such as Energy Efficient Ethernet [10]; it results in lowering the parameter a_0 defined in (II.5) and making the related function $f(x)$ almost linear. Such an approach does not need coordination among network devices but needs a careful evaluation about the energy overhead during power state transitions and about the impact on the traffic.

Figure III-2 shows the impact that performance scaling and idle logic technique have on the traffic. It can be seen that both such classes of techniques determine an additional delay on the traffic. Performance scaling (Figure III-2-c) obviously causes a stretching of packet service times (i.e., header processing time in a processing engine, or packet transmission time in a link interface), while the sole adoption of idle logic (Figure III-2-b) introduces an additional delay in packet service, due to the wake-up times. The objective of making the wake-up time as short as possible is in contrast to the objective of making the idle power as low as possible; in other words, in order to allow a quick wake up of the component some functionalities must remain active leading to higher idle power. Moreover, an optimization policy is generally needed to configure and control the usage of energy-aware capabilities and states with respect to the estimated workload and service requirements. The implementation of such methods might require a significant computation leading to additional energy consumption.

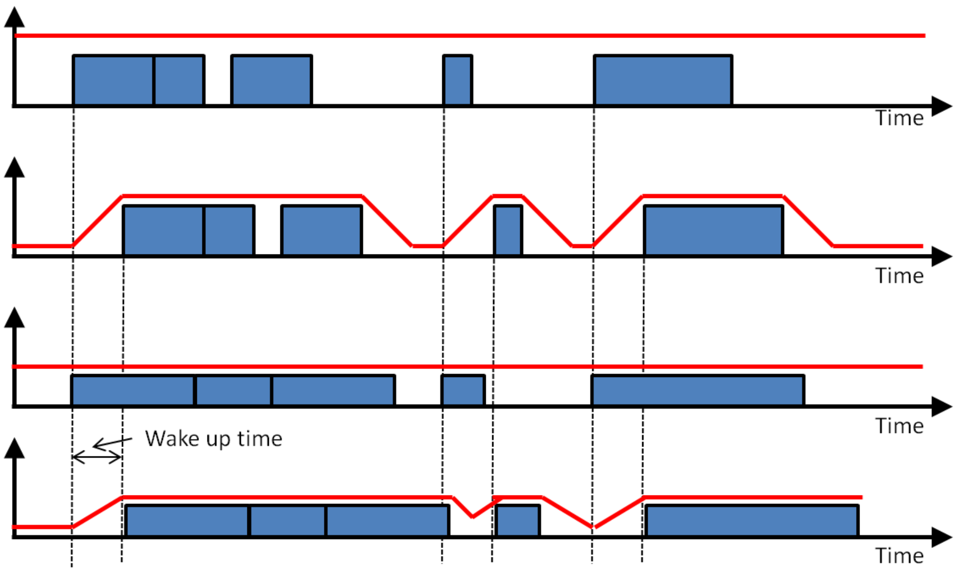


Figure III-2: Packet service times and power consumptions in the following cases: (a) no power-aware optimizations, (b) only idle logic, (c) only performance scaling, (d) performance scaling and idle logic. Source [10].

Traffic and topology control techniques aims at designing appropriate strategies, protocols and algorithms, integrating and exploiting the knowledge of the network topology, the carried traffic and the power behaviour of network devices. They can exploit sleeping/standby primitives, which allow devices turning part of them almost completely off, and entering very low energy states, while all their functionalities are frozen. Thus, sleeping/standby states can be thought as deeper idle states, characterized by higher energy savings and much larger wake-up times.

Related solutions generally require coordination among network devices in order to avoid service interruption when devices enter standby states. Specifically, this requires changes to the manner in which layer 2 and layer 3 protocols work. Furthermore, sleeping may affect layer 4 protocols such as TCP and we need to study this impact and develop appropriate solutions.

The main advantage of such an approach concerns the potential efficiency they lead to; in fact, if efficient mechanisms are adopted, it is possible to make the power consumption of the network almost proportional to the actual carried load. This is possible due to the very low power consumption of devices in sleeping mode and making full use of devices kept active. On the other hand, if appropriate mechanisms are defined, such an improvement in energy efficiency can be achieved without impairing network performance.

The main focus of the present thesis concerns:

1) **The EM-ND problem:** as above outlined such a problem deals with designing the virtual topology and mapping it on the physical topology, considering the peak traffic demand. It has the objective of reducing the peak network's power consumption.

2) **Traffic and topology control:** as above outlined such a strategy deals with making the power consumption of the network as much proportional to the carried traffic as possible, exploiting sleeping

capabilities of network devices and defining appropriate mechanisms, algorithms and protocol extensions.

The two identified strategies generally involve both the IP and the WDM optical layers. However, due to the different technology used at the two layers (the first operates in the electronic domain whilst the second in the optical one) and the different network service they provide (the IP layer is packet switched whilst the WDM optical network is a circuit switched layer) the related specific mechanisms result quite different. For such a reason, they will be separately treated in III.3 and III.4.

The next section, instead, gives an overview of the power consumption minimization problem in IP over WDM backbone networks and provides a mathematical formulation of the problem.

III.2 Energy Efficiency in IP over WDM Backbone Networks: Problem Statement

Let us consider the IP over WDM backbone network architecture described in II.3 and the related power consumption model. The power consumption minimization problem can be formulated as a MILP problem wherein the objective function accounts for the power consumption of each device deployed in the network. The inputs of the problem are instead:

i) **The traffic matrix:** it represents peak traffic relations, at the IP layer, among core nodes.

ii) **The physical topology:** the set of nodes and links as well as all devices deployed at the WDM optical layer.

iii) **The network model:** it includes the power consumption model as well as the capacity of each device.

The main decision variables are:

i) **The set of deployed devices at the IP layer:** it determines the virtual topology.

ii) **The routing of the traffic matrix on the virtual topology.**

iii) **The routing of the virtual topology on the physical topology.**

Concerning the WDM optical layer, in II.3.2 we outlined that the power consumption related to optical switching matrixes is negligible with respect to the total; thus, for sake of simplicity, it can be neglected in the problem. Instead, concerning transponders, because a transponder is needed for each line card deployed at the IP layer, their power consumption can be considered as an additional fixed term of the power of each line card. Finally, the power related to the electronic control system of each OXC must be treated as a fixed term that cannot be eliminated since each node must remain active in order to allow managing it. Thus, the only term of power consumption considered in the following is the power related to optical amplifiers.

Concerning IP routers, the fixed term of power consumption is represented by the power related to control plane operations (see Figure II-12) which, as for OXCs cannot be eliminated; what can be minimized, instead, concerns the power consumption of line cards and switching matrixes. As stressed in II.3.1, the power consumption of each device can be modeled as composed of a fixed term plus a variable load-dependent term (see eq. II.5). However, because current electronic devices show γ parameters close to one (see eq. II.5), the variable term can be neglected, thus considering that each device consumes always the same amount of power regardless of the actual supported load.

Thus, let us consider the following notation.

Indices

l and m	Indices of nodes in the physical topology G; the pair l,m represents a unidirectional link connecting two adjacent nodes.
k and λ	Indices of fibre of a link and of wavelength of a fibre, respectively.
i and j	Indices of the nodes in the virtual topology (IP layer). A virtual link (lightpath) connects two such end nodes. Physically, the nodes can be a pair of IP routers connected by the virtual link.
s and d	Indices of source and destination node of traffic flows in the traffic demand.

Sets and parameters

$G = (V,E)$	The physical topology, where V is the set of nodes and E the set of unidirectional links. $N = \|V\|$ and $L = \|E\|$ are the total number of nodes and links in the network respectively.
$K_{l,m}$	The set of available fibres on the link l,m. $F_{l,m}=\|K_{l,m}\|$ is the number of fibres deployed on the link l,m.
Λ	The set of wavelengths; $W = \|\Lambda\|$ is the number of wavelengths of each fibre.
$L_{l,m}$	The physical length of the link l,m.
$N_{l,m}$	The number of ILAs of each fibre of the link l,m. $N_{l,m}=\lfloor L_{l,m}/d^{amp}\rfloor$ where d^{amp} is the physical distance between two ILAs.
P^{ILA} and P^{E}	The power consumption of each ILA and the power consumption of each post-pre amplifier pair respectively.
$[T_{s,d}]$	The traffic demand indicating the amount $T_{s,d}$ of traffic requested between node s and d.

ΔC_{LC}	The capacity of each line card. It represents the amount of traffic carried by each wavelength.
ΔC_{SW}	The switching capacity of each switching matrix.
P_{LC}	The power consumption of each line card
P_{SW}	The power consumption of each switching matrix
LC_{MAX}	Maximum number of line cards that can be installed on an IP router.
SW_{MAX}	Maximum number of switching matrixes that can be installed on an IP router.

Variables

$f_{i,j}^{s,d}$	A binary variable equal to 1 if the virtual link *i,j* is used to transport the traffic flow between node *s* and *d*, otherwise is equal to 0.
$x_{i,j}$	An integer variable indicating the number of line cards on the link *i,j*.
y_i	An integer variable indicating the number of switching matrix on the node *i*.
$\varphi_{l,m,k,\lambda,}^{i,j}$	A binary variable equal to 1 if the wavelength λ of the fibre *k* of the link *l,m* is used to accommodate a lightpath between nodes *i* and *j*, otherwise 0.
$\varphi_{\lambda}^{i,j}$	An integer variable indicating the number of lightpaths between node *i* and *j* using the wavelength λ.
$\xi_{l,m,k}$	A binary variable equal to 1 if the fibre *k* of the link *l,m* is used to route at least one lightpath, otherwise 0.

The notation above leads to the following ILP formulation.

Minimize

$$P_{tot} = P_{IP} + P_{WDM} \qquad \text{(III.1)}$$

$$P_{IP} = \sum_{i \in V}\left(\sum_{j \in V}\left(x_{i,j} * P_{LC}\right) + y_i * P_{SW}\right) \quad \text{(III.2)}$$

$$P_{WDM} = \sum_{l \in V} \sum_{m \in V} \sum_{k \in K_{l,m}} \left(\xi_{l,m,k} * \left(N_{l,m} * P_{ILA} + P_E\right)\right) \quad \text{(III.3)}$$

Eq. III.1 is the objective function representing the total power consumption; it is composed of two terms: i) the term P_{IP} is related to the IP layer and is specified in III.2; it accounts for the power related to all line cards and switching matrixes in the network; ii) the term P_{WDM} is related to the optical layer and is specified in III.3; it accounts for the power related to all optical amplifier deployed along used fibres, i.e. fibres for which the variable $\xi_{l,m,k}$ is equal to 1.

Subject to

$$\sum_i (f_{i,j}^{s,d} - f_{j,i}^{s,d}) = \begin{cases} -1; \text{ if } j = s \\ 1; \text{if } j = d \\ 0; \text{if } j \neq s, d \end{cases} \quad \text{(III.4)}$$

Eq. III.4 are the classical flow constraints; it can be noted that, because $f_{i,j}^{s,d}$ are binary variables, each traffic flow *s*,*d* can be transported on only one path.

$$\sum_{s,d}\left(f_{i,j}^{s,d} * T_{s,d}\right) \leq x_{i,j} * \Delta C_{LC} \quad \text{(III.5)}$$

Eq. III.5 constrains the total traffic transported on a virtual link *i*,*j* to be less than the capacity deployed on that link that is given by the number $x_{i,j}$ of line cards deployed on it.

$$x_{i,j} = x_{j,i} \quad \text{(III.6)}$$

Eq. III.6 simply says that each line card is bidirectional.

$$\sum_{j,s,d}\left(\left(f_{i,j}^{s,d} + f_{j,i}^{s,d}\right) * T_{s,d}\right) \leq \ y_i * \Delta C_{SW} \tag{III.7}$$

Eq. III.7 constrains the total traffic switched by a node *i* to be less than the switching capacity given by the number of switching matrixes installed on it.

$$\sum_j\left(x_{i,j}\right) \leq LC_{MAX} \tag{III.8}$$

$$y_i \leq SW_{MAX} \tag{III.9}$$

Eq. III.8 and eq. III.9 respectively limit the total number of line cards and switching matrixes installed on a node.

$$\sum_\lambda\left(\varphi_\lambda^{i,j}\right) = x_{i,j} \tag{III.10}$$

$$\sum_{l,k}(\varphi_{l,m,k,\lambda}^{i,j} - \varphi_{m,l,k,\lambda}^{i,j}) = \begin{cases} -\varphi_\lambda^{i,j};\ \text{if } m = i \\ \varphi_\lambda^{i,j}; \text{if } m = j \\ 0; \text{if } m \neq i,j \end{cases} \tag{III.11}$$

Eq. III.10 constrains the total number of lightpaths from a node *i* to a node *j* to be equal to the number of line cards deployed on the related virtual link in the IP topology; this assures the actual mapping of the virtual topology on the physical topology. Eq. III.11, instead, represents the flow constraints for the optical network assuming no wavelength conversion functionality.

$$\sum_{i,j}\left(\varphi_{l,m,k,\lambda}^{i,j}\right) \leq 1 \tag{III.12}$$

$$\sum_{\lambda}\left(\varphi_{l,m,k,\lambda}^{i,j}\right) \leq M * \xi_{i,j,k} \tag{III.13}$$

Eq. III.12 constrains each wavelength channel on each fibre link to be assigned to at most one lightpath, whilst eq. III.13 constrains a fibre to be considered as used, and thus optical amplifier deployed along it to be accounted in the objective function, if at least one wavelength channel is used on it. This is assured by setting the constant M to a large enough value.

$$\sum_{\lambda,i,j}\left(\varphi_{l,m,k,\lambda}^{i,j}\right) \leq W \tag{III.14}$$

$$\sum_{k}\left(\xi_{l,m,k}\right) \leq F_{l,m} \tag{III.15}$$

Finally, eq. III.14 and eq. III.15 respectively limit the total number of wavelength used on a fibre and the number of fibre used on a link.

The formulation above leads to an NP-hard problem making the problem intractable for large networks. For such a reason, heuristic approaches are generally pursued to solve the problem in real networks. The mathematical formulation provided above, instead, is very useful to provide an absolute comparison parameter for evaluating the performance of a heuristic solution and it is used with this purpose in the present dissertation.

The power consumption minimization problem formalized so far considers the power consumption of both the IP and the optical layers; however, in many cases, the two network layers are administrated by different operators, each of which is interested in minimizing its own OPEX. Thus, in most cases, it is useful to divide the whole problem in two sub-problems, one related to the IP layer and one to the optical layer. The former aims at minimizing the term P_{IP} and does not consider the actual mapping of the IP topology on the physical topology. Thus it assumes that enough resources are available at the optical layer to establish all lightpaths requested by the IP layer. The sole input of the IP-related sub-problem is then the traffic matrix, whilst the optical topology is neglected. The WDM-related sub-problem, instead, takes in input the set of lightpaths requested by the IP layer, i.e. the set of $x_{i,j}$ variables, and deals with routing and assigning a wavelength to each of them.

Clearly, dividing the problem generally leads to a suboptimal solution. Specifically, let P_{ToT}^{Opt} be the power consumption given by the optimal solution of the original problem and $P_{ToT}^{Sub\_Opt}$ the total power consumption when the problem is divided in the two sub-problems:

$$P_{ToT}^{Opt} = P_{IP}^{Opt} + P_{WDM}^{Opt} \qquad \text{(III.16)}$$

$$P_{ToT}^{Sub\_Opt} = P_{IP}^{Sub\_Opt} + P_{WDM}^{Sub\_Opt} \qquad \text{(III.17)}$$

It results:

$$P_{ToT}^{Opt} \leq P_{ToT}^{Sub\_Opt} \qquad \text{(III.18)}$$

However, since the IP-related sub-problem does not depend on the optical layer it must results:

$$P_{IP}^{Sub\_Opt} \leq P_{IP}^{Opt} \qquad \text{(III.19)}$$

Then combining eq. III.19, eq. III.18, eq. III.17 and eq. III.16, it results

$$P_{WDM}^{Opt} \leq P_{WDM}^{Sub\_Opt} \qquad \text{(III.20)}$$

This means that the potential increase of power consumption due to treating the two sub-problems separately is related to the increase of the power consumed at the optical layer. This might occurs because the IP-related sub-problem does not consider such a power consumption term. However, such a potential sub-optimality is expected to be quite negligible for the following reasons. First, the power consumption related to optical amplifiers is estimated to be about 5% of the total power consumed by the core network (see II.3 and II.3.2). Moreover, since the power consumption of the IP layer is mainly related to the number of line cards, minimizing it means minimizing the total number of requested lightpaths and the amount of traffic offered at the optical layer. For such a reason, even in those cases where the same operator administrates both the network layers, treating the two sub-problems separately allows to make the whole problem easier to solve without moving away from the optimal solution.

The IP-related and the WDM-related sub-problems will be respectively discussed in III.3 and III.4 and the main solutions proposed in the literature will be described.

The energy consumption minimization problem formalized above can be applied to both the EM-ND and the EM-NO problems defined in III.1. Specifically, when the EM-ND problem is faced, the traffic matrix considered as input of the problem represents the peak traffic demand among IP routers. In this case the objective is to minimize the peak power consumption of the network and the applied strategy can be considered as a long period solution. In this case the problem can be solved off-line and the time needed to obtain the solution is not a stringent constraint; this would even allows to actually solve the MILP problem if the number of network nodes is not very large (generally limited to few tens of nodes). Moreover, since such a procedure is seldom performed (for instance once a year), the potential service interruption that might be experienced during the actual transition from the old to the new network state is not a big concern.

On the contrary, when the EM-NO is faced, the objective is to dynamically reduce the power consumption of the network, following the decreasing of the traffic occurring on a daily scale; thus, the problem should be solved a number of times during a day, considering a different traffic matrix each time. This imposes serious limitation on the amount of time that can be spent to find the solution, making practically infeasible solving the ILP formulation provided above even for small networks. Moreover, since the network is optimized much more often, proper mechanisms are needed to change the network state avoiding any possible out of service; furthermore, such a transition should be automatically performed by network nodes without the human participation. Ultimately, even if the two application scenarios happen to be quite similar from a mathematical point of view, they require different approaches in practice.

III.3 Energy Saving at the IP Layer

As far as the IP layer is concerned, the objective of the energy minimization problem is to minimize the number of line cards and switching matrixes. Let us analyze these two objectives separately. First, the traffic treated by each IP router can be divided in two components:

i) **the local traffic**: is the traffic generated or terminated by edge routers connected to the particular core router; such a traffic component is thought as originated/terminated by the node and is accounted in the traffic matrix;

ii) **the transit traffic**: is the traffic generated or terminated by edge routers connected to a different core router; such a traffic must be forwarded towards another core router in the network.

The number of line cards needed on a node depends on the total traffic treated by the node; since the local traffic is an input of the problem, a viable strategy is to reduce the total amount of transit traffic in the network. It can be noted that, the minimum number of line cards which each node must be provided with depends on the amount of local traffic. Each line card on a given node *i* constitutes a link towards another node *j* that will be used to transmit/receive the traffic flow between the two nodes. Other traffic flows towards nodes *k* that are not directly connected to the node *i* can be transmitted/received using the spare capacity on the various links *i*,*j*. Such traffic flows constitutes transit traffic for nodes *j*, thus increasing the total amount of treated traffic and then the number of needed line cards. However, as new line cards are added to treat transit traffic, new links are added to the IP topology that will be used to transfer local traffic, thus reducing the amount of transit traffic. Moreover, if links are added in a proper way, the average length of the shortest paths among nodes decreases thus allowing each traffic flow to traverse a smaller number of hops and to further reduce the amount of transit traffic. Intuitively, this process stops when the total capacity deployed in the network is sufficient to support both local and transit traffic.

Let us assume to have minimized the number of line cards; the number of switching matrixes needed in each IP router is due to the total amount of treated traffic. In order to reduce them, we should further reduce the amount of transit traffic in each node; however, this means adding new line cards in the network. Thus, we can conclude that, the objective of minimizing the number of line cards is coherent with the objective of reducing the number of switching matrixes since the former can be achieved by limiting the amount of transit traffic; however, once the first objective has been achieved, minimizing the number of switching matrixes means increasing the number of line cards and thus is in contrast with it.

Finally, it is worth remembering that the power consumption related to switching matrixes is much less than the power related to line cards (see II.3.1). For such a reason, in order to simplify the problem, the former power contribution can be neglected, reducing the whole problem to the sole objective of minimizing the total number of line cards.

The new simplified problem can be given as follows:

Minimize:

$$P_{IP}^{LC} = \sum_{i \in V} \left(\sum_{j \in V} \left(x_{i,j} \right) \right) \qquad \text{(III.21)}$$

Eq. III.21 represents the new objective functions that simply counts the total number of installed line cards.

Subject to:

Eq. III.4, eq. III.5, III.6, III.7, III.8, III.9.

III.3.1 *The Energy Minimized Virtual Topology Design Problem*

Considering the IP layer, the EM-ND problem deals with designing the virtual topology with the goal of minimizing the total power requested by the IP layer. The Virtual Topology Design (VTD) problem is also known as traffic grooming problem since it deals with grooming the smaller traffic flows of the higher network layer into the higher bit-rate traffic flows established at the lower layer. In IP over WDM networks this means aggregating the IP layer traffic into the lightpaths established in the optical network that directly interconnect IP routers. In doing this two opposite strategies can be distinguished: the link-by-link and the end-to-end grooming. In case of link-by-link, the grooming operation is performed at each intermediate node for any traffic flow. In this case, lightpaths are terminated at each intermediate node and all traffic is switched at the IP layer. Such a strategy is also referred to as non-bypass strategy since traffic is not allowed to optically bypass a node. On the opposite, the end-to-end grooming strategy consists of aggregating the whole traffic request between a node pair and routing it on a single lightpath, which directly interconnects the two end-nodes. Such a strategy is also known as single-hop lightpath bypass since each IP traffic flow is allowed to be routed on a single hop on the VT, i.e. on the IP topology. The resulting VT is the same as the physical topology, in the case of non-bypass, whilst, in the case of single-hop bypass, is composed of a link for each traffic request in the traffic matrix.

Although, in many cases, these two opposite strategies do not lead to the optimal solution of the VTD problem since they constrains the VT to a predetermined topology, comparing them in terms of energy consumption can give a general indication about how to construct efficient heuristics where both the non-bypass and the single-hop constraints are removed. A similar study has been performed in [21].

Results showed in that study demonstrate that as the total traffic offered to the network grows, the end-to-end strategy leads to a significant saving with respect to the link-by-link case. Such a result can be explained observing that, when the offered traffic is low, and in particular when each traffic relation is lower than the capacity of each lightpath, which is equal to the bit-rate of a line card, the end-to-end strategy leads to a considerable bandwidth waste since each lightpath carries only a small amount of traffic with respect to its capacity. Thus, line cards are underutilized leading to a considerable energy waste. Instead, as the offered traffic grows and each traffic relation becomes comparable with the line card capacity, the end-to-end strategy allows to fully use the capacity provided by each lightpath reducing the total consumption with respect to the link-by-link case. In particular, the higher the average bandwidth requested by a traffic relation is, the better the end-to-end strategy performs. An alternative strategy is to allow each IP traffic flow to be routed on a multi-hop path in the VT, removing any constraint on the VT itself. The problem has been faced in [22], [23]. In both papers the problem is formalized as a MILP problem and heuristic solutions are provided to solve the problem in reasonable time.

In [22] each IP router is assumed to consume an amount of power given by the number of line cards deployed on it and each line card is assumed to consume a fixed amount of power regardless of the actual carried traffic. Thus, the goal is to minimize the total number of line cards installed network's routers that is equivalent to minimize the total number of lightpaths. The proposed solution works in a greedy way. The first step is to order traffic relations in decreasing order of requested bandwidth; then each relation is processed to be routed on the current VT. The VT is constructed step-by-step starting from an empty topology, i.e. with no links. At each step the current traffic relation is routed on the shortest path with enough capacity, if any, otherwise a direct link is added in the VT connecting the two end-nodes and the traffic relation is routed on that link. The followed principle is to route traffic using the spare capacity as much as possible. Concerning the traffic flow ordering criterion, it has the objective of constructing direct links for higher traffic relations, trying to route smaller flows on multi-hop paths, trying to minimize the total network load. The proposed solution is compared with the direct-bypass and the non-bypass strategy and with the optimal solution given by the MILP formulation. Results are showed in Figure III-3. They are expressed as percentage saving with respect to the non-bypass strategy versus the average traffic between each node pair; the line card bit-rate is assumed to be 40 Gbits/s. Such results confirm that when the average bandwidth requested by each traffic relation is close to the capacity of each line card, the direct bypass strategy performs close to the optimal solution. At low load, instead, the proposed solution allows to save more energy than the direct-bypass strategy; however about a further 15% is needed to achieve the optimal solution.

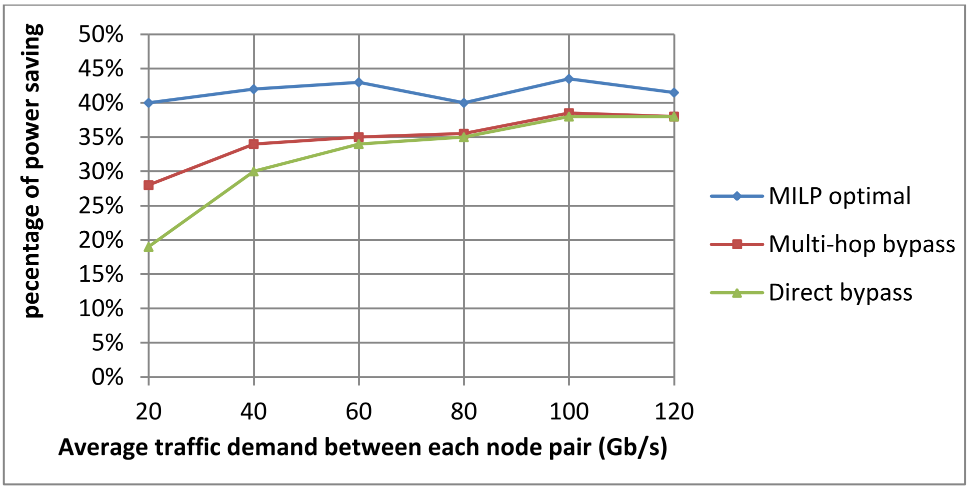


Figure III-3: Power consumption saving by various lightpath bypass approaches relative to the approach of non-bypass. Source [22].

The study in [23] considers a different power consumption model. It is composed of two different terms.

i) The term $P_{TX}(B_{TX})$ is related to the power required by a Transmitter-Receiver (TX-RX); it does not depend on the actual load but only on the bit-rate B_{TX}, i.e. a fixed amount of power is required for each TX-RX;

ii) The term $P_{SW}(\lambda)$ is related to the power required to electronically switch the traffic; it is assumed to linearly grow with the amount λ of switched traffic.

Regardless of the specific functionality which the two terms of power consumption are related with, the key point is that each lightpath established in the network can be assumed to require a fixed part of power plus a variable part proportional to the carried traffic. Clearly, the ratio between the two terms, say ν^0, has an important impact on the optimal VT. In this study two different solutions are proposed. The first one is an iterative greedy algorithm very similar to the one proposed in [22]. The only difference is that, when a traffic flow is processed to be routed, the power needed to route it on the shortest path with enough capacity is evaluated and compared with the power needed to add a new direct link and route the flow on it; then the least energy alternative is chosen. This makes sense because, due to the variable term of power consumption, a traffic flow is assumed to consume power even if it is routed on lightpaths that are already established. Authors also propose a solution based on a genetic algorithm. Results showed in this study are interesting since they are carried out as the ratio ν^0grows.

The parameter ν^0 is defined as follows:

$$\nu^0 = P_{SW}(\lambda=B_{TX}) / P_{TX}(B_{TX}). \quad \text{(III.22)}$$

Figure III-4 shows the average number of transmitters which each node is equipped with, which represents the average nodal degree, versus the ratio ν^0, for the optimal and the heuristic solutions. Results demonstrate that, as ν^0 grows, the optimal VT becomes more and more meshed. In fact, when the fixed term becomes negligible with respect to the variable one, the main goal is to minimize the total network load and the optimal solution tends to a full-meshed VT. For high values of ν^0 the solution can be easily found and the two proposed solution perform very close to the optimal one (see Figure III-5); instead, for low values of ν^0, the genetic algorithm outperforms the greedy algorithm.

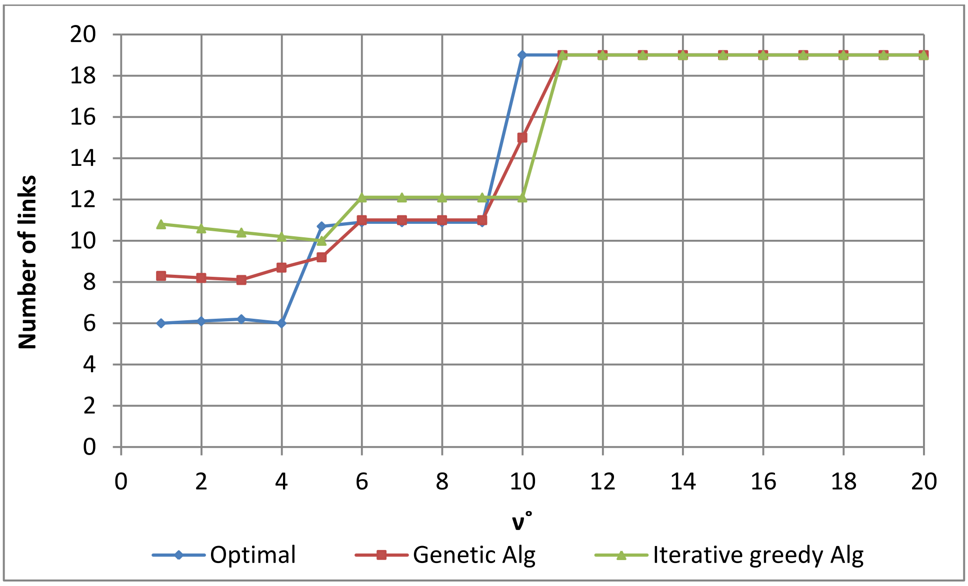


Figure III-4: Number of links Vs. v^0 for the optimal solution, the genetic algorithm and the iterative greedy algorithm; 20 nodes. Source [22].

It has to be outlined that current electronic devices, such as line cards, switching matrixes, etc., have a power consumption profile that is almost independent of the load; this correspond to a very low v^0. In this case the goal of the EM-VTD problem is to minimize the number of lightpaths trying to aggregate traffic as much as possible. In such a situation finding a good VT is a hard task especially when traffic relations are small compared with the capacity of a lightpath (B_{TX}). On the other hand, thanks to the introduction of mechanism such as Dynamic Voltage Scaling, Dynamic Frequency Scaling, Low Idle Power, etc., current electronic devices are moving towards more load-proportional power profiles; this will clearly improve the energy efficiency of the network and simplify the EM-VTD problem. However, since a full proportionality between the power consumption and the load is unlikely to be achieved also in the future, the EM-VTD problem remains a key point to improve the energy efficiency of the network. Another important indication is that as electronic devices will move towards more load-proportional power profiles, the optimal VT will require an increasing number of lightpaths for the same amount of traffic; this will lead to a growth of the share of power required by the optical network layer. Thus, improving the energy efficiency of the optical layer will become more and more important.

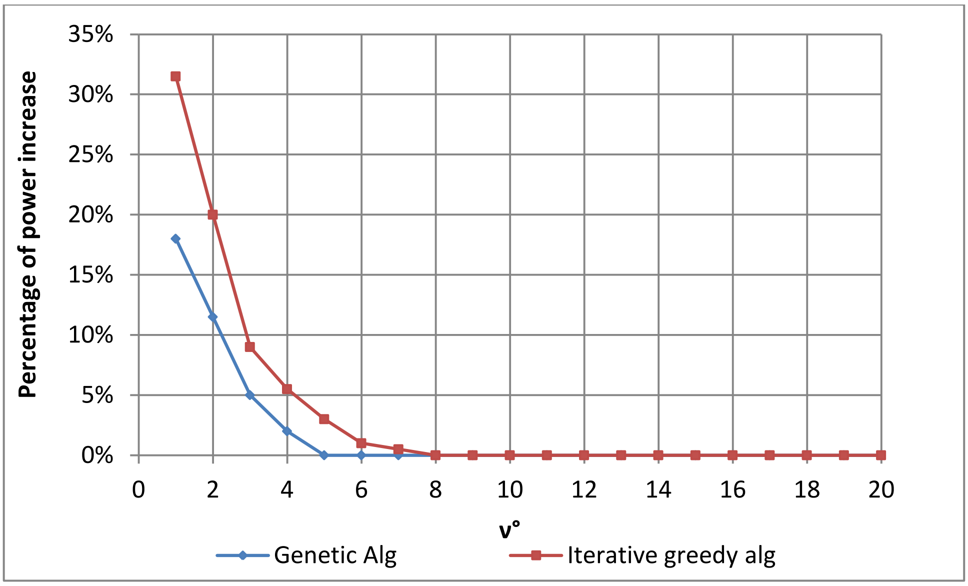


Figure III-5: Power consumption increase relative to the optimal solution vs. v^0 for the genetic algorithm and the iterative greedy algorithm; 20 nodes. Source [22].

III.3.2 *Energy-Aware Traffic Engineering*

Energy-Aware Traffic Engineering (EA-TE) strategies aim at bringing the network energy consumption nearly proportional to the amount of carried traffic; thus, they deal with the EM-NO problem. As well known, Internet design follows two main principles: redundancy and resource over-provisioning; the combined effect of these two principles enable Internet to support traffic variations and possible faults avoiding QoS degradations.

Taking in mind the traffic fluctuations, it results that network resources are underutilized for most of the time whilst their consumption is independent of the load.

EA-TE strategies generally aim at aggregating traffic flows over a subset (possibly the minimum one) of the network elements, allowing other links and interconnection devices to be put in power saving mode. These solutions should satisfy the traffic demand variations in time, preserve connectivity and QoS, for instance by limiting the maximum utilization over any link, and ensure a minimum level of path redundancy.

A generic block diagram of an EA-TE policy is shown in Figure III-6 [24].

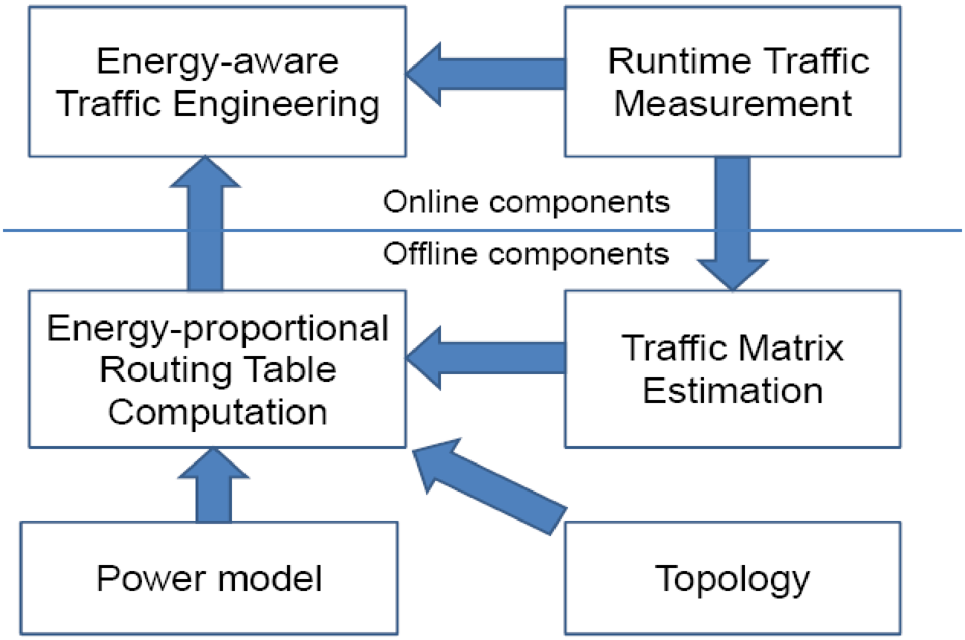


Figure III-6: Block diagram of a generic EA-TE policy.

In the off-line plane, a representation of the network topology, a power model of the network devices, and a traffic matrix estimation (if needed) are used to pre-compute the set of paths, so as to minimize the network energy consumption and the complementary set of links to be put in low power mode. The decisions made by the computation block are transferred to the network elements to be installed in their routing/forwarding tables. The traffic information are collected on-line and transferred to the processing functional blocks residing in the off line plane.

EA-TE strategies differentiate for the different ways adopted to implement the functional blocks of the generic scheme shown in Figure III-6. Different choices concern:

1) the centralized or distributed implementation of the energy-saving path computation function;
2) the algorithm used to determine the energy-saving path set;
3) the type of information utilized to compute the energy-saving paths;
4) the type of routing information and the mechanism used to exchange them between the network elements;
5) the path updating period.

As far as the application of EA-TE strategies is concerned, an interesting problem consisting in determining when the EA-TE strategies can be usefully applied and their expected potential benefits. In other words, it would be useful to have a rough preliminary estimate of the possible savings that the adoption of smart energy saving policies may entail and in which cases these advantages can be achieved. Some studies face this problem, e.g. [25].

These studies use a very simple network model, mainly based on purely topological properties, and consider, together with the benefit due to the link/node power-off, the average effect of traffic load over network links due to path rerouting. The energy consumption of a link interface is modelled by a constant part, that includes the fraction of the node energy cost due to the link, and a variable part that is proportional to the traffic loading the link.

Although the approach is oversimplified, it is able to capture some interesting properties on the field of applicability of EA-TE strategies.

It is to be observed that, the variable part considered in the energy consumption model means that power management capabilities provided by the HW level are implemented in link interfaces.

Lessons learned by the application of these prediction models are:

i) the advantage of application of sleeping techniques is increasing as the network size is increasing, i.e. EA-TE methods are the more convenient the more the network is large and the number of devices grows;

ii) EA-TE strategies are not convenient in small networks, where computation of energy-saving path sets is limited by the low value of path redundancy;

iii) there is a trade-off between EA-TE techniques and power management capabilities, if the device consumption is independent of the load, or lightly variable, the use of sleep modes is very effective; on the other hand, if efficient power management techniques are implemented in the network devices, the effectiveness of sleep mode approaches decreases.

A further aspect that has to be considered for EA-TE strategy application is that, unfortunately, putting interfaces of switches or routers to sleep can have serious drawbacks due to the network protocols operations. As discussed in [26], the main elements to be considered are the possible interactions of a specific EA-TE strategy with Layer 3 control protocols; specifically, a specific attention must be devoted to the reaction of routing protocols at a sleeping event of a network interface; as a matter of example, if a router put an interface to sleep, OSPF (Open Shortest Path First) would generate a flood of LSA (Link-State Advertisement) packets indicating the that link is down, followed by a re-computation of the Shortest Path First (SPF) algorithm by all routers. Analogously, Border Gateway Protocol (BGP) may suffer from route oscillations and, occasionally, forwarding loops.

An important aspect related to the application of an EA-TE strategy concerns what can be done in the network and what cannot; specifically, it is useful to fix some general rules about the possibility of changing the initial topology and routing. Let us consider a standard situation represented by a static network where the capacity of each link is not changed in time and the path of each traffic flow is kept fixed in the network; we refer to such a network as *Base Network* (BN). The BN may be given by the solution of the MILP formulation provided in III.3.1 or may be found using any other algorithm, such as the ones proposed in [22], [23]. The topology of the BN is completely described by the number of line cards deployed on each link i,j, say $N_{i,j}^{BN}$. Thus the issue is what the EA-TE mechanism can change and what cannot.

First, let the active capacity of a link be defined as follows:

$$C_{i,j}^{On} = N_{i,j}^{On} \cdot \Delta C;\ 0 \le N_{i,j}^{On} \le N_{i,j}^{Max};\ N_{i,j}^{On} \text{ Integer}; \qquad \text{(III.23)}$$

In (III.23) $N_{i,j}^{On}$ is the number of line cards kept active on the link *i,j*, whilst $N_{i,j}^{Max}$ is the maximum number of line cards that can be activated on that link.

According on the work in [27] three different scenarios can be distinguished. The most restrictive scenario is named *Fixed Routing Fixed Topology* (FRFT). In this scenario neither the routing nor the IP topology can be changed in time but only the capacity reserved for each traffic flow can be updated. In particular, keeping fixed the IP topology means that for each link *i,j* it must result $N_{i,j}^{Max}=N_{i,j}^{BN}$. In other words, the capacity that can be activated on each link is limited to the capacity deployed on that link for the BN. Since the routing cannot be changed, in this scenario, a line card can be switched off only if it belongs to a link composed of more than one line card. In fact, if traffic flows cannot be rerouted, links composed of only one line card cannot be empted and the related line card cannot be switched off.

The second scenario is named *Dynamic Routing Fixed Topology* (DRFT). It differs from FRFT because the path of each flow can be rearranged in time whilst the same restriction is kept for the topology. Unlike FRFT, in DRFT any line card can be switched off by rerouting the LSPs established on the related link on the rest of the network.

The last identified scenario is named *Dynamic Routing Dynamic Topology* (DRDT). In this scenario both the routing and the IP topology can be changed in time regardless of the starting network; in particular the LSR topology is not constrained to the BN topology. The only constraint is that the total number of line cards used on a node at any time must not exceed the number of line cards installed on that node for the BN. More in detail:

$$\sum_j N_{i,j}^{On} \leq \sum_j N_{i,j}^{BN} \qquad \text{(III.24)}$$

In DRDT the set of line cards deployed on a node *i* is shared among all possible *i,j* links and each line card can be dynamically assigned to a different link if unused. The practical realization of the DRDT scenario has the following implications. First of all, when a line card is switched off, all line card functionalities must be powered off, including those related to the physical layer. In particular, the lightpath connecting the line card pair must be torn down since they might be assigned to a different link in the future. For the same reason, the activation of a line card pair on a given link needs to set up a lightpath in the optical layer. This means that an Automatically Switched Optical Network (ASON) [45] is needed at the optical layer in order to allow to dynamically set-up or torn-down a lightpath; as a consequence, longer delays will be experienced in the network during the line card activation process due to signaling operations in the ASON and to the synchronization of physical receivers. However, such a delay is not a big concern since the activation of a new line card is done with the aim of optimizing the network and not as a consequence of a fault.

A number of papers have faced the EA-TE problem [28]-[33]. Taxonomy of the proposed solutions has to take into account at least the following aspects:

i) the algorithm used to find the best set of nodes and link to put in idle;

ii) the operation architecture (centralized or distributed);

iii) the network scenario, (e.g. pure IP or connection-oriented networks);

iv) routing strategy (TE based or classical shortest path routing).

The work in [28] faced the EA-TE problem for the first time; it proposes an iterative heuristic algorithm to decide the set of nodes and links to put in idle. The algorithm starts ordering nodes and links according to a given criterion, e.g. the load or the node degree; at each step the next element is selected and the algorithm checks if all the offered traffic can be supported putting the selected element in idle. Then the next element is selected. Such a kind of algorithm is suitable for a centralized scenario and uses a link/node-based approach, i.e. first the set of network elements to be put in idle is decided and then the traffic on them is rerouted on the rest of the network. The considered network scenario is a pure IP network and consequently the classical shortest path routing is assumed.

The work in [29] proposes a different solution based on the same approach in a quite similar network scenario; specifically, it exploits the algebraic connectivity to find the best set of links to be put in idle. The idea is to put in idle the links that have the lowest impact on the network considering the consequently decreasing of the algebraic connectivity. However, the heuristic does not consider the link load that probably is the main reason why it is not able to outperform the previous described solution. Combining the two aspects might result in a more efficient solution.

In [30] authors propose a link/node-based approach for a pure IP network. However, unlike the two aforementioned works, a traffic flow can be routed on a path different from the original shortest path by changing OSPF link weights. Two different heuristics are proposed: the first one is based on an iterative algorithm very similar to the one proposed in [28] except that is starts by looking for the set of OSPF link weights that minimize the total network load; then the iterative algorithm is run choosing an ordering criterion for nodes and links. Finally, links that must be put in idle are excluded from the active topology assigning them a very large weight that is the most straightforward solution to achieve such an objective. The second proposed heuristic is composed of two phases: in the first one it looks for the active topology minimizing the overall power consumption solving an ILP formulation that assumes fully splittable routing (instead of shortest path routing). The second phase tries to find a set of OSPF link weights such that the traffic demand can be routed on the active topology decided in the first phase, using the shortest path algorithm; if it does not succeed, the first phase is repeated increasing the traffic demand in order to obtain a topology with more active elements. Although the second heuristic outperforms the first one in case of moderate of high traffic load, it is very heavy and takes up to 10 hour to find a solution; this makes it infeasible when traffic changes on a daily scale.

All the previous works do not consider routing instabilities such as transient loops that may be experienced during routing change; in fact, as already outlined, the only way of excluding links from the shortest path tree (SPT) is to assign them a very high value. The work in [31] instead considers such an issue and proposes a loop-free distributed scheme to change OSPF weights. The routing is still based on the classical shortest path and optimal OSPF weights are computed by applying a Lagrangian relaxation to an ILP formulation that assumes shortest path routing and considers OSPF weights as decision variables.

The Energy Aware Routing (EAR) algorithm presented in [32] is able to switch off network elements exploiting a modified version of the OSPF protocol instead of acting on link weights. The EAR algorithm selects a subset of routers, Importers Router (IR), that do not calculate their own SPT but use the one of a neighbour router, the Exporter Router (ER); in this way, some links are excluded from the SPT of any node and can be switched off. The set of IRs and the relative ERs are chosen assuring loop-free routing and maximizing the number of links that can consequently be put in idle. An advantage of EAR is that it realizes the network rearrangement completely avoiding packet loss.

Considering a pure IP network makes necessary assuming an SPT-based routing; this, in turn, has an important implication on the optimal solution since limits the choice of the path which a traffic flow can be routed on. In fact, it can be noted that, however link weights are chosen, two paths from two different source nodes towards the same destination node can never diverge in the network. The same holds for the last discussed solution. Moreover, in such a kind of network, link load information is hard to be known by other network nodes or by a centralized management entity, making difficult assuring that no congestion is experienced when a network element is switched off. In fact, even if TE routing protocols can disseminate such information, they generally consider the reserved bandwidth as link load instead of the load evaluated by really monitoring the link.

In [33] a TE-based solution is proposed adopting a hybrid OSPF/MPLS scheme. The solution is based on an MILP formulation that does not constrain the routing to be SPT-based and assumes fully splittable routing. If a path is different from the one determined by shortest path routing, an MPLS tunnel is established in the network, otherwise the classical shortest path routing is kept for that flow. In order to decrease the complexity of the solution and to limit the maximum path length and delay, proper constraints are added to the MILP formulation so that only a subset of all possible paths between each source-destination pair is considered; they practically constrain each flow to be routed on the first *k* shortest paths. If *k* is large enough, it is possible to limit the complexity of the solution keeping the gap between the heuristic and the optimal solutions low. Unfortunately, authors show that, for a network of 20 nodes, limiting *k* to 10 assures a good efficiency of the solution but involves too long computation time (more than one hour). A possible solution is to stop the computation after a fixed amount of time (authors consider 300 seconds) and taking the best found solution; however, even if the additional degradation is limited (about 4%) the solution remains impracticable for even large networks.

Table III-1 summarizes discussed solution.

PA-R Technique	**Searching algorithm for link switch off**	**Operation architecture**	**Network scenario**	**Routing strategy**
L. Chiaraviglio, M. Mellia, and F. Neri [9]	Heuristic, Load based link ranking	Centralized	Pure IP	Shortest path
Cuomo, F.; Abbagnale, A.; Cianfrani, A.; Polverini, M. [10]	Heuristic, algebraic connectivity based link ranking	Centralized	Pure IP	Shortest path
Amaldi, E.; Capone, A.; Gianoli, L.G.; Mascetti, L. [11]	Heuristic, Load based link ranking & MILP based	Centralized	Pure IP	Shortest path - OSPF link weight variation
Lee S.S.W, Tseng P., Chen A. [12]	Lagrangian relaxation of MILP formulation	Centralized	Pure IP	Shortest path - OSPF link weight variation
Cianfrani, A.; Eramo, V.; Listanti, M.; Polverini, M. [13]	Heuristic, SPT exportation	Mixed (centralized/distributed)	Pure IP	Shortest path – SPT modification
Mingui Zhang; Cheng Yi; Bin Liu; Beichuan Zhang. [14]	MILP formulation	Centralized	IP/MPLS	Mixed shortest path or alternative routing

Table III-1: Taxonomy of EA-TE strategies.

III.4 Energy Saving at the WDM Optical Layer

The energy consumption issue of optical networks has been investigated less than the energy consumption of IP or electronic networks because of the lower consumption of photonics devices with respect to their electronic counterparts. In fact, as outlined in II.3, the power consumption of the optical layer is estimated to be at most 10% of the total power required by a core network. Most studies in this field have focused their attention on the estimation of the energy footprint of optical networks [14]-[16], [34]-[36], comparing different optical node architectures and technologies (e.g. optical packet-switched and optical circuit-switched nodes) with each other and with the electronic counterparts, but just a few studies have proposed saving strategies [37], [38].

However, there exist several reasons that make energy saving strategies an important topic also for circuit-switched optical networks. First, as the traffic grows, the energy-efficient design of the virtual topology leads to a growth of power consumed at the optical network because of the larger number of requested lightpaths. Moreover, as outlined in III.3.1, as electronic devices move towards more load-proportional power profiles, the EM-VTD problem leads to more meshed virtual topologies, with many links, i.e. many lightpaths, not fully loaded. Thus, the introduction of energy saving strategy in the optical network will become an important issue in the coming years for further improving the energy efficiency of the whole network.

Considering the network architecture described in II.3.2 and the problem definition provided in III.2, it is clear that the main way for reducing the power consumed by an optical network is to keep unused some fibre links in order to put in idle optical amplifiers deployed along such links. In this regard, the first important point concerns the applicability of a similar strategy; specifically, the main question is how to put in idle optical amplifiers deployed along unused fibre links. Such a problem can be easily overcome considering that the switch off of optical amplifiers could be easily performed in an automatic way by implementing a control device inside the amplifier that put it in idle state when the optical power inside the fibre goes below a given threshold. Some optical amplifiers implementing a similar mechanism for laser protection purpose are already commercially available [39].

Another important point concerns the considered scenario; specifically it is possible to distinguish between the Static Lightpath Establishment (SLE) and the Dynamic Lightpath Establishment (DLE) scenarios [47]. In a static scenario the set of lightpaths remains the same for a long time (several months) and thus the power consumption of the optical layer does not change in time. In this scenario, the traffic matrix in known in advance and no control plane is required in the network. In fact, lightpaths can be established via the management plane and the network can be re-optimized when the traffic matrix changes considerably. On the contrary, in the dynamic scenario connection requests dynamically arrive and the each lightpath is assumed to have a finite lifetime. In this scenario, the traffic matrix in not known in advance and a control plane is required in the network to set up and tear down lightpaths requested by upper layer nodes. A similar scenario may be obtained as a consequence of link switch-off procedures undertaken at the IP layer when even the physical layer functionality of IP router line cards is switched off during the switch off of a link. The direct consequence is that traffic variations experienced at the IP layer are reflected at the optical layer allowing the adaptation of the power consumption of the optical network to the offered load.

Finally the last point concerns the different strategies that it is possible to follow in order to achieve the goal, i.e. keeping unused as many optical fibre links as possible. In fact, at least two different strategies can be undertaken:

1) **Fibre switch-off:** this strategy consists of switching off some optical fibre links by rerouting the related lightpaths on the rest of the network. Such a strategy is useful in the DLE scenario when, due to traffic variation, the network state rapidly changes. However, a similar strategy must also take into consideration service continuity and assure no service interruption during lightpath rerouting.

2) **Power-Aware RWA (PA-RWA):** this strategy concerns the definition of RWA algorithms pursuing power consumption minimization. It can be applied to both SLE and DLE scenarios; in this regard it is possible to distinguish between Dynamic PA-RWA and Static PA-RWA. Unlike fibre switch-off, in this case lightpaths are never rerouted but they are established taking into account the power consumption of network elements and the state of the network, and aiming at minimizing the total needed power. However, other network performance metrics must be considered when a PA-RWA algorithm is defined. Specifically, when the SLE problem is faced, classical RWA algorithms aim to minimize the number of wavelengths W that are needed to support a given traffic matrix or, equivalently, to maximize the number of supported connections for a given W. Instead, in the DLE scenario, the objective is to establish lightpaths so that the total amount of blocked connections is minimized. Although both the objectives are not consistent with the problem of minimization of the energy consumption, a

viable solution should provide a good trade-off between these different performance metrics.

III.4.1 *Power Aware Routing and Wavelength Assignment*

Concerning the PA-RWA problem, in [37] authors propose two different heuristic algorithms to solve the PA-RWA problem, namely Most Used Path (MUP) and Ordered Lightpath-MUP (OL-MUP); both of them belong to the general family of least cost path algorithms; this means that the cost of a path is given by the sum of the costs of each link belonging to it and the path with the minimum cost is chosen. The basic idea of MUP is to assign to each fibre link a cost proportional to the power that it requires and dependent on its state. Specifically, if a link is already used, it gets a null cost since using it does not require any additional power; instead, if a link is unused, it gets a cost equal to the total power required by optical amplifiers deployed along it. Specifically:

$$LC_{i,j} = LF(l_{i,j})*PC_{i,j}. \tag{III.25}$$

$$LF(l_{i,j}) = \begin{cases} 0 \; if \; l_{i,j} > 0 \\ 1 \; Otherwise \end{cases} \tag{III.26}$$

In (III.25) the cost $LC_{i,j}$ of a link i,j is given by two terms: i) $PC_{i,j}$ is the power consumption of the link i,j that is due to optical amplifiers deployed along it; it can be computed according to eq. II.7. $PC_{i,j}$ is a static term that does not depends on the link load; ii) $LF(l_{i,j})$ is a variable term that depends on the load $l_{i,j}$ of the link i,j; the load, i.e. the variable $l_{i,j}$, is the number of used wavelengths.

OL-MUP differs from MUP since it orders lightpaths before to route them; specifically it selects the proper lightpath to be routed according to a particular criterion. For such a reason OL-MUP can only be applied to the

Static Lightpath Establishment (SLE) scenario since it requires that the traffic matrix is known in advance; on the contrary MUP can be used in both SLE and Dynamic Lightpath Establishment (DLE) scenarios.

In [37] the proposed algorithms are compared with the power-aware version of the classical shortest path algorithm (ShP); the last, in fact, is a simple shortest path algorithm where the cost of each link *i,j* is simply equal to $PC_{i,j}$. Results reported in [37] demonstrate the effectiveness of the proposed heuristics in reducing the power consumption of the network with respect to ShP; however no other performance metric has been evaluated.

In fact, a pure power minimization provisioning strategy may have an adverse impact on the length of the routed lightpaths, i.e., in order to avoid to unnecessarily power on network elements, the provisioned paths are, on average, longer than the one found with traditional shortest path solutions. This is in contradiction with the goal of classical routing and wavelength assignment (RWA) algorithms that tend to minimize the resource usage within the network, in order to minimize the blocking probability.

The study in [38] addresses this problem by investigating the impact that power minimization has on the overall network performance. With this purpose in mind, Authors present a non-conventional solution to the power-aware RWA problem. This approach, referred to as Weighted Power-Aware RWA (WPA-RWA), is based on the intuition that in some cases relaxing the power minimization constraint can have beneficial effects on the overall network performance, i.e., it can contribute to the reduction of resource fragmentation in the network and lower the blocking probability. The proposed algorithm leverages on a cost function that weights the power status of network elements versus information about the wavelength usage.

The proposed WPA-RWA strategy is based on a modified version of the k-shortest path algorithm [40], and it works as follows. For each connection request arriving at the network edge, up to k candidate paths are computed. The algorithm accounts for the current resource usage (i.e., wavelengths) in the network, where fiber links without available wavelengths are temporarily deleted from the network topology. When computing each candidate path, each fiber link *i,j* in the network is assigned a weight $LC_{i,j}$ given by eq. III.25, where the term $LF(l_{i,j})$ is computed as follows:

$$LF(l_{i,j}) = \begin{cases} \alpha \ if \ l_{i,j} > 0 \\ 1 \ if \ l_{i,j} = 0 \end{cases} \quad \text{(III.27)}$$

In III.27 α is a weighting factor with values between 0 and 1. Note that with values of α close to 0 WPA-RWA behaves as MUP, while for values of α close to 1 WPA-RWA tends to provision connection requests along shortest routes.

Figure III-7 and Figure III-8 show the values of the network blocking probability, and the average power saved per request, respectively. Both metrics are presented as a function of the network load. The average power saved per request is computed as the difference between the total network power consumption obtained when α=1 and the total network power consumption for any other given value of α.

The figures show that considerable power savings (up to 50%) are achievable, but at the expense of a relevant increase in the network blocking probability; moreover, it can also be noted that, as the traffic load increases, the power saving decreases down to 5% even for α=1 that corresponds to the MUP algorithm.

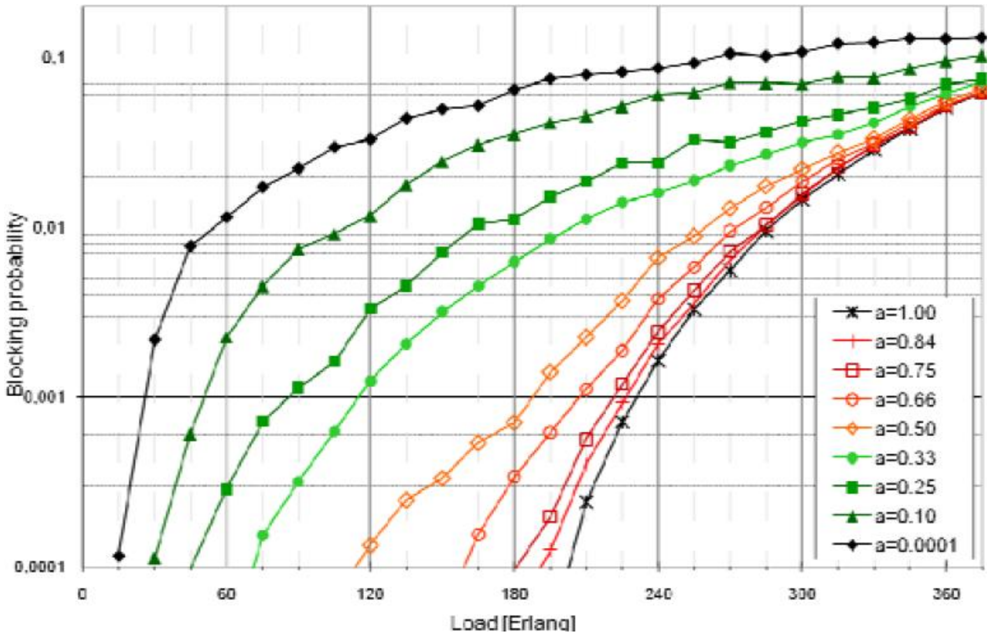


Figure III-7: Blocking probability vs. network load for different values of α. Source [38].

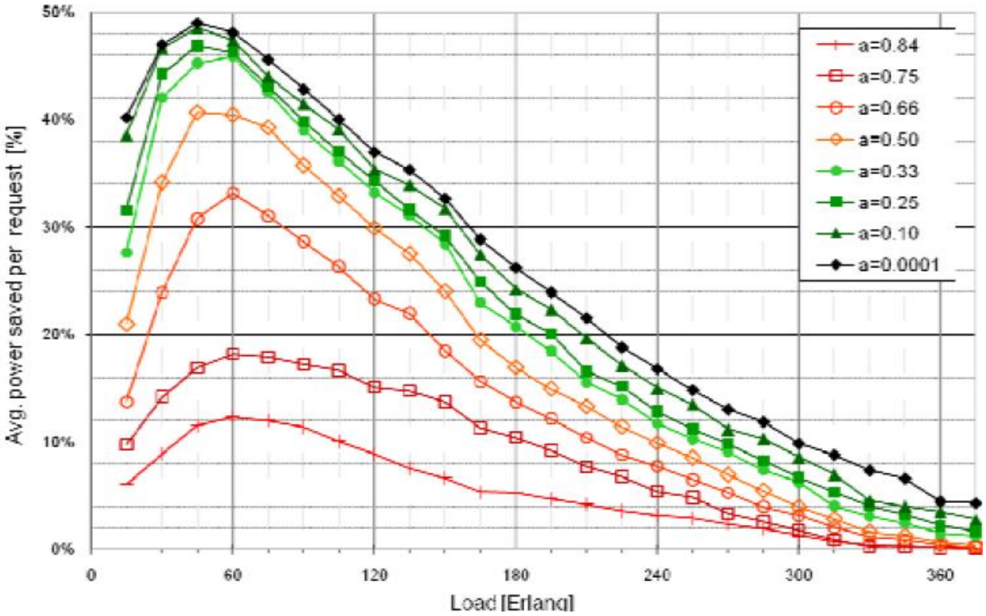


Figure III-8: Average power saved per requests vs. the network load for different values of α. Source [38].

Chapter IV

ENERGY SAVING IN BACKBONE IP OVER WDM NETWORKS: PROPOSED SOLUTIONS

IV.1 Energy Saving at the IP Layer

The energy efficiency of IP networks is the most important aspect for achieving energy conservation in telecommunication networks. We discussed the problem in III.3 where we outlined that the main objective is to minimize the total number of line cards needed to support the traffic request. Two different problems have been defined; the EM-VTD problem and the EA-TE. The former aims at reducing the peak power consumption since it considers a static traffic demand representing the peak demand among network nodes. It is a long-period optimization strategy that is generally performed off-line. We address this problem proposing a heuristic algorithm in IV.1.1. The latter, instead, aims at adapting the power consumption of the network to the offered load, which results quite variable in time on a daily scale. An EA-TE strategy is proposed in this dissertation and presented in IV.1.2.

IV.1.1 *The "Start –Single Hop and ReRoute" Algorithm for the EM-VTD problem*

The EM-VTD problem has been introduced in III.3.1 where has been outlined that it corresponds to a traffic grooming problem. Concerning the traditional traffic grooming problem, two different approaches can be found, namely the single-hop and the multi-hop lightpath bypass strategies [41], [42]. In the first case, the VT is composed of a link with sufficient capacity for each traffic relation in the traffic matrix, which is used to accommodate the related traffic flow. Formally, problem variables (see eq. III.4 and III.5) are fixed to the following values:

$$x_{i,j} = \text{Ceil}(T_{i,j} / \Delta C) \qquad \forall\ (i,j)\ \epsilon\ V\text{x}V. \qquad \text{(IV.1)}$$

$$f_{i,j}^{s,d} = \begin{cases} 1;\ if\ (i,j) = (s,d) \\ 0;\ if\ (i,j) \neq (s,d) \end{cases} \qquad \forall\ \text{(i,j),(s,d)}\ \epsilon\ V\text{x}V. \qquad \text{(IV.2)}$$

Clearly, such an approach leads to a considerable bandwidth waste when traffic requests are smaller than the capacity of each line card, whilst leads to the optimal solution when all traffic requests are larger than ΔC.

The multi-hop strategy, instead, allows a traffic flow to cross many virtual links in the IP topology, sharing the capacity of each virtual link with others traffic flows. The multi-hop strategy is very useful when traffic flows are small compared with the capacity of each line card since it allows to groom several flows on the same links, thus reducing the total number of links, and therefore of line cards, deployed in the network.

In III.3.1 the solution proposed in [22], here referred to as *Start-Empty*, has also been discussed and related results have been presented. The above cited algorithm is extended from the multi-hop lightpath bypass strategy and works in a greedy way by adding links to the VT step-by-step.

The algorithm proposed here, instead, is still based on the multi-hop grooming strategies but uses an innovative approach for constructing the IP topology and for deciding the traffic routing. Specifically, rather than constructing the IP topology adding a new link at each step, the proposed heuristic starts considering a full-meshed virtual topology and then tries to eliminate unnecessary links.

The novel heuristic algorithm proposed here is named *Start Single-Hop and Reroute* (Start-SH&ReR). The basic idea is to start considering the solution given by the single-hop lightpath bypass (i.e. a link for each traffic relation) and then trying to eliminate unnecessary links by iteratively rerouting traffic flows. To achieve such an objective, the classical shortest path algorithm is used to compute the path of each flow but link weights are computed using a particular cost function. Such cost function considers the state of each link, in terms of available capacity, and the capacity requested by the specific traffic flow.

The pseudo-code of the Start-SH&ReR algorithm is listed in Algorithm 1.

We define the variable $C_{i,j}^{Res}$ as follows.

$$C_{i,j}^{Res} = \sum_{s,d} (f_{i,j}^{s,d} * T_{s,d}) \tag{IV.3}$$

In (IV.3) $C_{i,j}^{Res}$ is the capacity reserved for all traffic flows established on the link *i,j*.

Lines from 1 to 6 correspond to the *Starting Phase*, which is equivalent to the single-hop bypass strategy described by eq. IV.1 and eq. IV.2. After this starting phase, the algorithm enters in the *Re-Routing Phase* (RRP) which can be executed several times. Each iteration of the RRP starts by ordering traffic flows according to a given criterion and choosing the specific cost function to be used. At each step, the next flow is selected and rerouted on the virtual topology using the cost function to compute link weights. Specifically, the rerouting process of each flow is composed of three phases:

i) flow removal, from line 12 to line 17;

ii) link weight assignment and path computation, from line 13 to line 20;

iii) flow establishment, from line 20 to line 28.

Algorithm1 – Start-SH&ReR algorithm's pseudo code

```
1:  /Starting phase/
2:  for all traffic requests s,d in trafficMatrix
3:      x_{s,d} = Ceil(T_{s,d}/ΔC);
4:      f_{s,d}^{s,d} = 1;
5:      C_{s,d}^{Res} = T_{s,d};
6:  end for
7:  /Rerouting phase/
8:  for n times /n iterations/
9:      []trafficRequestArray←ordering_criterion(trafficMatrix);
10:     costFunction←choose_cost_function();
11:     for all traffic requests s,d in trafficRequestArray
12:         /flow removal/
13:         for all links i,j of the path of the current request
14:             f_{i,j}^{s,d} = 0;
15:             C_{i,j}^{Res} = C_{i,j}^{Res} - T_{s,d};
16:             x_{i,j} = Ceil(C_{i,j}^{Res}/ΔC);
17:         end for
18:         /link cost and path computation/
19:         [][]linkCosts←cost_function(trafficMatrix);
20:         path_{s,d}←DijkstraAlg(linkCosts);
21:         /flow set up/
22:         for all links i,j of the new path of the current request
23:             f_{i,j}^{s,d} = 1;
24:             C_{i,j}^{Res} = C_{i,j}^{Res} + T_{s,d};
25:             x_{i,j} = Ceil(C_{i,j}^{Res}/ΔC);
26:         end for
27:     end for
28: end for
```

The *flow removal* phase simply updates the link state variables for each link *i,j* belonging to the current flow path. Specifically, the flow is removed putting the variable $f_{i,j}^{s,d}$ to zero for all links *i,j* belonging to the current path; then the reserved capacity $C_{i,j}^{Res}$ and the number of line cards $x_{i,j}$ of the same links are updated.

In the *link weight assignment and path computation* phase link weights are computed using the specific cost function (considering the new values of $C_{i,j}^{Res}$ and $x_{i,j}$ for all i,j links) and the new least cost path is computed. The path is computed considering the full network graph, i.e. every i,j links, regardless of the $x_{i,j}$ variable of each link. If the new path involves a link with no sufficient capacity, the number of line cards installed on that link is updated. In fact, the *flow establishment* phase sets up the flow by setting the variable $f_{i,j}^{s,d}$ equal to 1 for all links belonging to the new path; then the reserved capacity $C_{i,j}^{Res}$ and the number of line cards $x_{i,j}$ of the same links are updated.

The described algorithm can be seen as a general procedure and a number of different algorithms can be derived from it. In particular, the following parameters can be chosen: i) the traffic flow ordering criterion for the RRP; ii) the cost function for the RRP; iii) the number of iterations of the RRP. In particular, with respect to the number of iterations of the RRP, we observed that increasing it beyond four or five iterations does not lead to any performance improvement.

The flow ordering criterion and the cost function used in this work are presented in the next two subsections.

Ordering Criterion

The ordering criterion used in the algorithm determines the order in which traffic flows are processed to be rerouted. Here we describe the one that showed the best performance among those tested. It first orders nodes and then, for each node, orders the traffic flows towards other nodes. In particular, as node-ordering criterion, we consider the total traffic treated by a node, ordering nodes for increasing values of it. Traffic flows of the same node are then ordered following the same order computed for nodes. For instance, if the node order is *a-b-c-d*, then the flow sequence will be *a,b*; *a,c*; *a,d*; *b,c*; *b,d*; *c,d*.

Cost Functions

A class of cost functions named *Additional Power and Link Load* (AP&LL) is proposed. It considers the actual load when computes the cost of a link in addition to the number of additional LCs needed to support the specific traffic flow. Specifically, the link load is taken into account by the *Load Related Cost* (LRC) function depicted in Figure IV-1; it is based on the following considerations.

First, if a link has an available capacity close to ΔC its utilization should be avoided in order to try to eliminate one of its line card when other flow will be rerouted; second, if the available capacity of a link is close to zero, then it should not be used in order to avoid saturating it and then adding a new line card on that link during future steps. These two considerations are addressed by the decreasing and the increasing branches of the function, respectively. More in detail, the $\mathrm{LRC}_{\alpha,\beta_1,\beta_2}(\rho)$ function is defined as follows:

$$\mathrm{LRC}_{\alpha,\beta_1,\beta_2}(\rho) = \begin{cases} (1-\frac{\rho}{\alpha})^{\beta_1}; if\ \rho < \alpha \\ (\frac{\rho}{(1-\alpha)} - \frac{\alpha}{1-\alpha})^{\beta_2}; if\ \rho \geq \alpha \end{cases} \qquad \text{(IV.4)}$$

$$\rho = (C_{i,j}^{Res})\mathrm{mod}_{\Delta C} / \Delta C \qquad \text{(IV.5)}$$

The variable ρ is defined in (IV.5) and represents the normalized load of the last line card. The parameter α determines the point where the minimum cost falls. It can be observed that, by shifting the null point towards higher values, the function prefers saturating links rather than using those with available capacity, caring little of avoiding the addition on of new line cards. Since the goal of the RRP is to eliminate the most possible number of line cards, it is expected that higher values of α lead to higher efficiency. Nevertheless, we found out that varying such a parameter in the interval [1, 1/2] during different iterations of the RRP, leads to better performance and good trade-off between efficiency and network congestion.

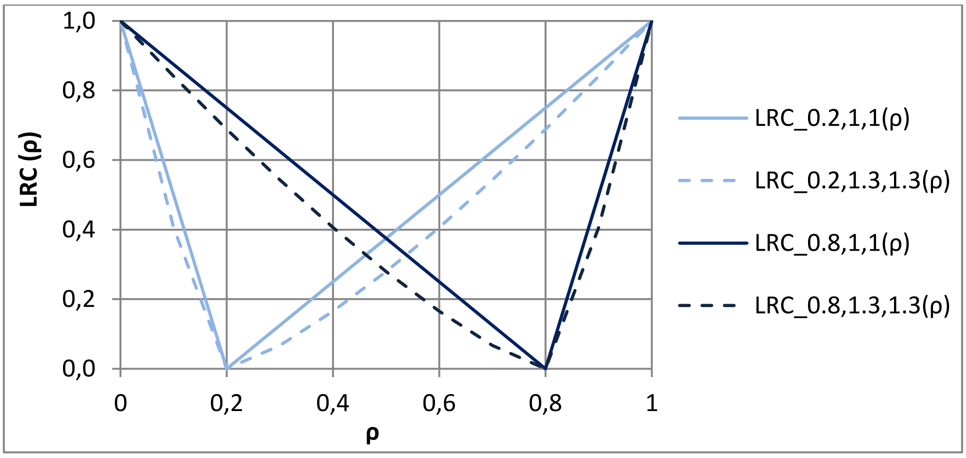


Figure IV-1: LRC functions for different values of the α, β_1 and β_2 parameters.

The parameter β_1 (β_2) rules the slope of the decreasing (increasing) branch, determining the average cost of a link: since ρ is always lower than 1, the higher β is the lower the average link cost is. It can be observed that the LRC function assigns a non null cost to each hop in the path, even if no additional line card is required. In this regard, the parameter β determines the average number of hops giving a cost equal to the cost of adding a new line card. The best performance is generally achieved with $\beta_1 = \beta_2 = 1$; however, when traffic flows are very smaller than the line card capacity, values of β lower than 1 can bring a little improvement.

The LRC function is used by the AP&LL function to compute the cost $LC_{i,j}$ related to a link *i,j* as follows:

$$LC_{i,j} = \frac{(\mathrm{LRC}_{\alpha,\beta_1,\beta_2}(\rho_{start}) + \mathrm{LRC}_{\alpha,\beta_1,\beta_2}(\rho_{end}))}{2} + N_{i,j}^{Add} \tag{IV.6}$$

$$\rho_{start} = (C_{i,j}^{Res}) \bmod_{\Delta C} / \Delta C;$$

$$\rho_{end} = (T_{s,d} + C_{i,j}^{Res}) \bmod_{\Delta C} / \Delta C. \tag{IV.7}$$

In (IV.6) $LC_{i,j}$ is the cost of the link *i,j* that depends on the particular traffic flow under rerouting. In fact, the terms ρ_{start} and ρ_{end} in (IV.7) can be seen as the capacity reserved on the last line card of the link *i,j* respectively before and after establishing the traffic flow. Finally, $N_{i,j}^{Add}$ is the number of additional line cards needed on the link *i,j* to accommodate the traffic flow.

Simulation Results

To evaluate the performance of the proposed solution randomly generated traffic demands have been considered. Each traffic demand is represented by an *N*x*N* matrix $[T_{s,d}]$, where *N* is the number of core routers in the network. Each element $T_{s,d}$ of the traffic matrix represents the traffic relation between nodes *s* and *d*. The parameter $T_{s,d}$ is randomly chosen according to a uniform distribution in the interval [*Min_T*, *Max_T*], where:

$$Max\_T = \gamma \cdot Min\_T = 30 \cdot SF \text{ Gbits/s.} \tag{IV.8}$$

$$SF = 2^{i};\ i = -4, -3, ..., 4. \tag{IV.9}$$

In (IV.8) the parameter γ represents the maximum variability among traffic flows, assumed to be equal to 5, whilst *SF* is a multiplicative parameter, named *Scaling Factor*, that is used to increase the total traffic treated by the network and it has been varied according to (IV.9). Moreover, it is assumed that for each node pair *s,d* a bidirectional traffic relation exists, so that $T_{s,d} = T_{d,s}$. We define the *Average Normalized Load* (*ANL*) parameter as ratio between the average capacity requested by a traffic flow and the capacity of a line card; in detail:

$$ANL = Avg\_T/\Delta C = Min\_T \cdot (\beta+1)/(2 \cdot \Delta C). \tag{IV.10}$$

The proposed algorithm has been compared with the single-hop bypass strategy and with the heuristic proposed in [22], here referred to as *Start-Empty*. Results related to the MILP formulation presented in III.2 are also provided for evaluating the energy efficiency ηe. Specifically, it has been defined as ratio between the power consumption related to the optimal solution, PC^{ILP}, and the power consumption determined by each heuristic PC^{i}, as in (IV.11).

$$\eta_e^{i} = PC^{ILP}/PC^{i} \tag{IV.11}$$

The proposed algorithm is used with the ordering criterion described above and performing five iterations of the RRP decreasing the parameter α from 0.9 up to 0.5.

Figure IV-2 shows the power saving, with respect to the single-hop heuristic, of the optimal solution, the *Start-Empty* and the *Start-SH&ReR* algorithms for different values of traffic load, for a network with 30 nodes; Figure IV-3, instead, shows the energy efficiency, as defined in (IV.11).

As it can be expected, the power saving decreases as the *ANL* grows, for both the heuristics and the optimal solution. In particular, all heuristics, included the single-hop bypass strategy, lead to the optimal solution when all traffic flows are greater than ΔC. The *Start-SH&ReR* algorithm outperforms the *Start-Empty* algorithm for any traffic load, especially for values of *ANL* lower than 0.45. For such values of traffic load, in fact, the energy efficiencies of the two algorithms are quite distant. In this zone, the power saving of the proposed algorithm with respect to the *Start-Empty* algorithm is equal to about 55%, for *ANL*=0.03, and decrease down to 27% for *ANL*=0.23.

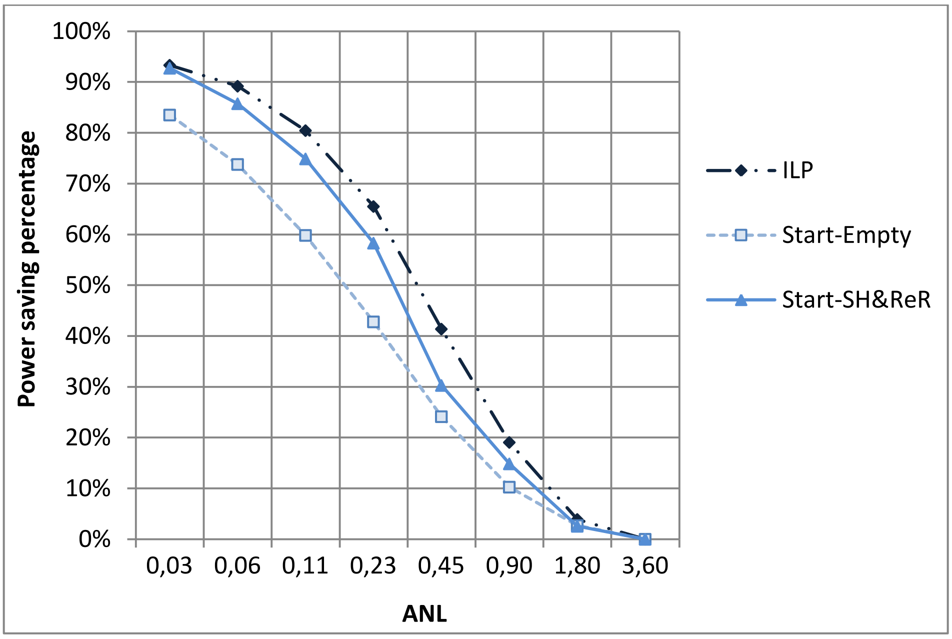


Figure IV-2: Percentage power saving of the Start-SH&ReR and the *Start-Empty* heuristics and the optimal solution with respect to the single-hop heuristic, versus *ANL*; N=30;

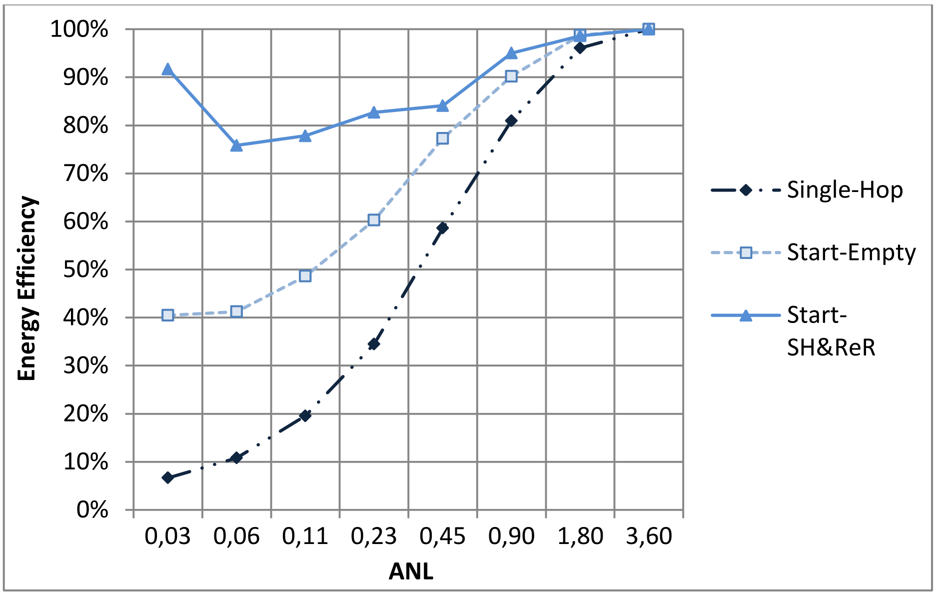


Figure IV-3: Energy efficiency of Start-SH&ReR, *Start-Empty* and single-hop heuristics, versus *ANL*; N=30.

For even higher loads, all heuristics rapidly tends to the optimal solution, and the proposed algorithm saves less than 10%. The performance difference is due to the opposite approach followed by the two algorithms. *Start-SH&ReR* decides the flow routing when the IP topology has not yet been entirely constructed; in this way, some flows are established on paths longer than needed. Moreover, the construction of a new link in the IP topology may be decided during the routing of a flow requesting less capacity than another one routed on a multi-hop path. The *Start-SH&ReR* algorithm, instead, starts considering the topology given by the single-hop heuristic and then tries to eliminate links by rerouting traffic on the rest of the network. Thanks to the AP&LL function the algorithm not just optimize the network consumption but even leads to a better resource usage.

We evaluated the distribution of the load experienced by links in the network; results are presented in Figure IV-4 for different values of traffic load. It can be seen that each heuristic has a characteristic behaviour until *ANL* is lower than 1.8; then all of them show the same distribution. Specifically, for the *Start-SH&ReR* algorithm the link capacity usage is never lower than 65% and the percentage of links experiencing a load higher than 85% rapidly decreases. As expected, the single-hop heuristic leads to have most links underutilized when *ANL* is low. Finally, the *Start-Empty* algorithm always saturates most links.

As additional performance parameter, we evaluate the path length distribution shown by the two multi-hop heuristics; results are shown in Figure IV-5.

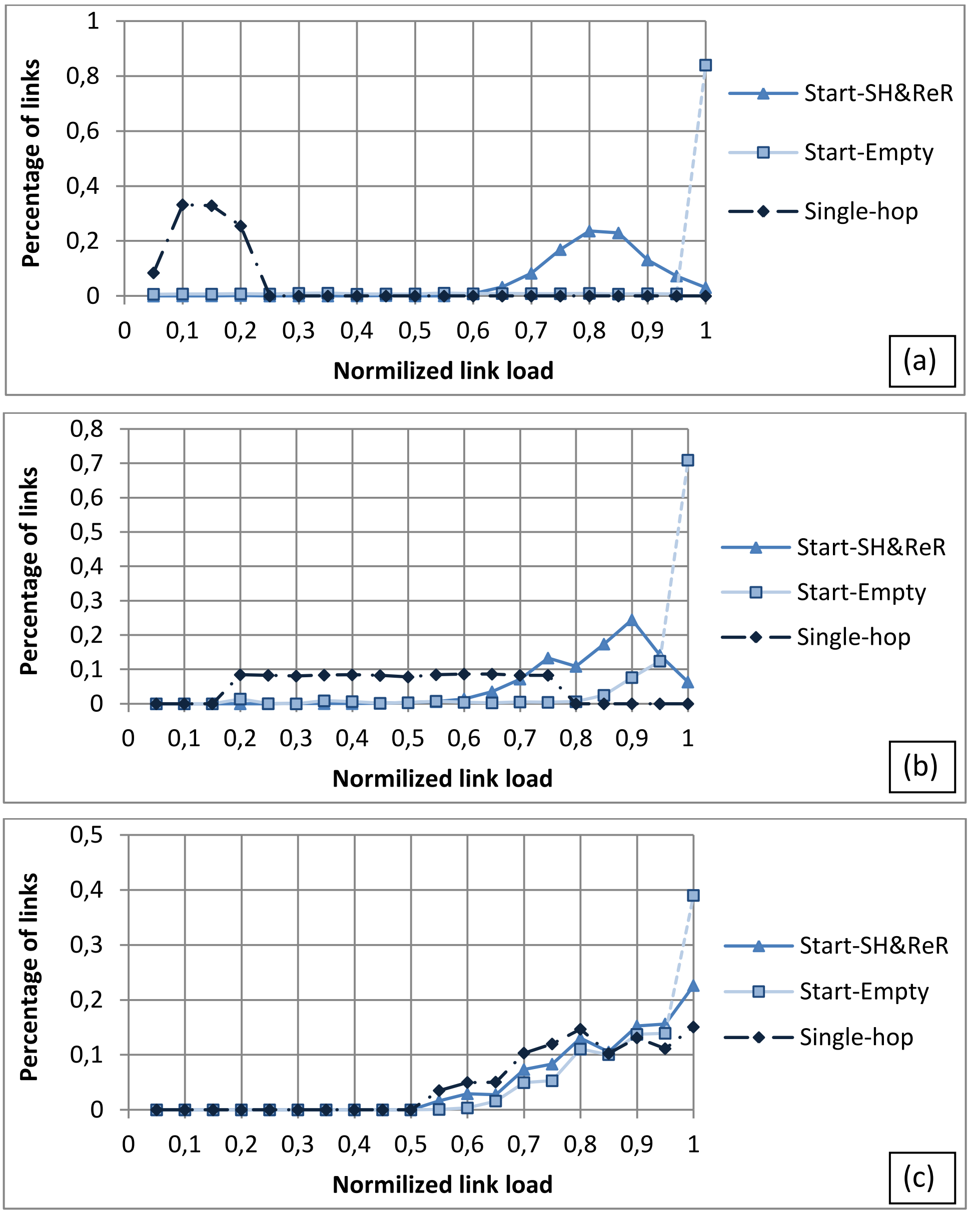


Figure IV-4: Capacity usage distribution of the Start-SH&ReR, the *Start-Empty* and the single-hop heuristics, for: a) *ANL*=0.11; b) *ANL*=0.45; c) *ANL*=1.8.

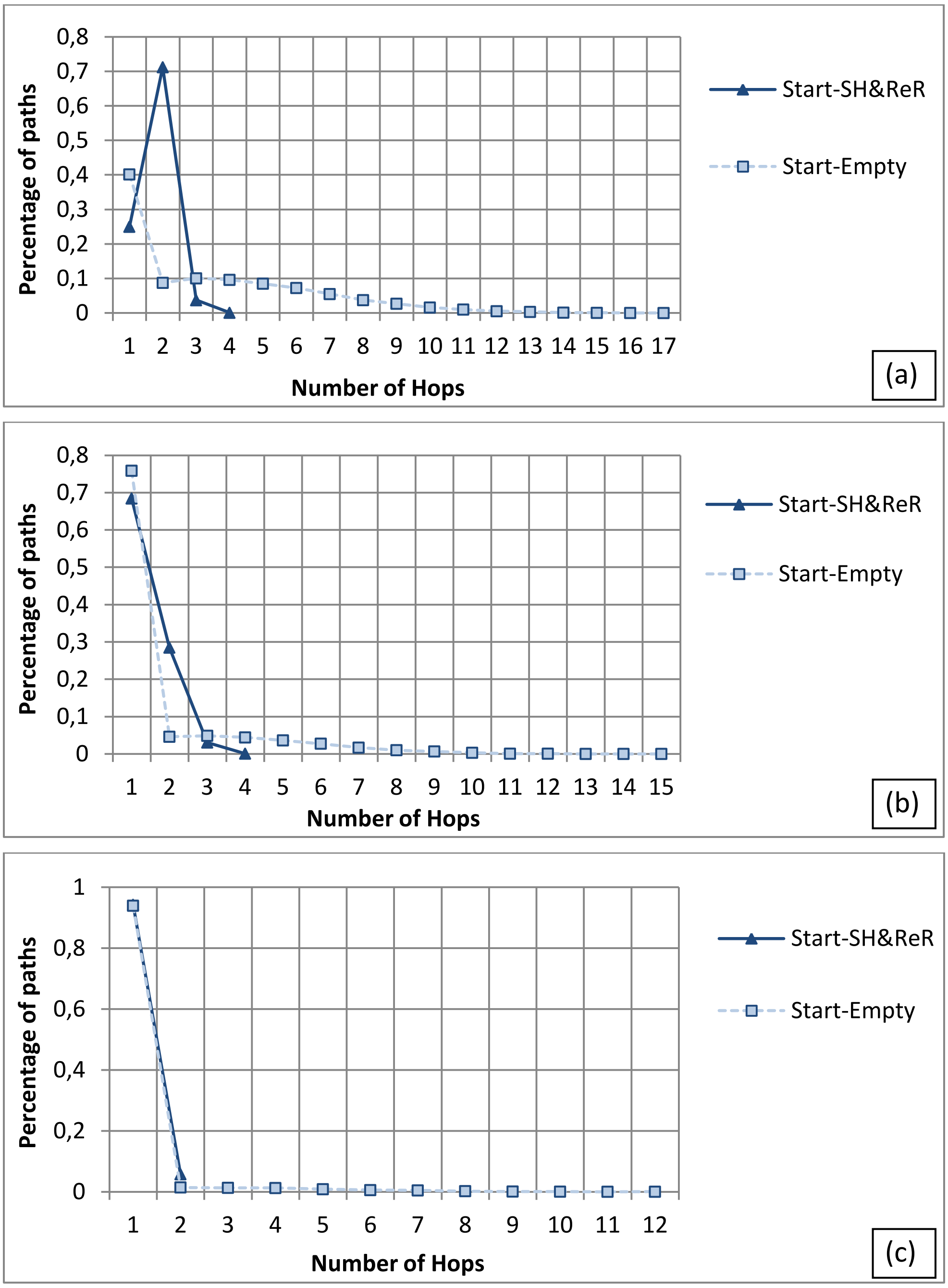


Figure IV-5: Path length distribution of *Start-SH&ReR* and *Start-Empty* heuristics, for: a) *ANL*=0.11; b) *ANL*=0.45; c) *ANL*=1.8.

The *Start-SH&ReR* algorithm allows to route flows on paths no longer than four hops, for any value of traffic load; clearly, when *ANL* grows, the percentage of single-hop paths increases. The very short tail of the path length distribution is mainly due to the use of the LRC function. In fact, it account the cost related to each hop, even if no additional interfaces are required for establishing a given traffic request. In other words, the cost function considers the cost related to the capacity usage, giving a different weight depending on the state of the link. Since the *Start-Empty* heuristic does not account the cost of each hop in the network, it routes traffic on longer paths, showing a distribution with a very long tail with some paths routed even on 15 hops. This determines a capacity waste, leading to a higher number of needed line cards and then to a higher power consumption.

As last performance analysis, we evaluate the energy efficiency of the proposed heuristic with respect to the network dimension, varying the number of nodes from *N*=10 up to *N*=35. Figure IV-6 and Figure IV-7 show results for *ANL*=0.11 and *ANL*=0.45. For *ANL*=0.11 the efficiency of the *Start-Empty* heuristic decreases more quickly and is equal to 50% for a network with 20 nodes. For *ANL*=0.45 the energy efficiency of the *Start-SH&ReR* heuristic, as well as the power saving with respect to the single-hop heuristic, remains almost constant as *N* grows and it is always higher than 80%. The efficiency of the *Start-Empty* algorithm instead slowly decreases with *N*.

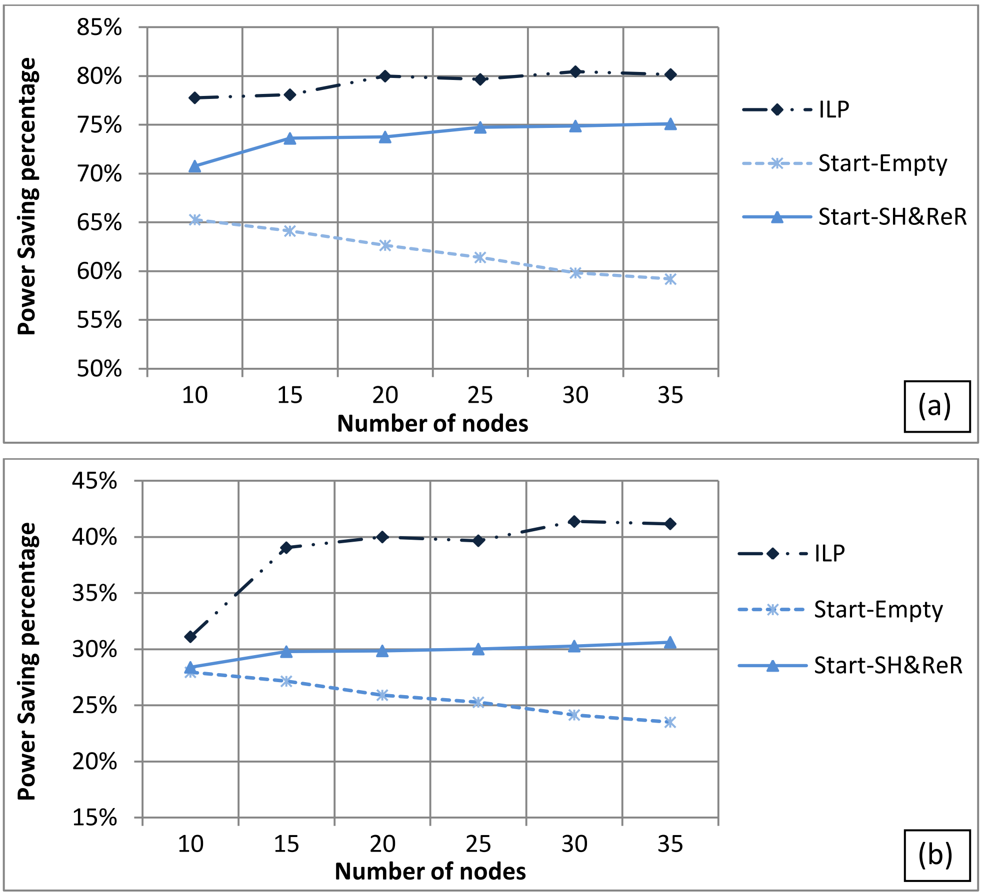


Figure IV-6: Percentage power saving of the Start-SH&ReR and the *Start-Empty* heuristics and the optimal solution with respect to the single-hop heuristic, versus *N*, for: a) *ANL*=0.11; b) *ANL*=0.45.

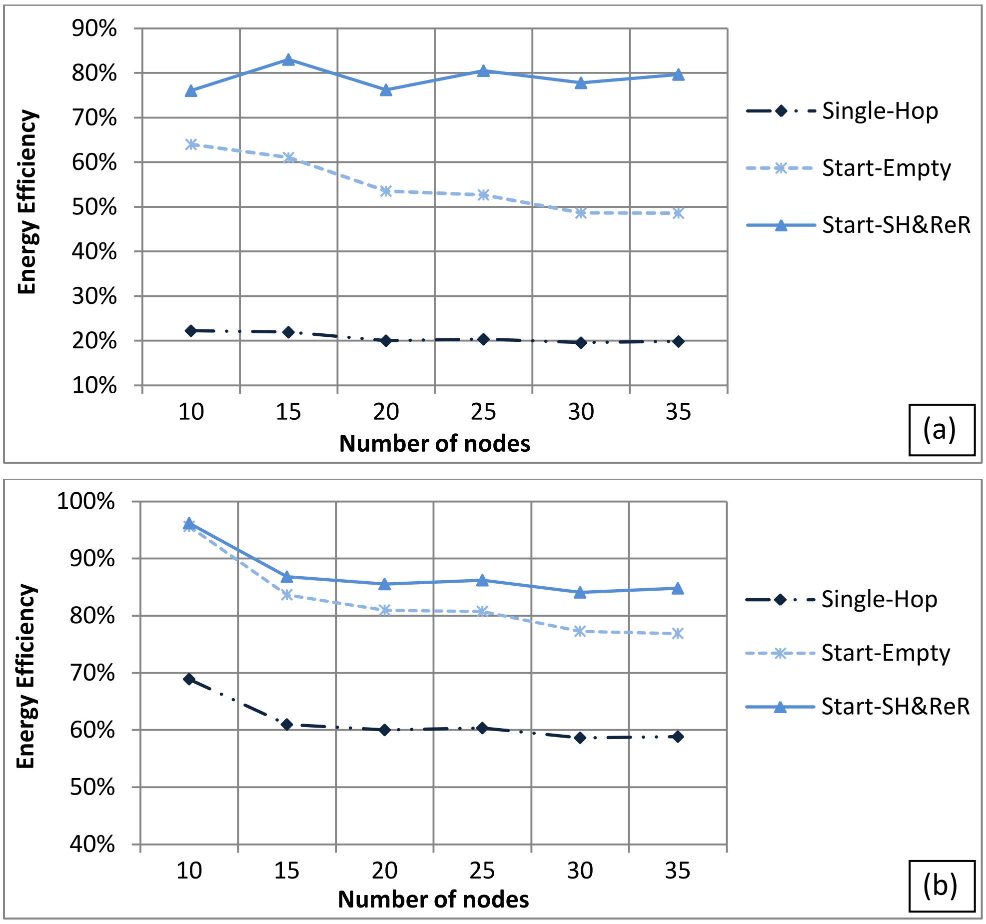


Figure IV-7: Energy Efficiency of Start-SH&ReR, *Start-Empty* and single-hop heuristics versus *N*, for: a) *ANL*=0.11; b) *ANL*=0.45.

IV.1.2 *"DAISIES": an EA-TE Strategy*

Although the EM-VTD strategy is able to drastically improve the energy efficiency of the network, it does not consider the variation experienced by traffic flows in time. In fact, as already discussed in III.1, since the traffic changes during a day following a typical day-night trend, it is possible to further reduce the network energy consumption putting in idle some network devices when they are not strictly needed pursuing energy efficiency in network operation. Specifically, at the IP layer, EA-TE strategies focus on aggregating traffic flows on a subset of links in the network in order to put in sleeping mode unused line cards (see III.3.2).

EA-TE strategies involve the definition of dynamic mechanisms to efficiently switch off nodes and links when the traffic demand slows down in order to make the power consumption of the network proportional to the carried traffic. However, besides the need of defining efficient heuristics to decide the set of devices to put in idle, such mechanisms should also consider several issues related to their practical applicability, such as routing instabilities, packet loss or out-of-order delivering, network congestion, etc., that may be experienced during the switch-off of network devices.

The EA-TE technique proposed here is named Distributed and Adaptive Interface Switch-off for Internet Energy Saving (DAISIES); it is presented with reference to the MPLS technology to explicitly signal the path of each traffic flow and to manage the traffic in the network. However, it must be pointed out that, for the purpose of our solution, any other packed-switched connection-oriented technology might be used, such as PBB-TE or MPLS-TP.

The proposed solution is based on a TE approach and requires that every IP traffic flow is carried by an LSP. In such a way, many practical issues can be overcome. First of all, it is possible to have complete knowledge about the load state of each link in the network since such information can be advertised by the OSPF-TE protocol considering the bandwidth reserved by each LSP. Moreover, thanks to the MPLS technology, it is possible to assure that neither packet loss nor out-of-order packet delivering is experienced during traffic rerouting. Unlike the solution proposed in [33], we propose a heuristic approach that can assure a good efficiency in a reasonable time. Two important aspects differentiate our solution from the ones discussed in III.3.2. The first one concerns the approach we adopt to optimize the network: instead of following a link-based approach, i.e. deciding to switch off a link and then performing the consequent traffic rerouting, we adopt a flow-based approach. This means that our solution looks for the best path for each traffic flow, considering the state of the whole network, whilst the actual link switch-off is a consequence of routing decisions. In fact, it mainly acts on the routing algorithm, which, by considering the state of each link in the network, implicitly decides the active topology. To do this, our solution uses a particular cost function into the Dijkstra algorithm to compute the weight of each link considering their state in terms of available capacity. Another important novelty of our solution is that it is completely distributed since each node in the network is responsible of the traffic flows of which it is the source; however, the path of each flow is computed considering the state of every link and with the aim at optimizing the whole network. In this way, the solution does not need an additional communication protocol to disseminate routing decisions from a centralized entity that performs the computation to nodes that must actually reroute the traffic and switch off links. The combination of the flow-based approach, the distributed architecture and the use of the MPLS technology also allows to locally manage the link switch-off process, in a fashion that results transparent to other network nodes. The proposed solution works as follows.

In the network, a Label Switching Path (LSP) is configured between each pair of Label Switching Routers (LSRs); in this way all IP packets coming from outside the network are carried by a given LSP. Thus, there are $N \cdot (N-1)$ unidirectional LSPs in the network, where N is the number of LSRs; these LSPs are routed on the active topology.

The MPLS layer plays two different roles. The first one is strictly related to the energy saving mechanism that will be described in the following; in fact, as already stressed, our solution uses dynamic load-dependent link weights inside the routing algorithm. A basic assumption is that the MPLS layer has to provide a full knowledge of the load state of each link. This information is advertised by the TE routing protocol in the whole network and can be used to rearrange the path of some LSPs (and thus the path of the traffic carried by them) to optimize the network. Without the connection-oriented intermediate layer we could not use the link load information into the routing algorithm and we would act on the link weights advertised by the IGP routing protocol to change the path of the traffic in the network, as proposed in [31], [32]. However, in such a way, routing instability might be experienced in the network during the convergence period of the routing protocol if no additional mechanism is defined; moreover, we could not decide to change the path of just some traffic flows since the whole SPT would be recomputed by the IGP routing protocol.

The second role executed by the MPLS technology is to decouple the LSR topology and the logical topology that is actually seen by IP entities. The last, in fact, is given by the set of Label Switching Paths (LSPs) established in the network and is independent from the LSR topology. Since the set of LSPs does not change in time the same occurs for the IP forwarding table, even if the path of an LSP is changed or a line card is switched-off; what changes when an LSP is rerouted or a line card is switched off is the MPLS forwarding table, whilst, at the IP layer, each destination network is still reached trough the same logical link that is represented by the LSP connecting the local node with the node which the destination network is directly connected to. In this way, the SPT has not to be recomputed when a line card is switched off or reactivated.

Each link i,j of the LSR topology is composed of a number of line cards of the same capacity ΔC. Each link is advertised in the network as a bundled TE-Link grouping the information about all physical resources (line cards), which the LSR link is composed of, in the same logical entity. This can be done by exploiting the link bundling concept provided by the MPLS technology [43].

The capacity of each TE-Link can be modelled as follows:

$$C_{i,j} = \Delta C \cdot N_{i,j}^{Max} \tag{IV.12}$$

In (IV.12) $C_{i,j}$ represents the maximum capacity that can be activated on a given link i,j that is given by the number $N_{i,j}^{Max}$ of line cards available for that link.

A TE-link is described in the network by its cost W, its capacity $C_{i,j}$, the parameter ΔC, that must be the same for each line card bundled in the same TE-link [43], and by its state; all this information must be advertised by the TE-routing protocol. In particular the state of a link is represented by the available capacity $C_{i,j}^{Avl}$ that is the capacity still available for setting up other LSPs considering idle line cards too:

$$C_{i,j}^{Avl} = C_{i,j} - C_{i,j}^{Res} \tag{IV.13}$$

In (IV.13) $C_{i,j}^{Res}$ is the capacity reserved for all LSPs established on the link i,j at a given time.

The basic idea of the proposed solution is to reroute each LSP a number of times during a day considering the state of every link and according to the variation of their requested capacities; the goal of the rerouting process is to empty as many line cards as possible so that the related nodes can switch them off. In this way, the active topology, i.e. the set of active line cards, is decided by the routing algorithm and the line card switch-off process can be locally managed.

Each IP layer traffic relation is carried by an $LSP_{s,d}$ established between nodes *s* and *d*. $C_{s,d}^{Req}$ (*t*) represents the capacity requested by the $LSP_{s,d}$ at the time *t* to carry the amount of traffic $f_{s,d}(t)$. Each node *s* in the network is responsible for monitoring the IP layer traffic entering in the $LSPs_{s,d}$, i.e. the LSPs of which it is the source. An update of the capacity requested by the LSP (LSP update in the following) is then locally generated whenever the traffic flow becomes lower or bigger than a prefixed threshold. Two different ways for computing the updating threshold will be proposed later. Clearly, in order to prevent a continual updating of LSPs, the traffic should be monitored on an adequate time scale (several seconds) so as to hide short-periods (milliseconds or less) traffic oscillations.

Whenever an update occurs for the $LSP_{s,d}$, the node *s* re-computes the path of that LSP by means of a least cost path algorithm in which a specific cost function is used to compute link weights on the basis of their state. The new computed path may be the same path which the LSP is already established on or a different path; in both cases, the same procedure is executed by the node *s* to update the capacity reserved by the LSP and its path. Such a procedure, here called LSP rerouting, may be performed by means of the MPLS make-before-break mechanism [44] which allows to reroute an LSP without neither experiencing flow interruption nor duplicating the reserved bandwidth. In detail, a new LSP is created and set up on the new path sharing the capacity with the old LSP; then the old one is torn down.

The line card switch-off/activation process is locally performed by each node involved in the LSP rerouting process. Specifically, each node *i* involved in the set up process of the new LSP checks if the LSP can be supported on the link *i,j* with the current active capacity; i.e. if $C_{i,j}^{on} \geq C_{i,j}^{Res} + C_{s,d}^{Req}$; if not, it activates a new line card on the link *i,j*. On the contrary, each node involved in the tear down process of the old LSP checks if the difference between the active and the reserved capacities, on the link *i,j*, is greater than ΔC; in this case, a line card can be switched off on that link. The switch-off process is performed so that the active available capacity $C_{i,j}^{Act\_Avl}$, i.e. the available capacity considering only active line cards, is always smaller than ΔC. It can be computed as follows:

$$C_{i,j}^{Act\_Avl} = C_{i,j}^{Avl} \bmod_{\Delta C} \tag{IV.14}$$

Finally, nodes advertise $C_{i,j}^{Avl}$ for all impacted links. Note that, for links belonging to both the new and the old path the available capacity changes only after the new LSP has been set up, if the LSP capacity is increased, or, if the LSP capacity is decreased, only after the old LSP has been torn down. This is due to the fact that the two LSPs share the capacity on the two paths. Figure IV-8 shows a finite state machine representation of DAISIES procedures.

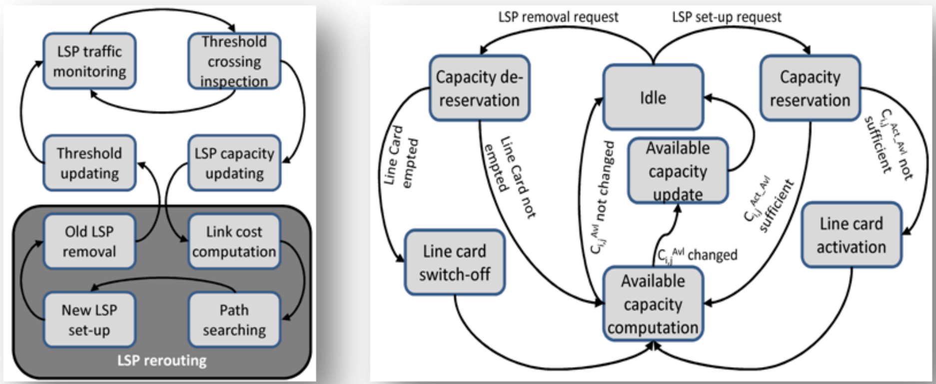


Figure IV-8: Finite state machine representations of a) LSP rerouting, b) line card switch-off/activation.

It can be noted that the LSP rerouting process performed by a node *s* may determine the switch-off and the activation of a line card on any link *i,j* in the network. This represents the main novelty of the proposed solution: traffic flows are not rerouted as a consequence of the decision to switch off a line card but the opposite occurs; i.e. the decision to reroute a traffic flow may determine the switch-off of a line card. In this way the active topology is decided by the routing algorithm and topology changes take place in a transparent way. Thanks to the make-before-break mechanism no packet loss is experienced during the rerouting of a traffic flow since the old LSP is torn down only when the new one has been established; moreover, out-of-order packet delivery can be avoided by starting the transmission of packets on the new LSP after the old one has been torn down, preserving the good operation of the TCP.

It is worth outlining that the time needed to perform a rerouting depends on the number of hops of the new and the old paths, on the line card bit rate, which determines the transmission delay of RSVP packets, and on the control plane load of each involved router; however, it is limited to some hundreds of milliseconds or less, since it practically corresponds to set up an LSP and tear down another one, and it is negligible with respect to the average time between two updates of the same LSP that is expected to be some tens of minutes or more.

The proposed solution can be applied to all the three scenarios defined in III.3.2. In the FRFT scenario the path of the LSP is not recomputed when an update occurs but just its capacity is updated. The LSP capacity updating, as well as any other LSP attribute updating, is performed by means of the make-before-break mechanism in an MPLS network; thus, the same procedure must be executed by the source node in all scenarios. In the DRDT scenario, line cards deployed on a node i can be assigned to any link i,j when needed. Thus, the capacity $C_{i,j}$ advertised for the TE-Link i,j, which represents the maximum capacity that can be activated on it, must be fixed to the total capacity installed on the node i. However, since the total capacity used on a node cannot exceed the installed capacity (see eq. III. 24), the available capacity of each link *i,k* must be decreased (increased) of ΔC by the node *i* whenever a line card is assigned to (removed from) a link *i,j*.

The main advantages of the proposed solution can be summarized in:

i) making the line card switch-off/activation process transparent to the routing protocol;

ii) avoiding packet loss and out of order delivering;

iii) automatically managing both traffic decreasing and increasing;

iv) requiring no change to current routing and signalling protocols.

Capacity Updating Methods

In order to manage the traffic variation, it is necessary to fix a set of thresholds beyond which the capacity of the LSP is actually updated. Specifically, at any instant an upper and a lower thresholds are considered. The upper threshold is represented by the current value of requested capacity, i.e. $C_{s,d}^{Req}$; in fact, soon as $f_{s,d}(t)$ becomes bigger than $C_{s,d}^{Req}$, the capacity requested by the $LSP_{s,d}$, and consequently the upper threshold, has to be updated to a higher value in order to not exceed the bandwidth reserved for the LSP in the network. When $f_{s,d}(t)$ crosses the upper threshold and $C_{s,d}^{Req}$ is updated, the lower threshold is updated to the previous value of requested capacity; instead, when $f_{s,d}(t)$ crosses the lower threshold, $C_{s,d}^{Req}$ is updated to the current value of the lower threshold, and the lower threshold is updated.

We propose two different methods to compute the new threshold values when these are crossed. The first updating method is named *Fixed*

Threshold (FT); it considers the parameter ΔF. When $f_{s,d}(t)$ crosses the upper (lower) threshold, $C_{s,d}^{Req}$ is increased (the lower threshold is decreased) of the quantity ΔF. The parameter ΔF determines the capacity updating rate. Specifically, a lower ΔF allows traffic variations to be better followed, thus increasing the efficiency of the mechanism; however, it implies a higher complexity in terms of computational and signalling loads. It can be noted that, when the FT method is used, the number of updates is different for each LSP; in particular, the larger the LSP is the higher the number of updates is. This might determine a significant performance reduction if, in order to decrease the complexity, the value of ΔF is chosen so large that smaller LSPs are updated too few times or not updated at all.

To overcome this problem we propose a different updating method named *Percentage Threshold* (PT); it considers the parameter $\%F$. When $f_{s,d}(t)$ crosses the upper (lower) threshold, $C_{s,d}^{Req}$ (the lower threshold) is scaled of a factor $1/(1\text{-}\%F)$ $(1\text{-}\%F)$. Note that the parameter $\%F$ plays the same role as the parameter ΔF in FT, determining the updating rate. However, when the PT updating method is used, the number of updates of an LSP depends only on the dynamics of the related traffic flow and not on its maximum dimension. If all traffic flows experience the same percentage variation during a day, as assumed in this study (see sec. VII), all the LSPs will be updated the same number of times.

In practice, it may be possible that some traffic flows experience a large variation whilst others vary much less or do not vary at all. In this case, other updating methods could be used, such as considering a different $\%F$ for each LSP or considering a maximum time interval beyond which the path of an LSP is recomputed even if the carried traffic has not changed. Such refined methods are not considered in this paper.

Cost Functions

In order to save energy, the path of an LSP has to be computed with the goal to minimize the total number of used line cards; to achieve such an objective, we use a simple least cost path algorithm that, in turn, uses a cost function to compute link weights. Such cost functions take in input the cost W, $C_{i,j}$, $C_{i,j}^{Avl}$, ΔC, and $C_{s,d}^{Req}$. Note that the cost of each link depends on the particular LSP under rerouting (the quantity $C_{s,d}^{Req}$ is considered). Since all line cards are assumed of the same type and capacity, in this paper the term W is assumed to be the same for all TE-links and it is implicitly assumed equal to 1; however, it might be fixed to a different value for each link if different kinds of line cards are available for each of them or might be decided on the basis of other considerations.

For links belonging to the path of the LSP under rerouting, the term $C_{i,j}^{Avl}$ that is used by cost functions is evaluated considering the LSP under rerouting as removed. More in detail, for any link i,j belonging to the current LSP path, it results:

$$C_{i,j}^{Avl} = C_{i,j}^{Avl'} + C_{s,d}^{Req'} \tag{IV.15}$$

In (IV.15) $C_{s,d}^{Req'}$ is the capacity currently reserved for the $LSP_{s,d}$, $C_{i,j}^{Avl'}$ is the available capacity advertised by the OSPF-TE protocol whilst $C_{i,j}^{Avl}$ is the available capacity actually considered for the link i,j. It must be pointed out that the Dijkstra algorithm, currently used by the OSPF routing protocol, can also be used in DAISIES without any modification. Finally, the number of hops is used by the routing algorithm to break ties.

As far as cost functions are concerned, it must be outlined that any plausible function should preliminarily evaluate whether the total capacity available on a link i,j, i.e. considering idle line cards too, is sufficient or not to support the specific LSP. In the last case the specific link must be excluded from the network graph that is used by the Dijkstra algorithm to search the path. Two different cost functions are proposed in the present work, namely OAP and AP&LL; both of them consider the quantity $N_{i,j}^{Add}$ that represents the number of line cards that have to be activated to set up a given $LSP_{s,d}$; it is computed as follows:

$$N_{i,j}^{Add} = \mathrm{Ceil}((C_{s,d}^{Req} - C_{i,j}^{Act\_Avl})/\Delta C). \qquad \text{(IV.16)}$$

The first proposed cost function is named *Only Additional Power* (OAP) since it only considers the additional power needed to support a given traffic request. Specifically, it counts the number of LCs that must be activated on a link *i,j* to establish a given LSP; thus it results $LC_{i,j} = N_{i,j}^{Add}$.

The OAP cost function is very straightforward: it chooses the shortest path with enough active capacity, if any, otherwise the path requiring the minimum additional power is chosen; this is a general principle that is often followed by power-aware routing algorithms, such as those in [22], [37]. However, in this way, the actual link load is not taken into account and OAP is not able to preventively reroute traffic to facilitate the line card switch-off or to avoid activating a new line card in the future. Moreover, by applying a null cost to links with sufficient capacity, LSPs could be routed on very long paths determining a potential bandwidth waste. We already showed the drawback of such an approach in IV.1.1 where the EM-VTD problem has been faced. For such a reason, the same class of cost functions used in the *Start-SH&ReR* algorithm, i.e. AP&LL, is also used in DAISIES to compute link weights. Specifically, the AP&LL function can be defined with reference to the new variable defined here as follows.

$$LC_{i,j} = \frac{(\mathrm{LRC}_{\alpha,\beta_1,\beta_2}(\rho_{start}) + \mathrm{LRC}_{\alpha,\beta_1,\beta_2}(\rho_{end}))}{2} + N_{i,j}^{Add} \qquad \text{(IV.17)}$$

$\rho_{start} = (\Delta C - C_{i,j}^{Act\_Avl})/\Delta C;$

$$\rho_{end} = ((\Delta C - C_{i,j}^{Act\_Avl} + C_{s,d}^{Req})\text{mod}_{\Delta C})/\Delta C. \qquad \text{(IV.18)}$$

In this context AP&LL is useful because it is able to reroute LSPs in a number of situations; in particular, it has been designed to preventively reroute LSPs in order to increase the probability to switch off a line card in the future. The novelty of this function is that it considers the cost related to utilization of the capacity in the network and not just to the additional power consumption.

Complexity and Scalability

The computational complexity of the proposed solution is mainly related to the total number of LSP updates generated in the network. In both DRDT and DRFT scenarios, the Dijkstra algorithm must be executed and link weights must be computed whenever an update occurs; instead, in FRFT the path does not have to be recomputed. Then, in all the scenarios, the rerouting/capacity updating process is performed by means of two different RSVP PATH/RESV message exchanges, one to set up the new LSP and one to tear down the old one. The RSVP message exchange represents an additional signalling overhead but also an additional computational load for nodes that must process such messages. In detail, each RSVP message is processed by H nodes, where H is the number of hops of the related LSP. Finally, once the new LSP has been established and the old one torn down, the new state of all the involved links has to be advertised by related nodes that must generate an OSPF LSA for each of those links; yet again, such messages represents an additional computational load and signalling overhead. The number of links that change their state after the rerouting of an LSP is at least equal to H_{old}, if the path is not changed, and at most equal to $H_{new} + H_{old}$, if the new and the old paths are link disjoint, where H_{new} and H_{old} are the number of hops of the new and the old path respectively.

Concerning the scalability of our solution, it can be noted that it is related to the number of LSPs in the network that grows as N^2, where N is the number of nodes in the network. Thus, fixed the average number of updates for each LSP, the total number of updates in the network grows as N^2. However, since the solution is completely distributed, the computational load that each node must support grows as N.

A possible way for improving the scalability of our solution is to exploits OSPF routing areas as follows. A full mesh of LSPs is configured inside each OSPF area, including the backbone area. Then inter-area LSPs are nested inside intra-area LSPs and only higher order LSPs are updated. If the maximum number of LSRs in each OSPF area is fixed to the value C, there will be a total of $N_a = N/C$ different areas. For each area an LSR is connected to the backbone area that will be composed of N_a routers. In this scenario, the computational load on each LSR not belonging to the backbone area remains constant as N grows since C LSPs are established in each area. Each LSR of the backbone area, instead, has to manage N_a-1 LSPs with the other backbone LSRs and C-1 LSPs with LSRs belonging to its own area. This means that the complexity continues to grow as N but it is reduced of a factor C.

Performance Evaluation

Simulations have been carried out using a proprietary software that computes the time instants when LSPs have to be updated considering a traffic matrix generated as described in the following, simulates the LSP rerouting and the consequent line card switch off/activation and computes the total energy consumed during a whole day; the software does not consider single packets flowing in the network but just aggregated traffic flows. At the starting instant, the software takes in input the solution given by the ILP formulation that represents the BN.

Each traffic demand used in simulations is represented by a NxN matrix $[f_{s,d}(t)]$, where N is the number of core routers in the network. Each element $f_{s,d}(t)$ of the traffic matrix represents the traffic relation, at the IP layer, between the nodes s and d and it is assumed variable in time so as to model the traffic variation during a day. For sake of simplicity a simple sinusoidal trend is assumed for each traffic relation in this study; then each traffic flow $f_{s,d}(t)$ is given by:

$$f_{s,d}(t) = f_{s,d}^{Tmax}\,[(1\text{-}\sigma)/2\cdot(1 + \sin(f_0 t)) + \sigma];\ \ f_0=1/24 \text{ hours.} \qquad \text{(IV.19)}$$

In (IV.19) $f_{s,d}^{Tmax}$ represents the peak traffic between the nodes s and d whilst σ represents the ratio between the off-peak and the peak traffic loads; it is assumed σ=0.2. The parameter $f_{s,d}^{Tmax}$ is randomly chosen according to a uniform distribution in the interval $[Min\_f^{\,Tmax}, Max\_f^{\,Tmax}]$, where:

$$Max\_f^{\,Tmax} = \gamma\cdot Min\_f^{\,Tmax} = 30\cdot SF \text{ Gbits/s.} \qquad \text{(IV.20)}$$

$$SF = 2^i;\ i = -4, -3, ..., 4. \qquad \text{(IV.21)}$$

In (IV.20) the parameter γ represents the maximum variability among traffic flows, fixed to 5, whilst *SF* is a multiplicative factor, named *Scaling Factor*, used to scale the total traffic offered to the network and it has been varied according to (IV.21). Moreover, it is assumed that for each node pair *s,d* a symmetric bidirectional traffic relation exists, so that $f_{s,d}(t) = f_{d,s}(t)$. We define the *Average Normalized Peak Load* (*ANPL*) parameter as ratio between the average peak capacity requested by a traffic flow and the capacity of a line card; in detail:

$$ANPL = Avg\_f^{Tmax}/\Delta C = Min\_f^{Tmax}\cdot(\beta+1)/(2\cdot\Delta C). \qquad \text{(IV.22)}$$

In (IV.22) *ANPL* is the average peak traffic normalized to the line card capacity.

We consider a network with 20 nodes. The capacity ΔC of each line card has been fixed to 40 Gbits/s; it is assumed that idle line cards have null power consumption; however, it is easy to evaluate results in case that the idle power is not null. In fact, if S is the percentage of power saving, it means that on average a percentage S of line cards are in idle state whilst a percentage equal to 1-S are active. Then, if $\varepsilon < 1$ is the fraction of power absorbed by a line card in idle state, the average power saving S' can be simply computed as $S'=S(1-\varepsilon)$.

The performance of the proposed solution has been evaluated as average power saving with respect to the case of a static *Base Network* (BN) (see III.3.2) designed by solving the MILP formulation provided in III.3. Moreover, in order to give an absolute benchmark, we consider a lower bound (LB) to the total energy consumed in a day evaluated as follows. The optimization problem is solved for five different time instants during a half day, scaling the related traffic matrix. Then a time-continuous power consumption profile is obtained by means of a polynomial interpolation and finally the daily average power consumption is computed. Such a lower bound has been used to evaluate the energy efficiency defined as ratio between the LB and the power consumption given by the proposed solution. The two const functions described above have been used in both DRDT and DRFT scenarios and the FRFT scenario has been evaluated too. For AP&LL we fixed $\alpha = 0.75$, $\beta_1 = \beta_2 = 1$.

Power Saving Vs. Traffic Load

The first set of results are related to the average power saving and the energy efficiency achieved by the proposed solution as the *ANPL* parameter varies; they are shown in Fig. 4 where the PT updating method is used with

%F=0,06. The *ANPL* parameter plays a key role for the performance of the proposed solution, as well as for any other link switch off mechanism, since it determines how meshed the topology of the BN is and consequently what percentage of links can be switched off. The lowest considered value is such that all the traffic originated and terminated in a node can be supported by just one line card; in this case the optimal topology of the BN is a star and no links can be switched off. The highest one, instead, is such that the smallest flow at its off-peak value is larger than the line card bit-rate; in this case the optimal solution is to keep the network always full-meshed adapting the number of line cards of each link when the related traffic flow slows down.

Observing results in Figure IV-9, it can be noted that the DRDT scenario leads to just a little improvement with respect to DRFT, confirming results obtained in [27]. This result is due to the fact that the topology of the BN is an optimal topology; thus, the probability that when the traffic slows down the optimal topology includes a link not belonging to the topology of the BN is low. In other words, we do not need to activate a link not belonging to the BN in order to optimize the network when the traffic slows down. This is also due to the fact that all traffic flows varies in the same way in our scenario; this means that the ratio between the capacities requested by any two traffic flows remains the same during a day. It is expected that a more appreciable difference between DRDT and DRFT might be experienced if the daily variation was different for each traffic flow. The percentage of power saving increases with *ANPL*, for all scenarios and for the LB. In fact, as *ANPL* grows, the BN becomes more meshed and the percentage of links that can be switched off without disconnecting the network increases.

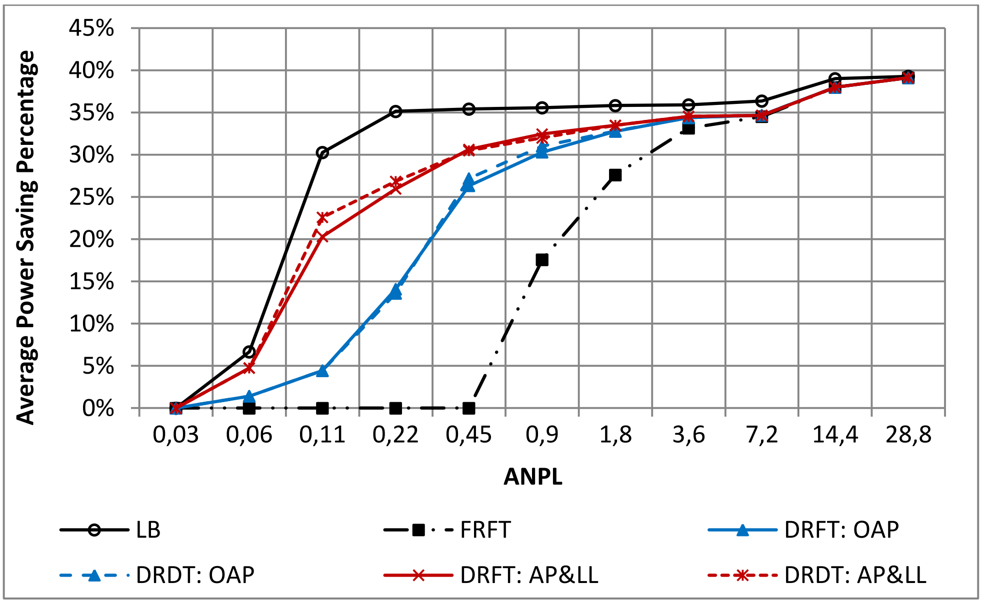


Figure IV-9: Average power saving vs. *ANPL*. PT updating method with *%F*=0.06.

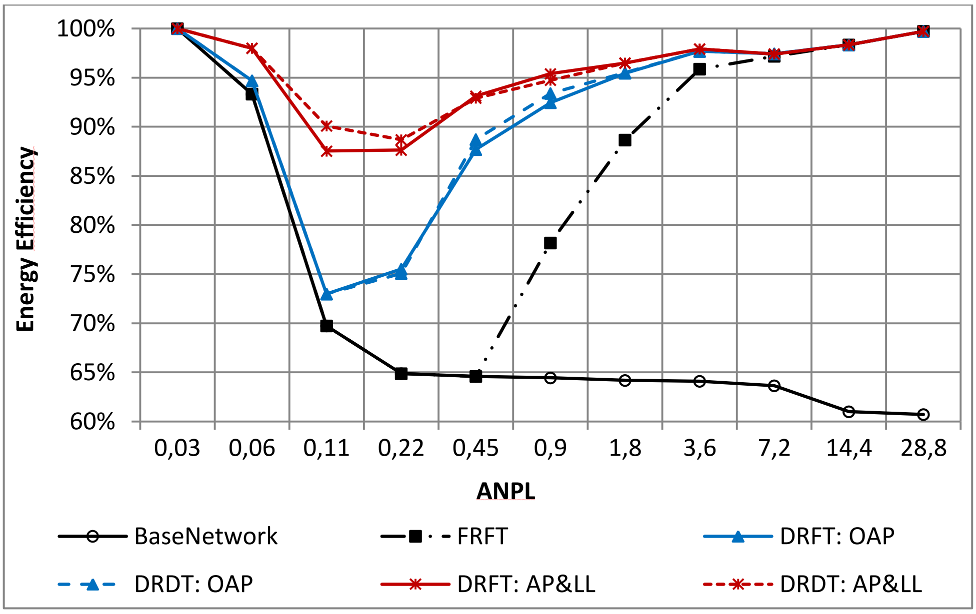


Figure IV-10: Energy Efficiency vs. *ANPL*. PT updating method with *%F*=0.06.

As a consequence, the energy efficiency of the BN monotonically decreases with *ANPL* and tends to 60% since each traffic flow requires during a day an average capacity equal to 60% of the peak capacity. For the heuristic solution, instead, the energy efficiency has a non-monotonic trend. In fact, when *ANPL* is very low, the maximum achievable power saving, i.e. the LB, is low and even the BN has high energy efficiency. In this situation there is only a little opportunity of saving energy. On the contrary, when *ANPL* is very high, the energy saving opportunity increases very much (the BN has a low energy efficiency) and a good solution is easy to be found since a considerable saving can be achieved by acting on links with multiple line cards. Instead, for medium values of *ANPL* a good solution is harder to be found even if the energy efficiency of the BN is low. Specifically, for $0.11 \leq ANPL \leq 0.45$, AP&LL performs much better than OAP whilst FRFT does not lead to any saving. In fact, in this region, all traffic flows are smaller than ΔC and only one line card is deployed on each link; for such a reason, it is not possible to switch off a line card without rerouting LSPs established on the related link. As far as the OAP performance is concerned, it can be observed that, since traffic flows are small compared with ΔC, most links support a lot of LSPs; for such a reason, in most cases, the rerouting of just one LSP does not lead to switch off a line card and thus OAP does not change the path of the LSP. The AP&LL function, instead, is able to preventively reroute an LSP even if this does not lead to a line card switch-off; in this way the probability to switch off a line card in the future, when other LSPs on the same link will be rerouted, is increased. The minimum value of energy efficiency achieved by AP&LL in this region is about 88%; thus, a further 12% of saving is needed to reach the lower bound. For higher values of *ANPL* ($0,9 \leq ANPL \leq 3,6$) OAP performs very close to AP&LL and its energy efficiency is between 93% and 97%; this occurs because some links support just one LSP and thus the related line card can be switched off by rerouting such an LSP when it requires less capacity. Other links are composed of two line cards so that one of them can be easily switched off by updating the LSP capacity. In fact, in this region, FRFT starts to save energy and quickly reaches the same performance as AP&LL and OAP. In fact, as the traffic increases, even more links are composed of more than one line card and the switch off is mainly due to the capacity updating rather than to the rerouting process. Finally, for *ANPL* $\geq$ 7.2, both cost functions and FRFT show the same performance. In fact, most traffic flows are larger than ΔC also during the off-peak time and rerouting them is never convenient.

Figure IV-11shows the average path length experienced by LSPs (LSP-APL) versus *ANPL*, for the BN, OAP and AP&LL in the DRFT scenario. It also shows the average value of the path length of each LSP weighted for its capacity. Such a parameter represents the APL experienced by unit of traffic in the network; thus, it can be seen as the APL experienced by each packet (Packet-APL) that is even more important than the LSP-APL since it gives a real indication of the performance of the network; moreover, it also determines the total network load. It can be seen that, for the BN, the LSP-APL slows down as *ANPL* grows. Specifically, it is equal to about two hops for *ANPL*=0.03, when the BN has a star topology, and decreases to one hop for *ANPL* larger than or equal to 3.6, when the BN is full-meshed. Both OAP and AP&LL determine a higher LSP-APL since, by switching off links, they decrease the connectivity degree of the network. Concerning the two cost functions, it can be noted that, for *ANPL*$\in$ [0.03, 0.45], AP&LL determines a higher LSP-APL with respect to OAP. This is a consequence of the higher energy-efficiency achieved by AP&LL in this region. For *ANPL* $\in$ [0.45, 7.2] such a trend is inverted even if AP&LL still outperforms OAP in terms of energy efficiency. The Packet-APL is lower than the LSP-APL for both OAP and AP&LL and for the BN. This is a consequence of the fact that larger LSPs are generally routed on shorter paths.

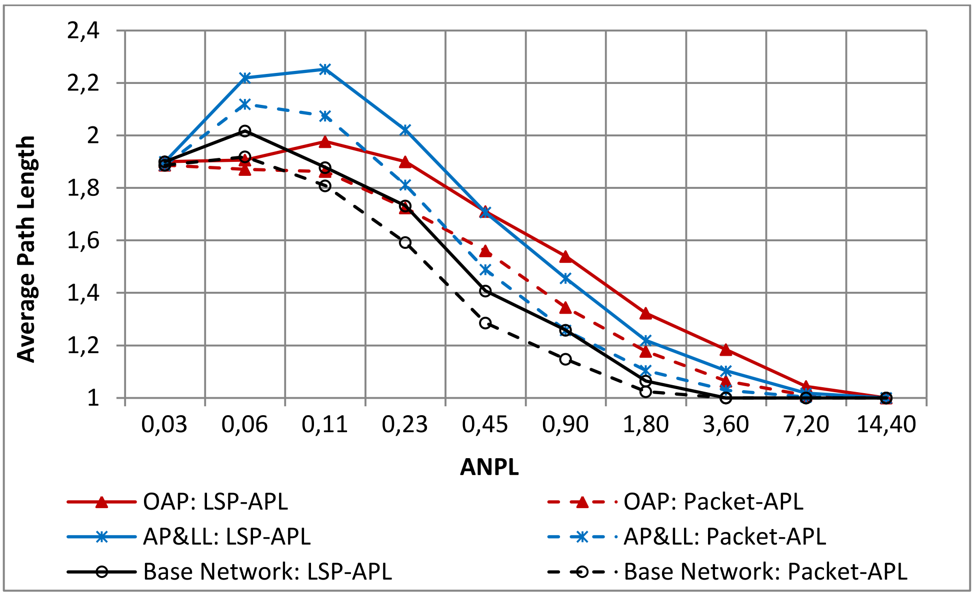


Figure IV-11: LSP-APL and Packet-APL vs. *ANPL* for OAP, AP&LL and the Base Network.

As already stressed, the Packet-APL determines the total network load; this, jointly with the total capacity kept active in the network, determines the average link load (ALL) experienced in the network which is another important parameter of the network performance. However, the ALL assumes two opposite meanings depending on the performance parameter which we are interested in. Classical power-unaware routing strategies aim at maximizing the load that the network can support. This means that, fixed the offered load, i.e. the traffic matrix, they aim at minimizing the resultant network load and, given that all the capacity deployed in the network is always kept active, at minimizing the ALL. Such a goal can be easily achieved by means of the shortest path routing algorithm. Power-aware routing strategies, instead, aim at minimizing the capacity kept active in the network thus increasing the ALL; generally, the higher the energy efficiency is the higher the ALL is. When comparing different power-aware routing solutions, the congestion level can be used only as a secondary performance parameter; in fact, if two solutions experience a different performance in terms of energy efficiency, the congestion level cannot be used as a comparison parameter. However, if their performance is the same, the solution that leads to a lower congestion is to be preferred since reducing the congestion level means reducing the queuing delay experienced by packets in the network. However, since the queuing delay grows more than linearly with the load of a link, the minimization of the ALL does not assure the minimization neither of the average nor of the maximum queuing delay. In order to limit the maximum queuing delay experienced by a packet in a router, a common practice is to use a maximum link usage threshold, say α. This is equivalent to consider the capacity $\Delta C^1=\alpha\cdot\Delta C$ as the line card capacity. In this case, the performance of the proposed mechanism can be evaluated as follows. First, observe that the performance depends only on *ANPL* and not on the actual level of traffic and ΔC; moreover, considering the threshold α is equivalent to consider a line card of capacity $\Delta C_1=\alpha\cdot\Delta C$ since the same value of α would be also considered for the BN and the power absorbed by each line card does not change. Then, let $P_1(TL_1)$ and $P_2(TL_2)$ be the power consumption in case of ΔC_1 and ΔC_2, respectively, with $TL_1=Avg\_f_1^{Tmax}/\Delta C_1$, $TL_2=Avg\_f_2^{Tmax}/\Delta C_2$ and $\Delta C_2=\alpha\cdot\Delta C_1$. Since $P_1(TL_1)=P_2(TL_2)$ for $TL_1=TL_2$, it results $Avg\_f_1^{Tmax}=Avg\_f_2^{Tmax}/\alpha$ that corresponds to consider the traffic level increased of a factor 1/α.

Figure IV-12 shows the ALL versus *ANPL* for the *BN*, FRFT and the two cost function in the DRFT scenario. Concerning the BN, the ALL tends to 60% as *ANPL* grows. In fact, the average offered load is equal to 60% of the peak load and for high *ANPL* values the ALL is close to 100% during peak hours. FRFT increases the ALL with respect to the BN only for *ANPL*

> 0.45, when it is able to switch-off line cards and thus to decrease the active capacity. Concerning the two cost function, instead, for *ANPL* ≤ 0.45 AP&LL leads to a higher ALL that is a direct consequence of the higher power saving; however, for higher values of *ANPL* it reduces the ALL even if it still saves more energy than OAP. This is possible because it reduces the Packet-APL too.

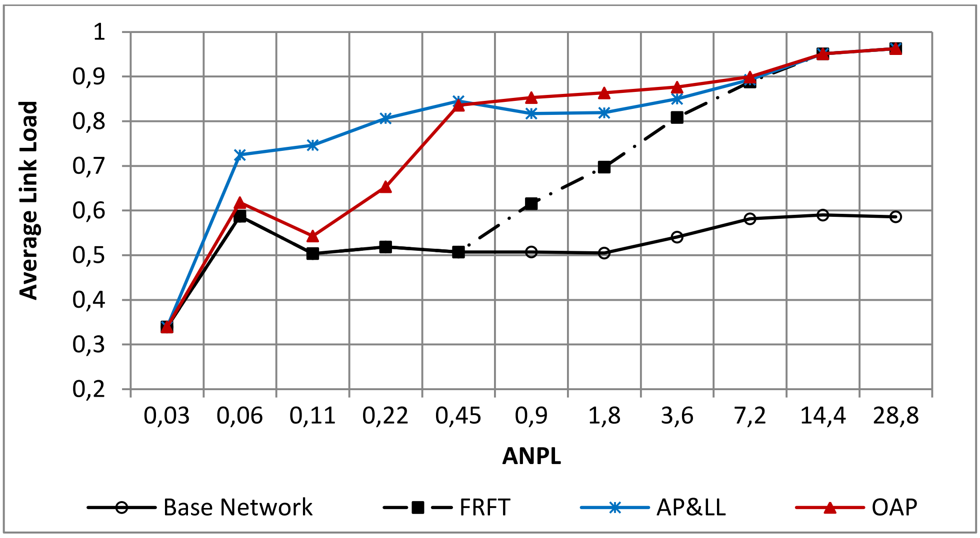
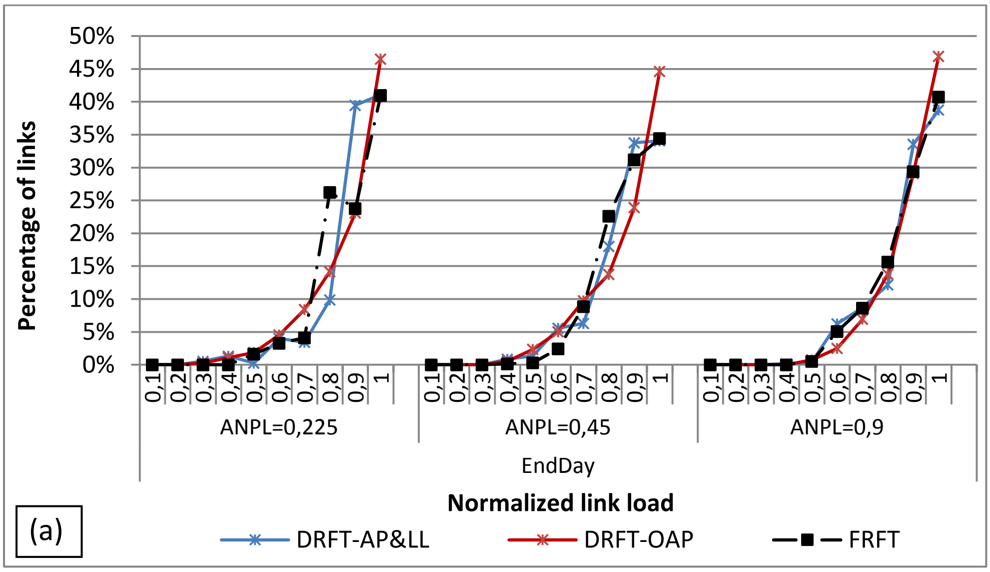


Figure IV-12: Average Link Load vs. *ANPL*, for OAP, AP&LL, FRFT and the Base Network.

Figure IV-13 shows the link load distribution (LLD) experienced by the two cost functions when the traffic load is at the peak and the off-peak values, for different *ANPL*. It can be seen that, during peak hours, the two distributions are very similar to each other. The only difference is that OAP leads to a little higher percentage of links at the maximum load with respect to AP&LL and FRFT; in fact, thanks to the increasing branch of the AP&LL function, it is unlikely that many links experience a load close to 100%. For *ANPL* ≤ 0.45, OAP is not able to switch off many line cards when the traffic slows down and the LLD related to the off-peak traffic shifts towards lower values. When *ANPL* is high and OAP performs close to

AP&LL in terms of energy efficiency, it leads to a higher percentage of links experiencing a maximum load during off-peak hours.



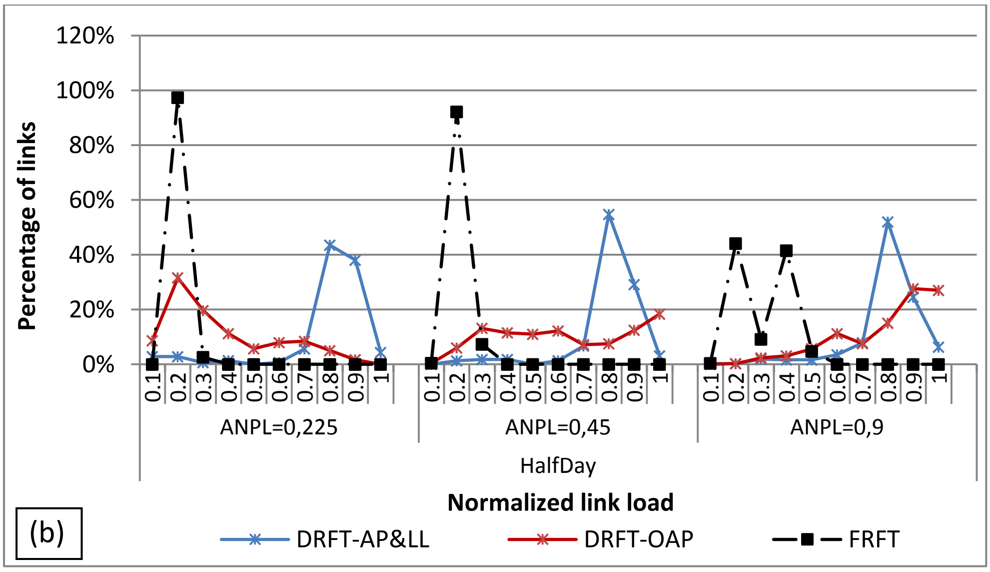


Figure IV-13: Comparison of the Link Load Distribution of FRFT, OAP and AP&LL for different *ANPL* values: a) traffic at the peak and b) traffic at the off-peak.

Efficiency Vs. Complexity

The complexity associated with the proposed solution has been discussed in Sec VI.D where we outlined that it is related to the additional RSVP and OSPF signalling needed to perform the LSP rerouting mechanism and to the execution of the Dijkstra algorithm needed to compute the new path. Thus, the total complexity is mainly due to the number of updates generated in the network which has an important impact on the performance of our solution too.

Before analysing the relationship between the efficiency and the complexity of DAISIES, it is worth outlining the impact that the traffic load, the updating method and the particular cost function used have on the signalling overhead. The number of RSVP PATH/RESV messages generated during each rerouting depends on the number of hops of the new and the old paths of the particular LSP; in case of PT, the average number of such messages is equal to two times the LSP-APL since every LSP is updated the same number of times. In case of FT, since the number of updates of each LSP is proportional to its capacity, the average number of RSVP messages generated for each LSP update is related to the Packet-APL and thus it is lower than in case of PT. Clearly, as outlined in the previous section, both the LSP-APL and the Packet-APL depends on *ANPL* and the cost function. The number of OSPF LSA, instead, also depends on how similar the new and the old paths of a given LSP are. The average number of RSVP and OSPF messages generated for each LSP update is shown in Figure IV-14.

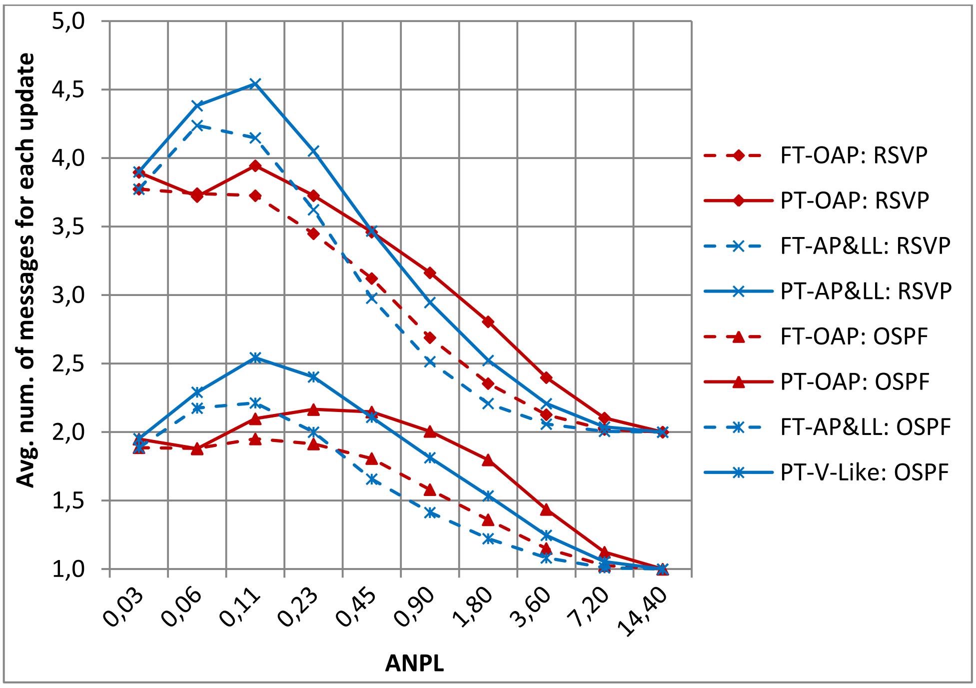


Figure IV-14: Average number of RSVP and OSPF messages generated during each LSP update vs. *ANPL*, for OAP and AP&LL and for the PT and the FT updating methods.

The complexity associated with the proposed solution mainly depends on the total number of LSP updates which in turn can be adjusted by tuning the updating threshold parameter, i.e. the ΔF or the *%F* parameter defined before. In fact, such parameters regulate the average number of updates generated for each LSP whilst the total number simply depends on the number of LSPs. If too many updates are generated, the impact on the router performance may be non-negligible. However, the higher the number of LSP updates is, the better DAISIES work. Figure IV-15, Figure IV-16 and Figure IV-17 show the average energy saving and the energy efficiency versus the average number of updates of each LSP for *ANPL*=0.23, *ANPL*=0.9 and *ANPL*=3.6, respectively. It can be seen that as the updating rate decreases below to about 12 updates for LSP, the performance of the proposed mechanism rapidly slows down, loosing most its efficiency. On the contrary, increasing it beyond about 22 updates for LSP does not lead to a significant gain. Such a trend is more evident for large *ANPL* values: for *ANPL*=3.6 passing from 14 to 80 updates for each LSP makes the energy saving grow of just a further 5%. As expected, the PT updating method outperforms FT, especially when the number of updates is low; however, for high *ANPL* values no difference is experienced between them. For *ANPL*=0.9 the performance of FRFT are almost constant if LSPs are updated at least 5-6 times during a day. In fact, for this *ANPL* value, only a small fraction of links is composed of two line cards; thus, updating LSPs when such line cards have already been powered off is quite useless.

The work has been published in [58].

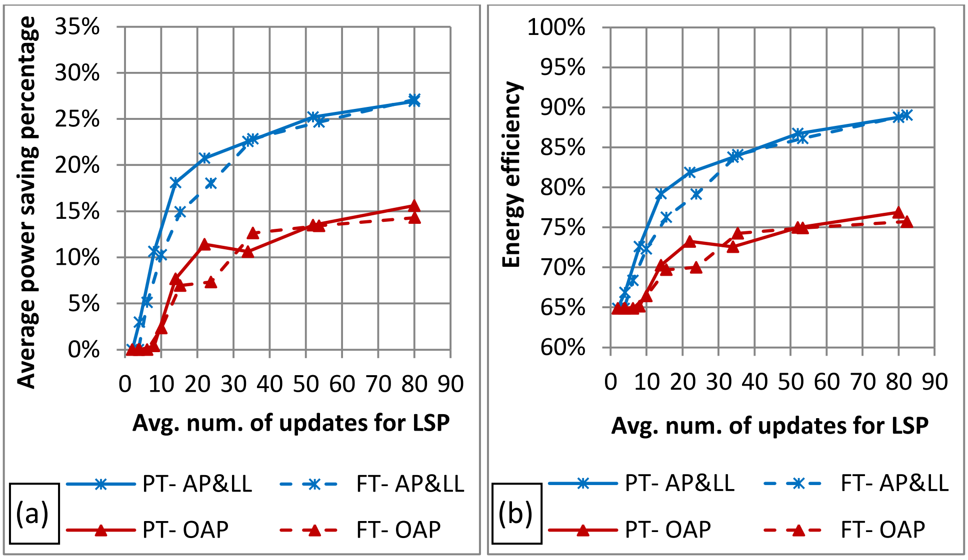


Figure IV-15: (a) Average power saving (b) energy efficiency vs. the average number of updates for each LSP, for OAP and AP&LL and for the PT and the FT updating methods, *ANPL*=0.23.

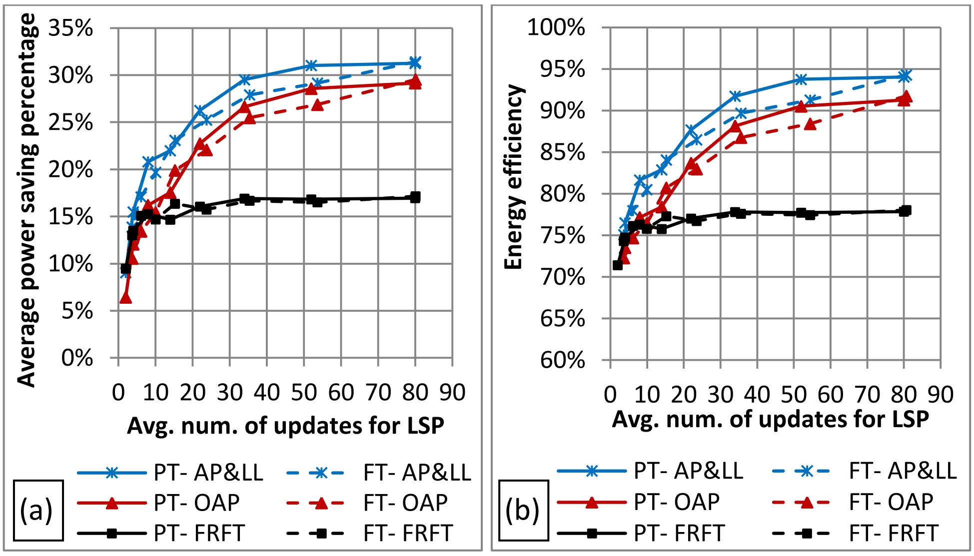


Figure IV-16: (a) Average power saving (b) energy efficiency vs. the average number of updates for each LSP, for OAP and AP&LL and for the PT and the FT updating methods, *ANPL*=0.9.

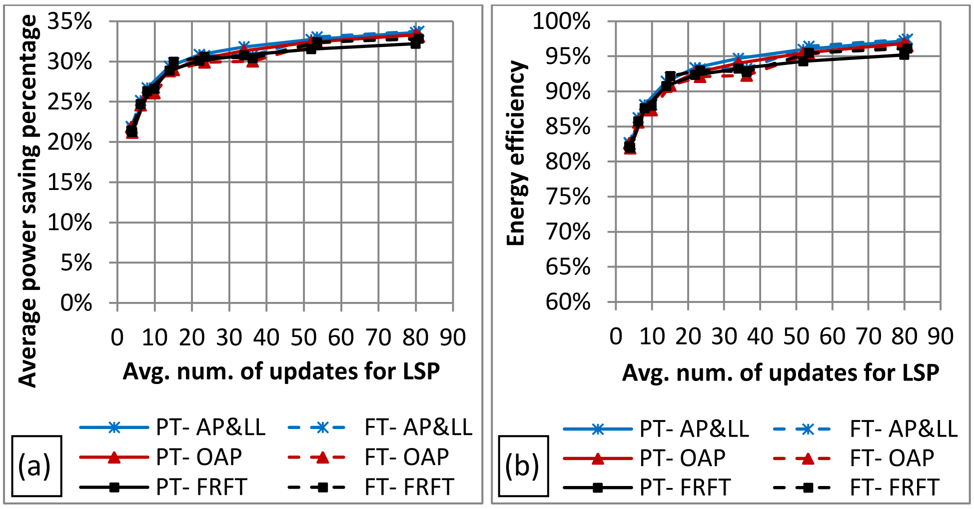


Figure IV-17: (a) Average power saving (b) energy efficiency vs. the average number of updates for each LSP, for OAP and AP&LL and for the PT and the FT updating methods, *ANPL*=3.6.

IV.2 Energy Saving at the Optical Layer

In III.4 the energy saving issue in a transparent circuit-switched optical network has been analyzed. Specifically, it has been outlined that the main objective is to keep unused as many optical fibre links as possible in order to put in idle state optical amplifiers deployed along them. Moreover, two different strategies have been identified: fibre switch-off and Power-Aware RWA.

Both strategies have been considered here and two different solution have been proposed to achieve energy saving in optical networks. They are presented in the next two subsections.

IV.2.1 *An Iterative Greedy Algorithm for Switching Off Optical Fibres*

The approach proposed here is based on an iterative greedy algorithm derived from the approach proposed in [28]. The basic idea is to order network elements according to a given heuristic criterion and to select, at each step, the next element trying to switch it off. The network graph is modified eliminating the selected element; then, if flow and capacity constraints are met, the selected node or link is powered off and the traffic is rerouted on the modified network graph. It is assumed here that only links can be switched off whilst nodes cannot because every node is, with high probability, source and destination of traffic flows.

Network Model

The network model presented in II.3.2 is considered. It is a transparent circuit-switched optical network without wavelength conversion functionality. Each optical fibre link is a bidirectional link composed of F fibres for each direction (F>0).

Connection requests are assumed to be known in advance; each connection is fulfilled by setting up a primary path, the working path, and a disjoint backup path used for protection purpose. Each path is set up by means of a lightpath, from the source node to the destination node, using the same wavelength on each link across the path. Direct and inverse paths are co-routed. Primary paths are protected by means of backup paths using Shared Backup Path Protection (SBPP): a wavelength is shared between different backup paths whose correspondent primary paths are link-disjoint.

To solve the Routing and Wavelength Assignment (RWA) problem, the Iterative Two-Step-Approach (ITSA) [46] is used for the routing problem while the wavelength is assigned in two different ways for the working and protection paths. The ITSA algorithm provides a pair of two disjoint paths on the basis of the network state but does not provide any solution to the WA problem under wavelength continuity constraint. First Fit (FF) algorithm [47] is used to assign the wavelength to the working path, whilst, for the backup path, the wavelength that is sharable on the most links is assigned.

Heuristic Algorithm

The flow diagram of the proposed algorithm is shown in Figure IV-18. It starts by ordering fibres according to a given criterion; at each step the next fibre is selected as the candidate to be switched off. This means that all connections established on such a fibre have to be rerouted. Connections are rerouted one by one using the GMPLS graceful shutdown mechanism [48] to avoid flow interruption during the rerouting process. If all connections using the selected fibre can be rerouted the algorithm restarts by re-ordering fibres; otherwise the next fibre is selected. The algorithm ends when all fibers have been inspected.

It is worth outlining that, even if working and protection paths of the same connection cannot be both affected by the fiber switch off process, both the primary and the back-up paths are recomputed by means of the RWA algorithm adopted.

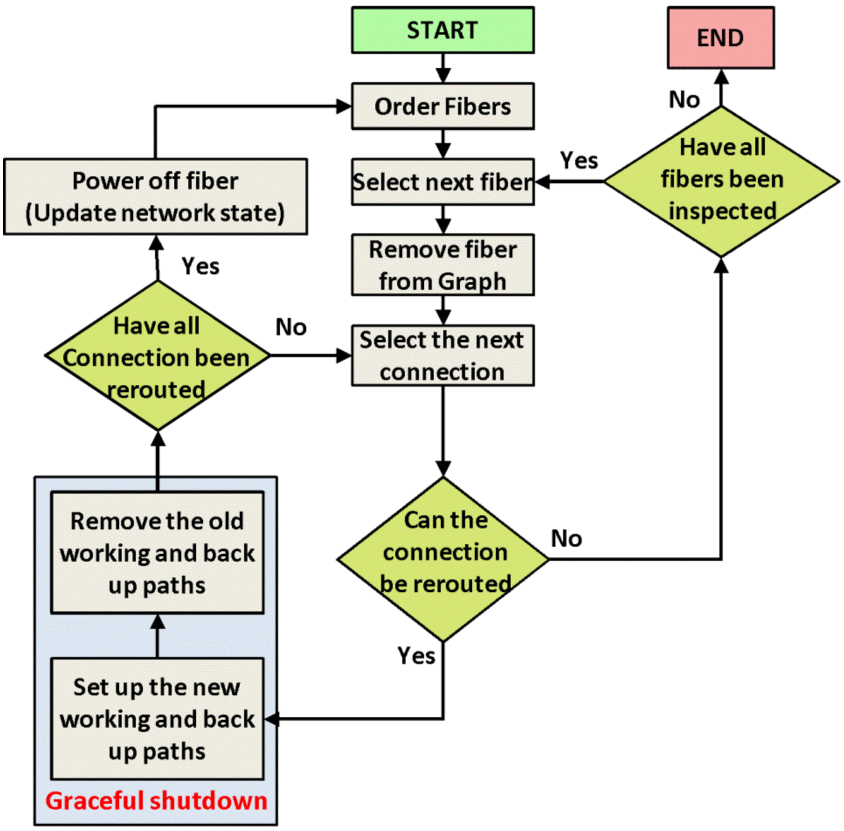


Figure IV-18: Switch-off algorithm flow-chart

The GMPLS graceful shutdown mechanism works as follows. The node initiating the shutdown procedure communicates to other nodes the unavailability of the resource under shutdown; then, for each affected connection, a reroute request is sent to the relative source node that will looks for a new path by avoiding the resource under shutdown. Flow interruption is avoided by means of the GMPLS make-before-break procedure [44]: a new set-up request is sent along the new computed path before to tear-down the connection from the old one. On common links between the old and the new path, the new set-up request is recognized as a reroute request of the old connection allowing the new connection to share resources with the old one; if the new connection can be established, the old one is torn down, otherwise an RSVP-PATH message is sent along the old path indicating that the reroute cannot be performed.

Ordering criteria here proposed are described in the following.

- **Random:** the Random criterion simply orders fibres in a random way; it is provided as comparison for other ordering criteria.

- **Least Flow (LF):** the probability that a fibre can be switched off depends on the number of wavelengths used on that fibre; for such a reason the LF criterion orders fibres in increasing values of their supported load.

- **Most Power (MP):** the amount of power that can be saved when a fibre is switched off depends on the number of EDFAs deployed along it; specifically, it is given by eq. II.7. Then the MP criterion orders fibres in decreasing values of their $P_{i,j}^{Fibre}$. It can be observed that different fibres of the same link have the same $P_{i,j}^{Fibre}$ factor; then the LF criterion is used to break ties in MP.

- **Maximum Power Saving Factor (MaxPSF):** the switch-off of a fibre imposes that an alternative path between the same node pair has to remain active in order to keep the network connected; then, for each fibre *k* of the link *s,d* the Least Cost Alternative Path (LCAP) of that fibre can be defined as the path connecting nodes s and d, that does not use the fibre k of the link s,d, whose sum of $P_{i,j}^{Fibre}$ terms is minimum. The cost of such a path is called Alternative Path Cost (APC). It is defined as follows: let $p_{s,d,k}^{n}$ be the n-th alternative path of the fibre k of the link s,d, $P_{s,d,k}=\{p_{s,d,k}^{n}\}$ be the set of all possible $p_{s,d,k}^{n}$ paths, $(i,j)_{s,d,k}^{n}$ a link belonging to $p_{s,d,k}^{n}$ and $L_{s,d,k}^{n}=\{(i,j)_{s,d,k}^{n}\}$ the set of all links belonging to $p_{s,d,k}^{n}$. Then $APC_{s,d,k}$ is given by (IV.23).

$$APC_{s,d,k} = \min_{n}\{\sum_{(i,j)\epsilon L_{s,d,k}^{n}}(P_{i,j}^{Fibre})\} \quad \text{(IV.23)}$$

The Power Saving Factor (PSF) is defined as the difference between the power absorbed by a fibre and its APC:

$$PSF_{s,d,k}= P_{s,d}^{Fibre}-APC_{s,d,k} \quad \text{(IV.24)}$$

The MaxPSF criterion order fibres in decreasing values of their PSF. It can be noted that if more than one fibre is still active on the link s,d then the LCAP of any active fibre of that link is simply given by another fibre of the same link. Then the MA and LF criteria are respectively used as first and second tie breaker parameters for the MaxPSF ordering criterion.

Performance Analysis

In order to evaluate the performances of the proposed solution, randomly generated network topologies and traffic demands have been considered.

Physical network topologies used in simulations are randomly generated by fixing the number of nodes N, the number of wavelengths of each fibre *W* and the connectivity degree *D*, defined as $D=2L/N(N-1)$, where *L* is the number of links in the network. It can be observed that the nodal degree *K*, i.e. the average number of links for each node, can be computed as $K=D(N-1)$.

Network nodes are uniformly distributed in a given area. Links between nodes are generated as follows. First, for each node, at most two links are generated by choosing among the shortest ones until the network graph is connected. Then other links are added so that two node-disjoint paths are available between each node pair. Finally, other links are added until the connectivity degree reaches a value equal to the pre-fixed *D* value; such links are chosen so as to minimize the ratio between the physical distance between two nodes and the physical length of the shortest path connecting such nodes. An example of a random generated physical topology is shown in Figure IV-19.

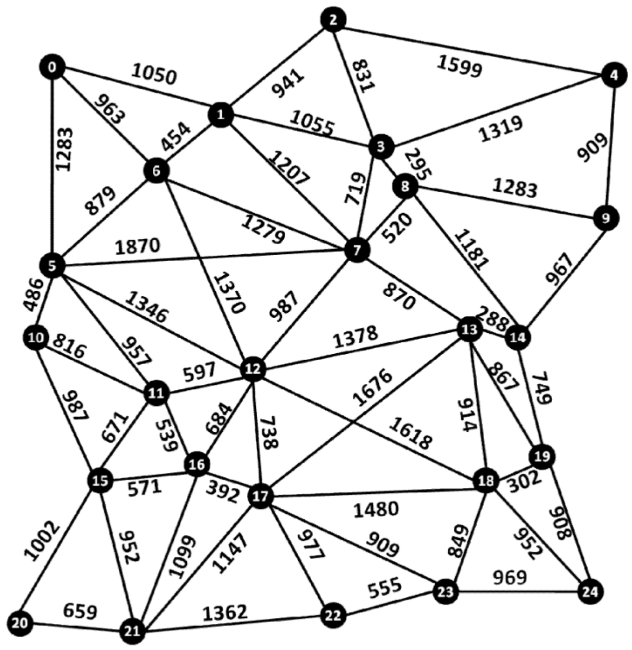


Figure IV-19: An example of a randomly generated physical network topology; *N*=25; *D*=0.3

The traffic demand considered here is represented by an NxN matrix $[F_{s,d}]$, where N is the number of nodes in the network. Each element $F_{s,d}$ represents the number of requested connections between nodes s and d and is assumed variable in time according to a sinusoidal function:

$$F_{s,d}(\mathrm{t})=F_{s,d}^{Min}[(\gamma-1)/2(1+\sin(2\pi f_0 t))+1]; \quad \mathrm{f}_0=1/24\mathrm{h},\ \gamma \text{ integer.} \qquad \text{(IV.25)}$$

The parameter γ represents the ratio between the peak and the minimum traffic loads during a day. It is assumed that during night the traffic load is 20% of the peak traffic, so that $\gamma=5$.

$F_{s,d}(t)$ is evaluated only for those values of t for which it results an integer: i.e., values of the variable t, such that:

$$F_{s,d}(\mathrm{t})=m\cdot F_{s,d}^{Min}\ m=1,\ldots,\gamma. \qquad \text{(IV.26)}$$

It results:

$$F_{s,d}^{Min} \leq F_{s,d}(\mathrm{t}) \leq \gamma\cdot F_{s,d}^{Min} \qquad \text{(IV.27)}$$

$F_{s,d}^{Min}$ represents the minimum number of connections requested between nodes s and d (i.e. the number of connections requested during the night), and it is randomly chosen between 0 and $Max\_F_{s,d}^{Min}$ according to the following probabilities:

$$\mathrm{Prob}\{F_{s,d}^{Min} > 0\}=1/2. \qquad \text{(IV.28)}$$

$$\mathrm{Prob}\{\ F_{s,d}^{Min} = n|\ F_{s,d}^{Min} > 0\}=1/Max\_F_{s,d}^{Min};\ \ n=1,\ldots,\ Max\_F_{s,d}^{Min}. \qquad \text{(IV.29)}$$

First he peak traffic matrix is considered, i.e. $F_{s,d}= \gamma\cdot F_{s,d}^{Min}$ and the network capacity (i.e. the number of fibres of each link) is dimensioned with respect to such a traffic matrix; then the power that can be saved by running the proposed mechanism is evaluated considering the traffic matrix scaled by a factor *m* as in (IV.26).

A network with *N*=20, *W*=40 and *D*=0.25, extending across an area of 4000 km x 3200 km is considered. Figure IV-20 shows the power consumption of all network devices versus the traffic load without applying the proposed algorithm.

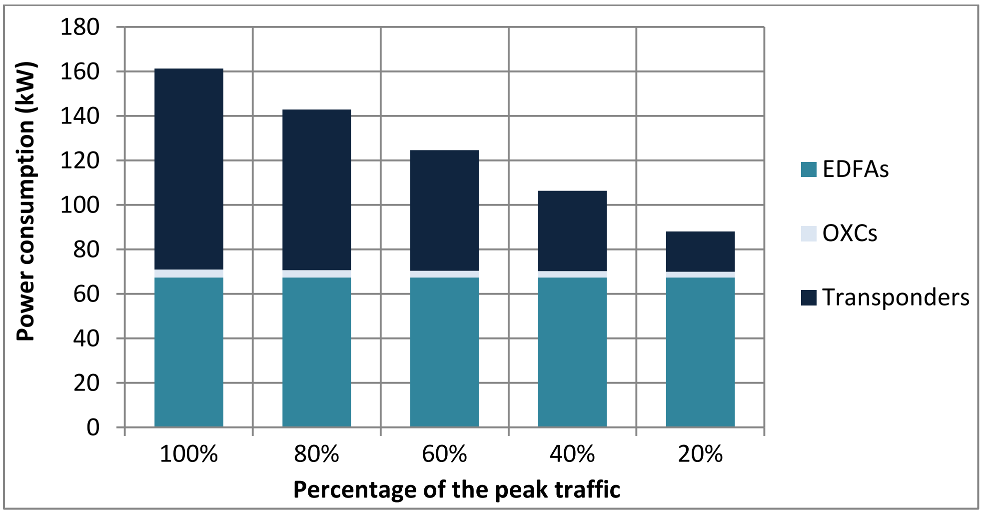


Figure IV-20: Network devices power consumption vs. traffic load with no optimization, *N*=20, *W*=40, *D*=0.25.

The power consumption is divided in three components corresponding to the three typologies of active devices in the network. Moreover it is assumed that, even if no optimization is performed, the number of transponders in power on state is equal to the number of flows in the network. It can be seen that, when traffic is at the peak level, the total power consumption of the network is about 160 KW. The most part of the energy is consumed by transponders and optical amplifiers that consume 56% and 42% of the total respectively, whilst OXCs consume just 2% of energy. However, when traffic decreases, the percentage of energy consumed by optical amplifiers becomes predominant reaching a value of 76% when traffic is at the 20%.

Instead, by running the proposed energy saving mechanism, the power absorbed by optical amplifiers can be significantly reduced. Figure IV-21 shows such a power saving considering optical amplifiers only; it can be seen that MaxPSF outperforms all other criteria, saving from 22% of power, when traffic is at its maximum, up to 80% for traffic at its minimum. The largest difference among heuristic performance is about 8% between MaxPSF and Random.

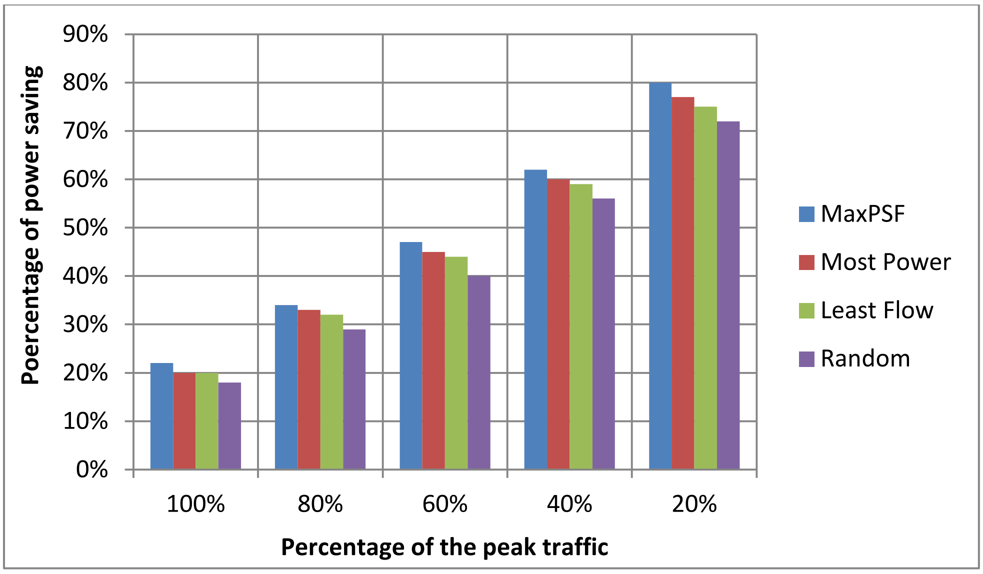


Figure IV-21: Link ordering criteria comparison. *N*=20, *W*=40, *D*=0.25.

Another interesting aspect concerns the impact that the proposed mechanism has on the physical length of lightpaths. For this purpose we evaluated the path length distribution (PLD) for both working and protection paths. Figure IV-22 shows the PLD when the proposed algorithm is not applied, when it is applied and the traffic load is at its peak (100%, indicates as First Opt.) and when it is applied with the minimum traffic load (20%, indicated as Last Opt.).

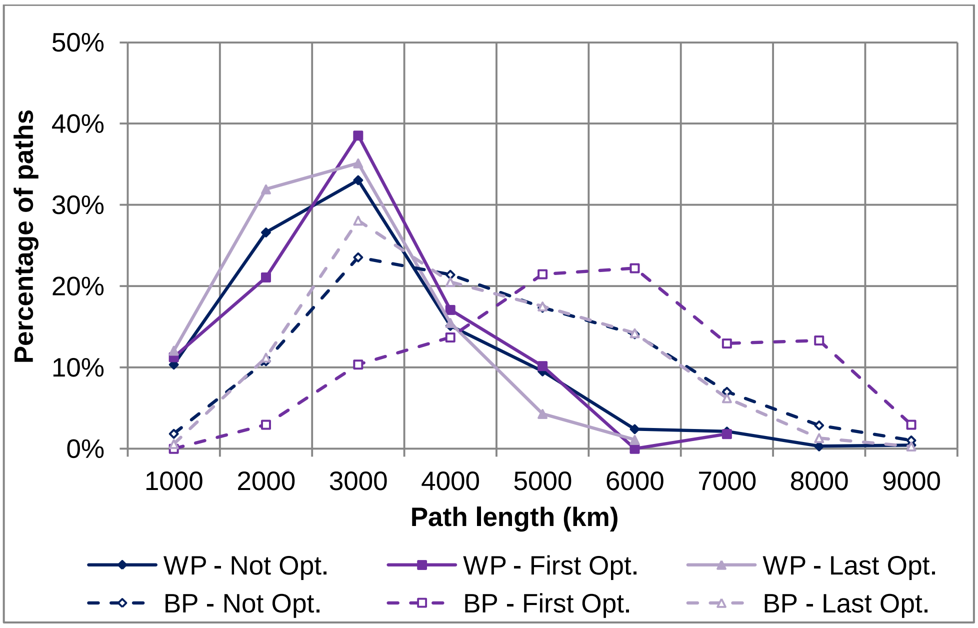


Figure IV-22: Comparison of the path length distribution without running the energy saving algorithm and performing link switch off with traffic at 100% and 20%.

It can be seen that the proposed mechanism allows to shorten the tail of the PLD for working paths, for both peak and minimum traffic load; instead, for protection paths, the average path length experiences a little increasing when the traffic is at its peak. Results have been presented in [55]-[57].

IV.2.2 *Power Aware RWA Algorithm*

The RWA problem has been well studied in the literature and several ILP formulations and heuristic algorithms have been proposed to solve it [47], [49]-[53]. Two different scenarios can be defined: the Static Lightpath Establishment (SLE) and the Dynamic Lightpath Establishment (DLE) problems [47]. In case of SLE, the traffic matrix is known in advance; classical RWA algorithms aim to minimize the number of wavelengths W that are needed to support a given traffic matrix or, equivalently, to maximize the number of supported connections for a given *W*. On the contrary, in the DLE scenario the traffic demand is not known in advance and connection requests dynamically arrive. The objective is to establish connections so that the total amount of blocked connections is minimized. Both the objectives are not consistent with the problem of minimization of the energy consumption. In fact, some algorithms, such as LCP, equally distribute the traffic over the network links, whilst others, such as ShP, aim to minimize the number of used wavelength channels by choosing the path with the minimum number of hops. Instead, in order to minimize the power consumption, the total length of used fibres should be minimized.

In III.4.1 the MUP algorithm and a modified version of it have been presented; here, instead a new PA-RWA algorithm is proposed.

Load Based Cost Routing Algorithm

The novel routing algorithm here presented is named Load Based Cost (LBC) and takes into account both the power consumption and the load status of each link. It is a least cost path algorithm but uses a time variant cost function to dynamically compute link costs accordingly to their current sate.

The cost $LC_{i,j}(t)$ associated with the link *i,j* at the time *t* is computed as follows:

$$LC_{i,j}(t)=PC_{i,j}\cdot LF_{i,j}(t). \qquad \text{(IV.30)}$$

Wherein: i) $PC_{i,j}$ is the Power Cost of the link *i,j*; it takes into account the power dissipation due to the optical amplifiers deployed along each fibre span of the optical link; this term is given by (2) and does not depend on the traffic load; ii) $LF_{i,j}(t)$ is named Load Factor of the link *i,j* at the time *t* and is computed on the basis of the link state in terms of number of used wavelengths. The term $LF_{i,j}(t)$ is determined through a cost function constituted by an inner function, named Fibre Cost Assignment (FCA) function, and an outer function, named Fibre Selection (FS) function; the FCA function computes the term $LF_{i,j}^{k}(t)$ representing the Load Factor related to the fibre k, whilst the FS function provides the term $LF_{i,j}(t)$ related to the whole optical link on the basis of the Load Factor of each fibre. More in detail:

$$LF_{i,j}(t)=\mathrm{FS}(LF_{i,j}^{1}(t), LF_{i,j}^{2}(t),...,LF_{i,j}^{k}(t),...,LF_{i,j}^{K}(t)). \quad \text{(IV.31)}$$

$$LF_{i,j}^{k}(t)=FCA(L_{i,j}^{k}(t)). \quad \text{(IV.32)}$$

In (IV.32) $L_{i,j}^{k}(t)$ is the load, i.e. the number of used wavelengths, of the fibre k of the link i,j at time t.

With respect to FCA functions, we observed that, by applying a null cost to used fibres, the average path length increases very much; this leads to bandwidth waste and, consequently, to an increase of the total number of needed fibres. Moreover, when fibre loads are high, the probability to find an available wavelength along the whole path rapidly decreases due to the wavelength continuity constraint; this could determine the need to use a new fibre, on some links of the chosen path, during the wavelength assignment phase. The previous observations suggest to use a cost function that increases with the load of each fibre. On the other hand, when the fibre load is low, the probability that the fibre could be powered off is high and thus a higher cost should be assigned to such a fibre in order to avoid its utilization.

The previous observations have suggested the definition of a class of functions, named V-Like functions, composed of two branches: the first one decreases in the interval $[0, W/2]$, where W is the number of wavelengths of each fibre; the second one increases in the interval $[W/2, W]$. The general form of the V-Like functions is given by eq. IV.33; the γ parameter varies the shape of the two branches. We tested values of γ between 1 and 4; the curve that showed the best efficiency was for γ=2. V-Like functions are shown in Figure IV-23, where W=20 is assumed. Note that the LRC function defined in IV.1.1 have been derived from V-Like functions.

$$LF_{i,j}^{k}(t) = |2\cdot L_{i,j}^{k}(t)/W-1|\hat{}(-\gamma\cdot \mathrm{sign}(2\, L_{i,j}^{k}(t)/W-1)). \qquad \text{(IV.33)}$$

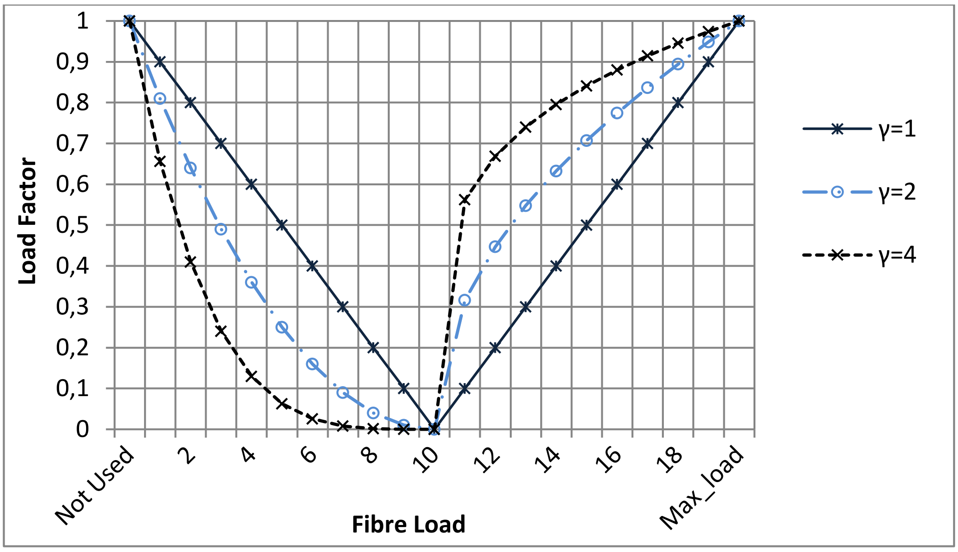


Figure IV-23: V-Like functions for γ=1, γ=2 and γ=4.

The shape of the curve has been suggested by the following considerations. With respect to the decreasing branch, it can be observed that the probability to power off a fibre exponentially decreases with the number n of wavelengths used on that fibre; in fact it is the probability that *n* connections end before that a new connection is established[1]. On the increasing branch, instead, the function rapidly grows in order to avoid long paths with several links supporting a medium-high load.

The FS function is used to obtain a cost that is representative of the whole link, starting from the cost of each fibre. Several options are possible such as considering the highest cost or the lowest one, or combining all fibre costs in some way. We tested the Min(·), Max(·) and Avg(·) functions; however, only results related to Avg(·) are here presented since it showed the best performance.

Algorithm 1 – LBC link cost computation

```
for all links in the network
    unused_fibre = false;
    for all fibres of the current link i,j
        if  0 < L_{i,j}^k < W  then
            LF_{i,j}^k = FCA(L_{i,j}^k);
            add the term LF_{i,j}^k to the fibre_cost_list;
        else if  L_{i,j}^k = 0  then
            unused_fibre = true;
        end if
    end for
    if fibre_cost_list is not empty then
        LF_{i,j}=FS(fiber_cost_list);
    else if unused = true then
        LF_{i,j} = 1;
    else
        LF_{i,j} = ∞;
    end if
    LC_{i,j} = PC_{i,j}·LF_{i,j};
    set LC_{i,j} to link i,j;
end for
```

[1] The statement "exponentially decreases" holds only if the probability that a lightpath is accommodated on a given fibre does not depend on the number of connections that are already established on that fibre.

The FS function does not consider $LF_{i,j}^{k}(t)$ terms equal to 1, i.e. terms related to unused or fully loaded fibres; if all terms are equal to 1, two cases can be distinguished: 1) an unused fibre exists; 2) there are no unused fibres. In the first case the term $LF_{i,j}(t)$ is set to 1; in the second case it is set to infinity. Algorithm 1 reports the pseudo-code of the link cost computation method.

Wavelength Assignment Algorithms

The wavelength assignment problem has been extensively studied in the literature. Several heuristics, like First Fit (FF), Most Used (MU), Least Used (LU), Least Loaded (LL), Relative Capacity Loss (RCL), etc., have been proposed and their performance evaluated [47], [52], [53]. In particular, FF, besides showing a good performance in terms of blocking probability, has a low computational complexity and can be easily implemented since requires that each node handles only local information. This makes it the most used wavelength assignment algorithm in both real networks and research papers.

FF simply assigns a sequence number to wavelengths, say from 0 to *W*, and then chooses the smallest one that results available on the whole path. When applied to a multi-fibre WDM network, FF looks for an available wavelength considering both used and unused fibres. In this way all fibres of a link are likely to be used prior to choose a higher wavelength. In the present paper FF has been used with LCP in order to provide an upper bound to the total consumed energy; in fact, since LCP tends to equally load network links and FF to use all fibres of a link, it is very likely that all fibres in the network are used even in case of low traffic load. Moreover, since FF does not care of the consumed energy, it does not make sense using it with power aware routing algorithms like MUP, LBC or even the power aware version of the ShP algorithm here considered.

FF Derived Wavelength Assignment Algorithms

In order to avoid energy waste, a modified version of FF was proposed in [37], called Two Phase-FF (TP-FF) here. It works as follows: in the first phase it works as FF but considers only used fibres; if a block occurs, i.e. if an available wavelength cannot be found considering just used fibres, it considers unused fibres too.

We propose a different version of the FF algorithm, named Least Additional Power-FF (LAP-FF). It differs from TP-FF since, when a block occurs in the first phase, the wavelength requiring the minimum amount of additional power is selected, that is the wavelength that requires to use a set of new fibres whose sum of $PC_{i,j}$ terms is minimum.

Finally, when the chosen wavelength is available on more than one fibre of the same link, all previously defined algorithms select the first fibre.

Least Cost Wavelength Algorithm

Traditional Wavelength Assignment (WA) heuristic algorithms aim to use the minimum number of wavelengths to support a given traffic matrix. In order to achieve such an objective, an efficient overall criterion is to use, as most as possible, wavelengths that are already used in the network. Two among the most efficient heuristics, FF and MU, follow this general principle [47].

The wavelength assignment algorithm here proposed, instead, aims to keep active the minimum number of optical amplifiers. In doing this, several issues must be considered. First, when a connection request arrives, the wavelength requiring the smallest amount of additional power should be used; this is the overall criterion followed by LAP-FF. However, in such a way, the load of other involved fibres would not be taken into account. In fact, since least loaded fibres are likely to be powered off, wavelengths on such fibres should not be used if other wavelengths are available.

Starting from the considerations above, we developed a novel power aware WA algorithm, named Least Cost Wavelength (LCW). It works as follows.

For each link i,j belonging to the found path, the algorithm computes the cost $WC_{i,j}^{\lambda}(t)$ of each wavelength λ on the basis of the state of the link i,j; then the algorithm looks for the wavelength that minimizes the sum of $WC_{i,j}^{\lambda}(t)$ costs on the whole path. The basic idea is that each fibre has its own cost, whilst wavelengths assume the cost of the fibre which they result available on.

The cost $FC_{i,j}^{k}(t)$ assigned to the fibre k of the link i,j at the time t is composed of two terms:

$$FC_{i,j}^{k}(t) = PC_{i,j} \cdot LF_{i,j}^{k}(t). \tag{IV.34}$$

Wherein: i) $PC_{i,j}$ is the Power Cost of the link i,j, as defined in (2); ii) $LF_{i,j}^{k}(t)$ is named Load Factor of the fibre k of the link i,j at time t and it is computed by means of a Fibre Cost Assignment (FCA) function on the basis of the fibre state in terms of number of used wavelengths.

$$LF_{i,j}^{k}(t) = FCA(L_{i,j}^{k}(t)). \tag{IV.35}$$

$L_{i,j}^{k}(t)$ is the number of wavelengths used on the fibre k of the link i,j.

As mentioned above, the cost of each wavelength is given by the cost of the fibre which it is available on. However, when a given wavelength λ is available on more than one fibre of the same link, a Fibre Selection (FS) function is used as follows:

$$WC_{i,j}^{\lambda} = PC_{i,j} \cdot FS(LF_{i,j}^{1}(t),..., LF_{i,j}^{k}(t),..., LF_{i,j}^{K}(t)). \tag{IV.36}$$

$LF_{i,j}^{k}(t)$ terms inside the *FS* function are related to fibres that are already used, i.e. terms $LF_{i,j}^{k}(t) \neq 1$, and which the wavelength λ is not used on. If the wavelength λ is used on all used fibres two cases can be distinguished: 1) an unused fibre exists on that link; 2) all fibres are already used. In the first case, the FS function returns a value equal to 1 and the term $WC_{i,j}^{\lambda}$ assumes the value $PC_{i,j}$; in the second case an infinity cost is assigned to the wavelength λ since it is not available on that link.

With respect to the cost of each fibre, it has to be noted that, when the wavelength assignment algorithm is run, links composing the path have already been decided. For such a reason, the objective of the WA algorithm is to choose the wavelength avoiding using slightly loaded fibres. In this way, in fact, the probability to power off a fibre can be increased. For such a reason any plausible FCA function should be a monotone decreasing function in the interval [0,*W*]. We tested functions with linear, exponential and hyperbolic decrease; they are shown in Figure IV-24 for *W*=10. The one that showed the best performance was the Hyperbolic(·) function that computes costs as follows:

$$LF_{i,j}^{k}(t) = 1/(L_{i,j}^{k}(t) + 1). \qquad \text{(IV.37)}$$

The Hyperbolic(·) function addresses the issue that the probability to switch off a fibre decreases more than linearly with the load. In the following, only results related to the Hyperbolic(·) function are shown.

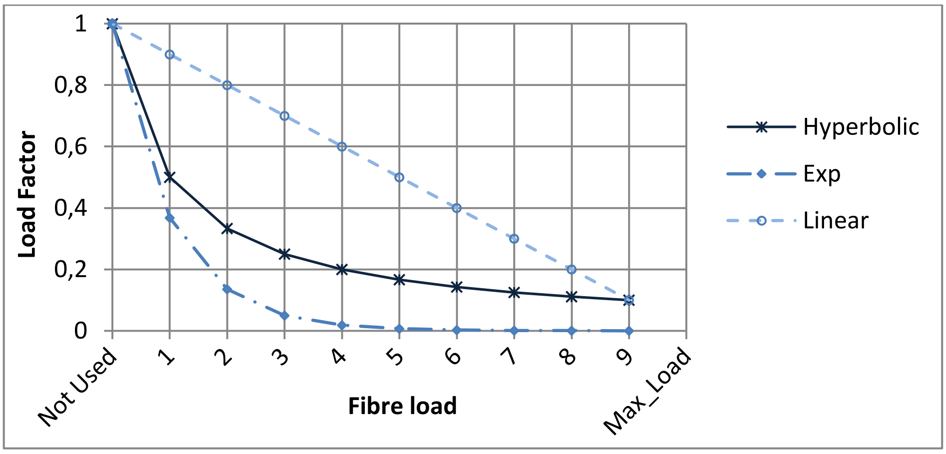


Figure IV-24: FCA functions used for fibre cost calculation in the LCW algorithm.

As FS functions we propose the Parallel(·) function, defined as follows:

$$\text{Parallel}(x_1, x_2,..., x_N)=1/(1/ x_1+1/ x_2+,...,+1/ x_N). \quad \text{(IV.38)}$$

The Parallel(·) function has the objective to assign a lower cost to wavelengths that result available on more than one fibre of the same link. In fact, by selecting a similar wavelength, it will be still possible to accommodate a future request on the same wavelength of the same link.

Once the wavelength has been decided and it is available on more than one fibre of the same link, the algorithm always selects the most loaded fibre among suitable ones. Finally, FF is used to break ties.

Algorithm 2 shows the pseudo-code of the wavelength cost computation method of the LCW algorithm. It has to be observed that, if only one fibre pair is deployed for each bidirectional link, all discussed algorithms exactly work as FF.

Performance evaluation

The performance analysis has the objective to evaluate the feasibility of the proposed RWA algorithm. In particular, two different performance metrics are considered here:

- the improvement, in terms of power consumption, of adopting the proposed RWA algorithm with respect to other RWA algorithms;
- the price paid in terms of blocking probability with respect to the LCP|FF alternative.

Physical network topologies used in simulations are randomly generated as described in IV.2.1.

Results presented in the following have been obtained by combining all routing algorithms with all WA algorithms in order to better understand the impact that routing and WA algorithms have on the performance of the whole RWA mechanism. Specifically, the considered routing algorithms are the following:

1) Least Congested Path (LCP)
2) Power-Aware Shortest Path (ShP)
3) Most Used Path (MUP)
4) Load Based Cost (LBC)

The considered WA algorithms are the following:

1) First Fit (FF)
2) Two Phases-First Fit (TP-FF)

3) Least Additional-Power First Fit (LAP-FF)

4) Least Cost Wavelength (LCW)

Moreover, the optimal solution given by the formulation provided in III.2 has also been evaluated.

Algorithm 2 – LCW wavelength cost computation

for all links in the path
    for all wavelengths λ
        for all fibres of the current link i,j
            unused_fibre = false;
            if current λ is available **then**
                $LF_{i,j}^{k}$ = FCA($L_{i,j}^{k}$);
                add the term $LF_{i,j}^{k}$ to the fibre_cost_list;
            else if $L_{i,j}^{k} = 0$ **then**
                unsed_fibre = true;
            end if
        end for
        if fibre_cost_list is not empty **then**
            $WC_{i,j}^{\lambda} = PC_{i,j}$·FS(fiber_cost_list);
        else if unused_fibre = true **then**
            $WC_{i,j}^{\lambda} = PC_{i,j}$;
        else
            $WC_{i,j}^{\lambda} = \infty$;
        end if
        set $WC_{i,j}^{\lambda}$ to the current λ on the current link;
    end for
end for

Power consumption

In Figure IV-25 all routing algorithms are compared for each WA algorithm. The percentage power consumption with respect to the LCP|FF alternative can be observed on the right axis. As it can be expected, LCP consumes much more power than others, since it always tries to equally distribute the traffic on all links. LBC shows the best performance with any WA algorithm, whilst MUP performs worse than ShP. LBC saves about 45% of energy with respect to LCP and between 24% and 32% with respect to MUP, depending on the used WA algorithm. Instead, when compared with ShP, just a decrease between 5% and 10% of the consumed energy is achieved. ShP, in fact, shows a good performance in terms of energy consumption even if it was not designed to achieve such an objective. This is due to the fact that, by applying a cost equal to $PC_{i,j}$, ShP aims to minimize the number of optical amplifiers used in the network; moreover, when multiple fibres are needed for each link, the wasted energy regards only one used fibre of each link. In other words, said $F_{i,j}$ the number of fibres used by ShP on the link i,j, $F_{i,j}$-1 fibres are strictly needed whilst the $F_{i,j}$-th could be powered off and the traffic on it routed on some other links.

On the contrary, MUP does not show an excellent performance even if it was designed to minimize the consumed energy. MUP suffers of too long paths that cause capacity waste leading to a higher power consumption. The main drawback of MUP is that, when at least one fibre on a number of links connecting all network nodes have been used, other links are very unlikely to be used. This occurs because, when a link with a high load is selected to accommodate a lightpath, it is very likely that a block occurs during the wavelength assignment phase, determining the need to use a new fibre on some links of the chosen path; in such a way, the cost of a link is unlikely to return high. In other words MUP does not takes into account that, even if a wavelength is available on an used fibre of a link, the wavelength continuity constraint could determine the need to use new fibres, especially when the path is long and some links are highly loaded. LBC, instead, takes into consideration the link load by means of the FCA function and prefers to use a new fibre rather than to route a lightpath on a path much longer than the shortest one.

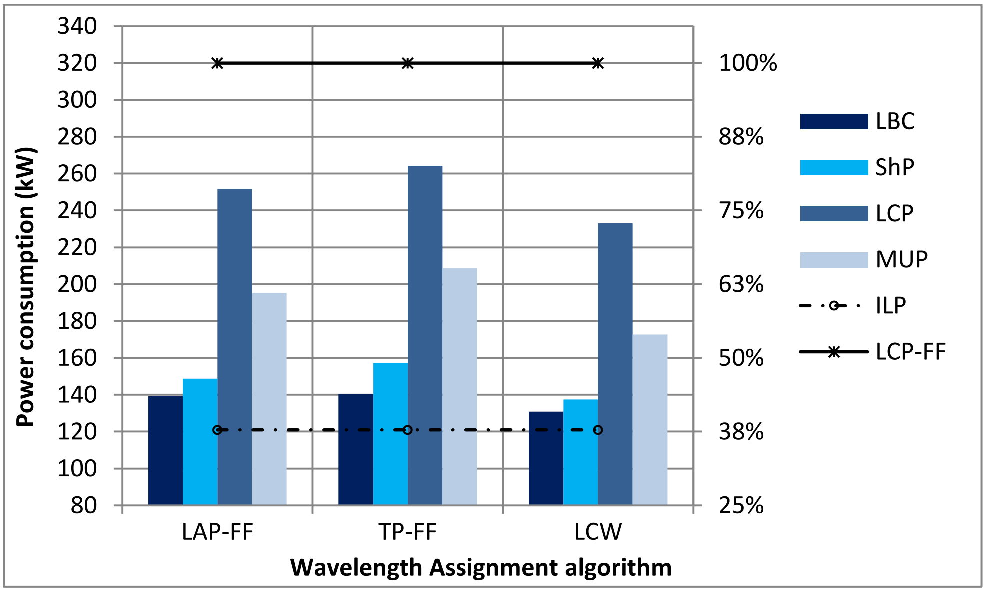


Figure IV-25: Routing algorithm performance comparison for TF-FF, LAP-FF and LCW wavelength assignment algorithms.

By observing the distribution of the average number of used fibres for each link, shown in Figure IV-26, it can be seen that MUP keeps completely unused a higher number of links with respect to LBC but, at the same time, it uses much more fibres on some links and never uses only one fibre on a link. For ShP, such a distribution is very similar to that of LBC, whilst for LCP it is very tight around its medium value and links are never unused.

Figure IV-25 also shows results related to LCP|FF, representing the upper bound, and to the MILP formulation, representing the lower bound. LBC|LCW allows to save about 60% of energy with respect to the LCP|FF alternative and its performance is very close to the lower bound, with a percentage difference of about 10%.

By comparing the WA algorithms it can be seen that LCW shows the best performance for each routing algorithm and LAP-FF always outperform TP-FF. When LCP or ShP is used to solve the routing problem, the percentage save of LCW with respect to TP-FF and LAP-FF is around

12% and 7% respectively; when using MUP, instead, it increases to 17% and 12%. Finally, when LBC is used, LAP-FF performs very close to TP-FF and the LCW improvement is just 6%.

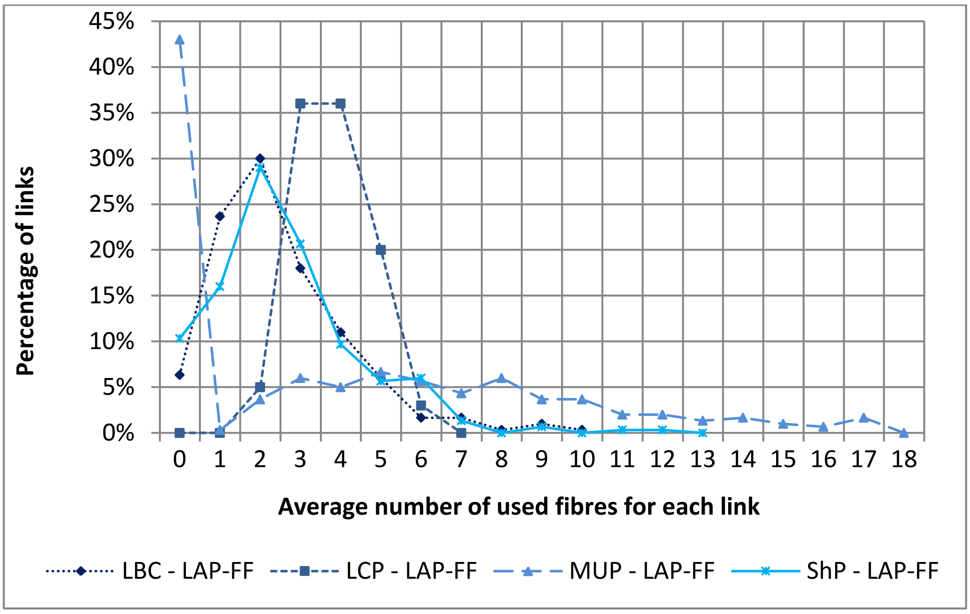


Figure IV-26: Distribution of the number of used fibres for each link, for all routing algorithms; LAP-FF used as WA algorithm.

It must be pointed out that results here presented are not in contrast with those presented in [37] since different values of traffic load are considered in the two works. In [37] the total amount of traffic in the network is equal to *N*(*N*-1)/4 bidirectional lightpaths, where *N* is the number of nodes, resulting in an average traffic load of 0.5 Erlang between each node pair whilst the number of wavelengths of each fibre is equal to 128. In the next subsection the impact of the traffic load is evaluated.

Impact of traffic load

In order to evaluate the impact that the amount of traffic treated by the network has on RWA algorithms, a set of simulations have been carried out by varying the traffic load from 0.3 up to 20 Erlang between each node pair.

The same network parameters of previously analyzed studies have been used. We define the energy efficiency η_e of a given algorithm as ratio between the power consumption given by the optimal solution, obtained by resolving the ILP formulation, and the power consumption obtained by using that algorithm. Results are presented by comparing the energy efficiency of the different algorithms in order to better outline performance differences and the distance from the lower bound. The absolute values of power consumption can be derived from Table IV-1, reporting the power consumption given by the optimal solution, and the energy efficiency shown in Figure IV-27.

Average traffic load between each node pair (Erlang)	0.3125	0.625	1.25	2.5	5	10	20
Power Consumption (KW)	10.3	14.6	22.1	34.1	55.9	98.8	176.9

Table IV-1: Values of power consumption of the optimal solution.

By observing results presented in Figure IV-27, where routing algorithms are compared with each other for all WA algorithms, it can be seen that MUP performs better than ShP when the traffic load is low. In particular, when TP-FF is used and the traffic load is equal to 0.3 Erlang, MUP reduces the power consumption by a factor of 5 with respect to ShP, that is approximately the same result obtained in [37]. However, as the traffic load grows, MUP power consumption rapidly increases and exceeds ShP when the traffic load reaches 2.5 Erlang. LBC, instead, allows to reduce the power consumption for both low and high traffic loads; it shows an energy efficiency between 85% and 95% when used with LCW or LAP-FF leading to a power saving between 25% and 35% with respect to the MUP|TF-FF alternative.

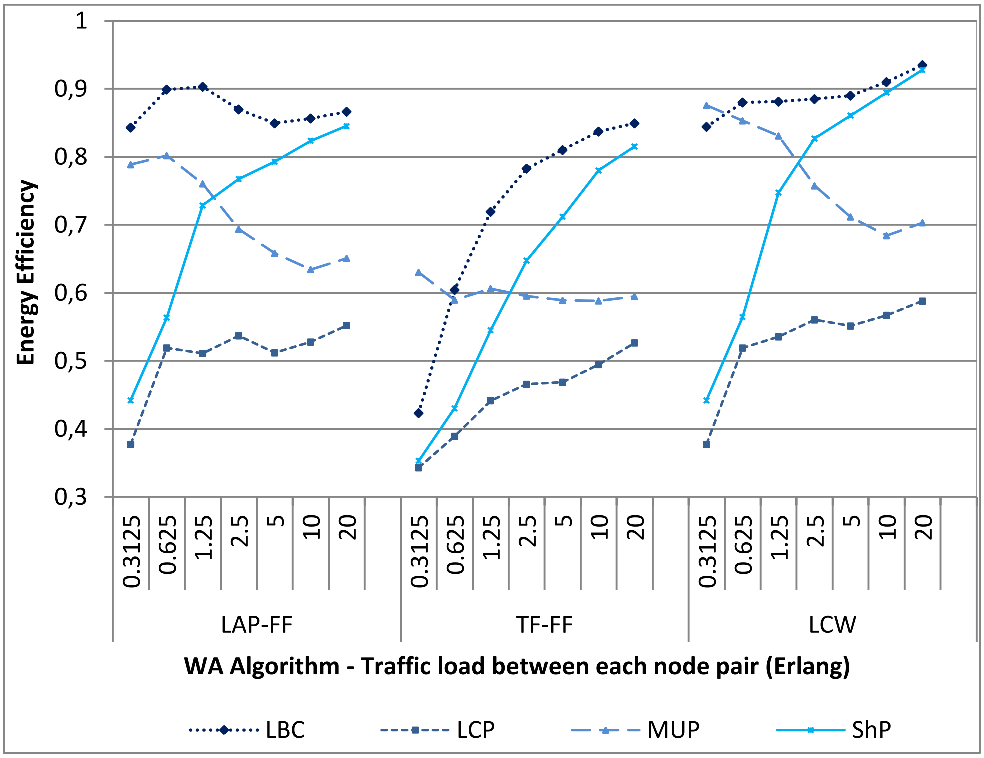


Figure IV-27: Routing algorithm performance comparison versus traffic load, for LAP-FF, TP-FF and LCW wavelength assignment algorithms.

When the load is high, ShP performs very close to LBC; this occurs because, since the number $N_{i,j}$ of used fibres grows, the power reduction opportunity decreases to $1/N_{i,j}$, as explained in the previous section. This can be also explained by observing the path length distribution shown in Figure IV-28. MUP behaves very similar to LCP, showing a long tail and an average path length equal to about 4.2 hops. On the contrary, ShP shows a short tail and an average path length equal to about 2.5 hops. Unlike the other algorithms, LBC adapts itself to the load state; for low load, it behaves like MUP and shows an average path length equal to about 3.7 hops; for high load, it practically shows the same distribution as ShP and the average path length is just 2.6 hops.

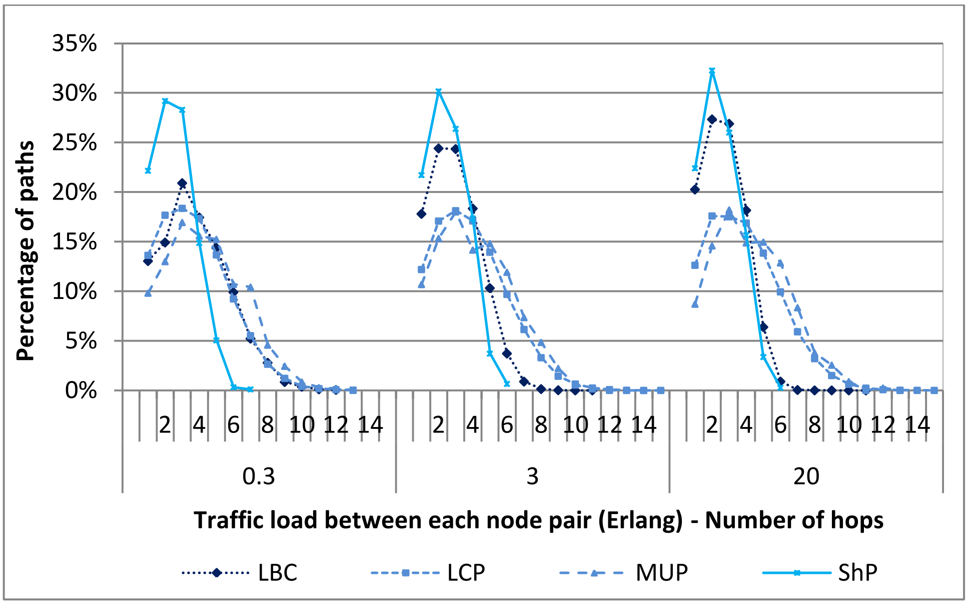


Figure IV-28: Comparison of path length distributions for all routing algorithms, for low (0.3 Erlang), medium (3 Erlang) and high (20 Erlang) traffic loads. LAP-FF is used as wavelength assignment algorithm.

In Figure IV-29, WA algorithms are compared with each other for all routing algorithms. LCW always shows the best performance; the power saving percentage with respect to LAP-FF is negligible at low loads whilst grows up to about 10% at high load. TP-FF, instead, consumes much more power than others at low load but shows almost the same performance of LAP-FF when the load is high. Moreover LBC suffers more than other routing algorithms the utilization of TP-FF at low loads.

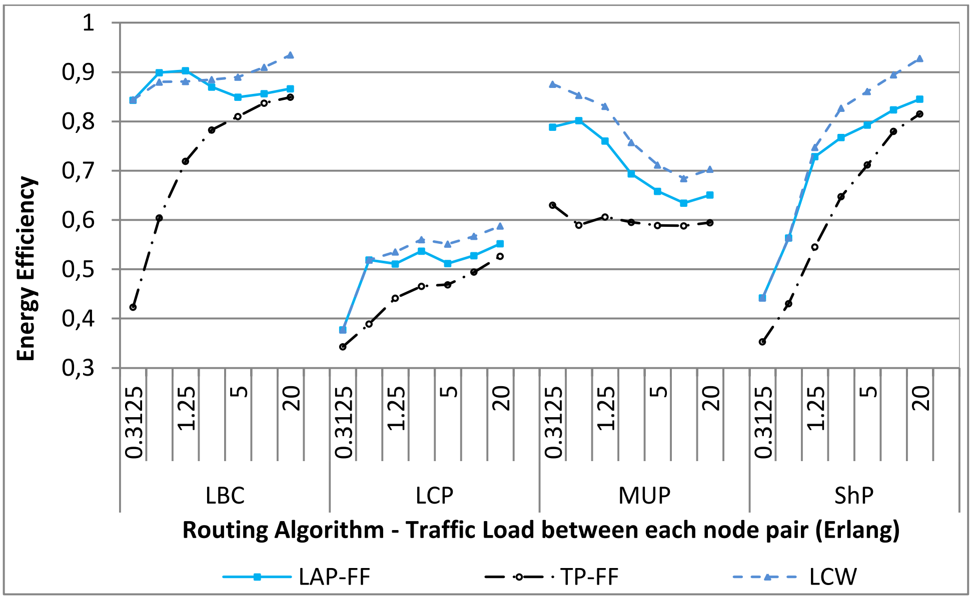


Figure IV-29: Wavelength assignment algorithm performance comparison versus traffic load, for LCP, ShP, LBC and MUP routing algorithms.

Blocking Probability

The performance of classical RWA algorithms is generally evaluated in terms of supported traffic fixing the amount of resources available in the network or, equivalently, in terms of resources needed to support a given traffic matrix. Network resources are represented by the number of wavelengths *W* of each optical fibre and, in case of multi-fibre links, the number of fibres of each WDM link. When the DLE problem is faced, since the traffic request is not known in advance, the blocking probability is taken into account as performance metric.

Even if the RWA algorithm proposed in this paper has not been designed to achieve such an objective, it is interesting to analyze its performance also in terms of blocking probability in order to provide an evaluation of the price that must be paid to achieve the energy efficiency shown in the previous section.

A set of experiments have been carried out fixing *N*=12, *W*=40 and *D*=0.3; the number of fibres has been fixed to 5 for all links and the blocking probability has been evaluated for different values of traffic load. We choose a lower value of *N*, with respect to the power consumption analysis, in order to speed up each run, given that, in this case, we perform a much higher number of runs in order to obtain a stable value for low blocking rates. All routing and all WA algorithms have been considered; results are shown in Figure IV-30.

As it can be expected, LCP shows the lowest blocking probability; however LCW outperforms FF that shows the same performance of TP-FF and LAP-FF. Even if just a little difference is shown between the two WA algorithms, such a result demonstrates that it is possible to use LCW with LCP without any impairment in terms of blocking probability with respect to FF. Moreover, since LCW, as FF, needs that nodes handle only local information, the only price that must be paid is due to the higher computational complexity that it requires. However, it can be observed that $WC_{i,j}^{\lambda}$ terms used by the algorithm can be updated by each node whenever a new lightpath is established; in such a way no processing delay will be experienced during the set-up procedure of the next connection.

The second important result is that LBC outperforms all other routing algorithms except than LCP. Furthermore, as the traffic load grows, it performs very close to LCP|FF. More in detail, for a traffic load of about 25 Erlang between each node pair, they practically experience the same blocking probability (about 5%); if a lower blocking probability is desired in the network, for example about 10^{-3}, the supported traffic varies from about 13 Erlang, for LBC|LCW, to about 15 Erlang for LCP|FF, i.e. a percentage difference of about 15%.

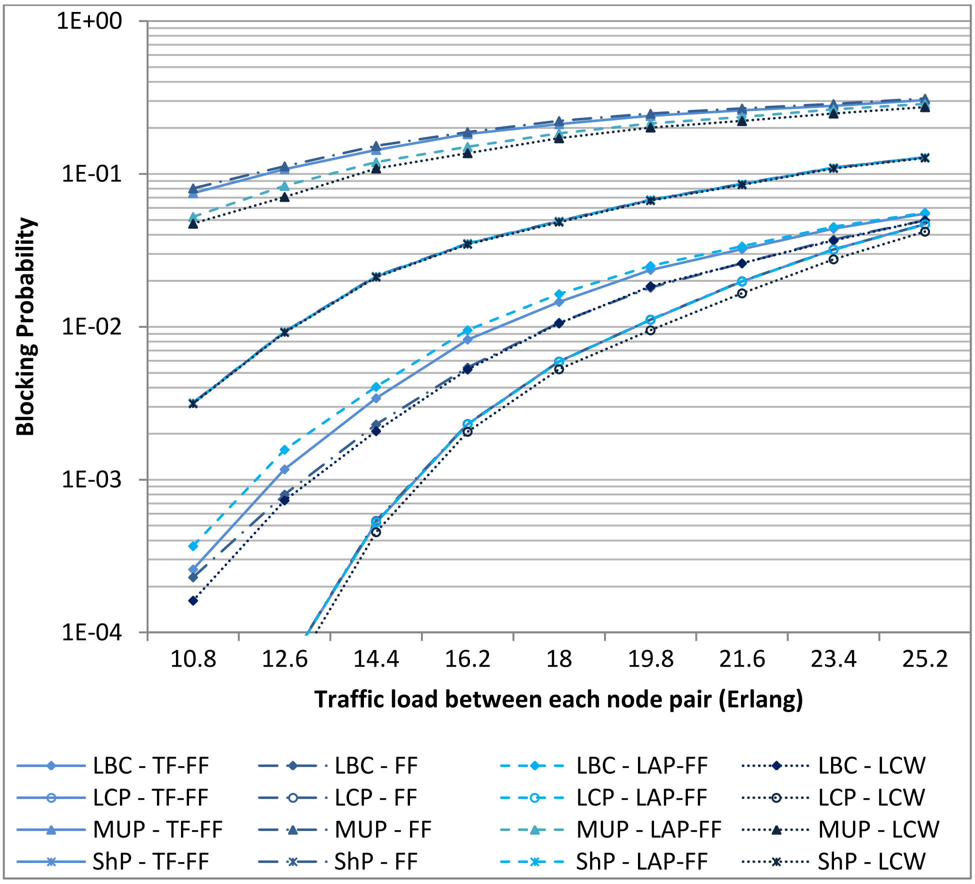


Figure IV-30: Blocking probability versus traffic load for all routing and wavelength assignment algorithms. *N*=12; *W*=40; *D*=0.3; fixed number of fibres *F*=5.

Moreover, it has to be noted that, even if LCW leads only to a little improvements in terms of power consumption with respect to LAP-FF when used with LBC, it leads to a more relevant improvement in terms of blocking probability.

Results also show that ShP and MUP experience a blocking probability much higher than LBC and LCP; ShP does not experience any difference with respect to the used WA algorithm, whilst, for MUP, LCW shows the lowest blocking probability, FF the highest and LAP-FF outperforms TP-FF. This can be explained by observing that, since LCW and LAP-FF aim to saturate already used fibres prior to use a new fibre, they increase the probability to take the cost of a link back to a high value, allowing MUP to use a larger number of links.

It is worth outlining that, even if MUP and ShP could be valid alternatives to LBC for networks with respectively a low and a high number of fibres per link, their poor performance in terms of blocking probability drastically reduces their applicability; thanks to its low blocking probability instead LBC is a viable solution to improve the energy efficiency of optical networks.

The results of the present study have been presented in [54], [59].

Chapter V
CONCLUSION

Energy saving in telecommunication networks is gaining increasing attention in the networking research community. This dissertation focused on that topic considering the backbone section of the network. We developed algorithms and mechanisms to build energy efficient next generation backbone networks that leverage on new concepts and methodologies to make current routing algorithms, signalling protocols and networking technologies green. In this Chapter the main contributions are summarized.

V.1 Summary of the Research Contribution

Chapter II discussed energy consumption trend and the need of "green" networking providing a comprehensive summary of the main research efforts on that subject. Moreover, we also developed an energy consumption model for IP over WDM backbone networks which was used to formalize the energy consumption minimization problem and to assess the performance of the proposed solutions.

Chapter III discussed the different strategies that can be undertaken for improving network energy efficiency emphasizing those based on energy-efficient network design and on traffic and topology control. The chapter also provided a mathematical formalization of the energy consumption minimization problem in IP over WDM networks and discussed possible ways for simplifying the problem which was treated separately for the two network layers. Specifically, concerning the IP layer, the specific strategies of EM-VTD and EA-TE were discussed separately and several solutions proposed in the literature were presented. Then, two possible approach aimed at achieving energy saving in the optical layer were considered, namely fiber-link switch-off and PA-RWA.

Chapter IV presented the proposed solutions and discussed the related performance. The main findings are summarized in the following.

V.1.1 *Energy Minimized Virtual Topology Design*

Concerning the EM-VTD problem we developed a heuristic algorithm named *Start-Single Hop and ReRoute* (*Start-SH&ReR*). The proposed algorithm follows an opposite approach with respect to other algorithms proposed in the literature; in fact, rather than considering an empty starting topology and adding links step-by-step, the *Start-SH&ReR* algorithms starts considering the solution provided by the *Single-Hop lightpath bypass* strategy and then tries to eliminate unnecessary links by rerouting traffic flows on the rest of the network. To achieve such an objective the algorithm uses a particular cost function to compute link weights into the shortest path routing algorithm named Additional Power and Link Load (AP&LL).

The performance of the proposed algorithm were evaluated and compared against the optimal solution, considered as a lower bound, the *Single-Hop lightpath bypass* strategy, considered as an upper bound, and another heuristic recently proposed in the literature [22] here referred to as *Start-Empty*.

Among the main achieved results are the following.

i) The energy efficiency (comparison with the optimal solution) of the *Start-SH&ReR* algorithm is between 75% and 100% depending on the traffic load.

ii) *Start-SH&ReR* outperforms *Start-Empty* allowing energy saving between 50% and 25% when every traffic relation between IP routers is lower than the line card capacity. For higher values of traffic load both algorithms and the *Single-Hop lightpath bypass* strategy perform close to the lower bound.

iii) *Start-SH&ReR* allows to better use the capacity deployed in the network; specifically, the percentage of links experiencing maximum load is less than in case of *Start-Empty* but, at the same time, link are never underutilized.

iv) *Start-SH&ReR* is able to limit the maximum path length to 4 hops whilst *Start-Empty* routes some flows on paths as long as 16 hops. This result is mainly due to the use of the AP&LL cost function.

V.1.2 *Energy Aware Traffic Engineering*

Concerning the EA-TE strategy, we developed a distributed solution named *Distributed and Adaptive Interface Switch-Off for Internet Energy Saving* (DAISIES). The mechanism follows a "flow-based" approach since it mainly acts on the routing algorithm which, by taking routing decisions, determines the traffic removal from some links and their consequent switch-off. The opposite approach, which can be referred to as "link-based", involves instead the definition of heuristic criteria to identify the set of links to kept active (or to switch off) and just subsequently the actual traffic rerouting.

DAISIES leverages on the MPLS technology and requires that an LSP is established between any node pair whose traffic relation is not null. The main advantages deriving from adopting the MPLS technology are the following.

i) Having complete knowledge about the load state of each link in the network. Such information can be advertised in the network by the TE routing protocol on the basis of the capacity reserved for each LSP.

ii) Avoiding both packet loss and out-of-order packet delivering during the rerouting of a traffic flow by using the MPLS make-before-break mechanism to reroute LSPs.

iii) Decoupling the logical topology seen by IP entities, which is equivalent to the set of LSPs, from the actual topology at the MPLS layer.

The basic idea of DAISIES is to reroute each LSP when the carried traffic changes, fixing an upper and a lower threshold. Such an operation is performed by the Ingress node exploiting the knowledge of the load state of every link and by using a specific cost function into the classical Dijkstra algorithm. The switch-off of a line card, as well as its reactivation, is then performed by each node in the network respectively when no capacity is reserved on it and when some additional capacity is needed to fulfill an LSP set-up request. As cost functions, we used the AP&LL function and a cost function based on a classical approach and here named Only Additional Power (OAP).

Among the main advantages of the proposed solution are the following.

i) Making the line card switch-off/activation process transparent to the routing protocol and to other network nodes;

ii) Avoiding packet loss and out-of-order packet delivering;

iii) Automatically managing both traffic decreasing and increasing;

iv) Requiring no change to current routing and signalling protocols.

The performance of the proposed solution was evaluated; the main results can be summarized as follows.

i) The energy efficiency of the proposed solution is between 87% and 100% depending on how meshed the *Base Network* is.

ii) When the *Base Network* is little meshed AP&LL performs much better than OAP.

iii) The complexity of the proposed solution can be kept low while still achieving a good level of energy efficiency.

V.1.3 *Energy Efficiency in Transparent Optical Networks*

Two different solutions were proposed for saving energy in transparent optical networks.

The first one is based on an iterative greedy algorithm that a-posteriori tries to aggregate lightpaths on a subset of optical fibres links, re-optimizing the network when the traffic decreases. The algorithm simply orders fibre links according to some heuristic criteria and tries to switch off each link by rerouting lightpaths on the rest of the network. The proposed mechanism exploits the GMPLS graceful shutdown mechanism to avoid traffic flow interruption during lightpath rerouting. Results showed that the application of the proposed algorithm allows to achieve fully proportionality between traffic load and power consumption.

The second proposed solution is a Power Aware Routing and Wavelength Assignment (PA-RWA) algorithm that tries to set up lightpaths in such a way that the total consumed energy is minimized. Specifically, a routing algorithm, named *Load Based Cost* (LBC), and two different wavelength assignment algorithms, namely *Least Additional Power-First Fit* (LAP-FF) and *Least Cost Wavelength* (LCW), were developed. LBC belongs to the general class of least cost path algorithms but uses a load-dependent cost function to compute link weights. LAP-FF is derived from the FF heuristic whilst LCW is an innovative WA algorithm that works as least cost path routing algorithm by assigning a cost to each wavelength for each link in the path and then looking for the wavelength that minimizes the total cost on the whole path. The LBC algorithm has been compared with the following routing algorithms:

1) Least Congested Path (LCP).
2) Power-Aware Shortest Path (ShP).
3) Most Used Path (MUP) [37].

LAP-FF and LCW have been compared with:

1) First Fit (FF).
2) Two Phases-First Fit (TP-FF).

Moreover, the optimal solution of the problem has also been evaluated.

The main achieved results can be summarized as follows.

i) Both LBC and LCW perform better than the other considered algorithms, showing energy efficiency between 85% and 95%.

ii) LCW performs a little better than LAP-FF which in turn outperforms TF-FF.

iii) LBC is the only algorithm among considered ones that performs close to LCP in terms of blocking probability, especially when used with LCW.

iv) LCW performs better than FF in terms of blocking probability.

REFERENCES

[1] M. Pickavet, W. Vereecken, S. Demeyer, P. Audenaert, B. Vermeulen, C. Develder, D. Colle, B. Dhoedt, and P. Demeester, "Worldwide energy needs for lCT: the rise of power-aware networking," *International Symposium on Advanced Networks and Telecommunication Systems (ANTS'08)*, Bombay, India, Dec. 2008.

[2] M. Gupta and S. Singh, "Greening of the Internet," in Proceedings of SIGCOMM'03: Applications, Technologies, Architectures, and Protocols for Computer Communications, Karlsruhe, 2003, pp.19–26

[3] W. Vereecken, L. Deboosere, D. Colle, B. Vermeulen, M. Pickavet, B. Dhoedt, and P. Demeester, "Energy efficiency in telecommunication networks," *European Conference on Networks and Optical Communications & Optical Cabling and Infrastructure (NOC'08)*, July 2008.

[4] M. Pickavet, W. Vereecken, S. Demeyer, P. Audenaert, D. Colle, C. Develder, and P. Demeester, "Contribution and role of network architectures in the footprint reduction of ICT," *NOC/OC&I'09*, Valladolid, Spain, June 2009.

[5] P. Leisching and M. Pickavet, "Energy footprint of ICT: Forecasts and network solutions," *OFC/NFOEC'09, Workshop on Energy Footprint of ICT: Forecast and Network Solutions*, San Diego, CA, Mar. 2009.

[6] [Online] "Minnesota Internet Traffic Studies", Available: http://www.dtc.umn.edu/mints/2002-2008/analysis-2002-2008.html.

[7] Neilson, D.T., "Photonics for Switching and Routing," IEEE Journal of Selected Topics in Quantum Electronics (JSTQE), vol.12, no.4, pp.669-678, July-Aug. 2006.

[8] M. Bohr, "A 30 Year Retrospective on Dennard's MOSFET Scaling Paper," IEEE SSCS Newsletter, vol. 12, no. 1, pp. 11-13, Winter 2007.

[9] IEEE STANDARD 802.3az-2010 - IEEE Standard for Information technology-- Local and metropolitan area networks-- Specific requirements-- Part 3: CSMA/CD

Access Method and Physical Layer Specifications Amendment 5: Media Access Control Parameters, Physical Layers, and Management Parameters for Energy-Efficient Ethernet.

[10] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti “Energy Efficiency in the Future Internet: A Survey of Existing Approaches and Trends in Energy-Aware Fixed Network Infrastructures”, *IEEE Communications Surveys & Tutorials*, vol.13, no.2, pp.223-244.

[11] European Commission DG INFSO, ”Impacts of Information and Communication Technologies on Energy Efficiency,” final report, Sept. 2008, http://ec.europa.eu/information society/newsroom/cf/itemdetail.cfm ?item id=4441.

[12] J. Baliga, R. Ayre, W. V. Sorin, K. Hinton, and R. S. Tucker, “Energy Consumption in Optical IP Networks”, IEEE/OSA J. Lightw. Tech., vol 27, no. 13,pp. 2391 – 2403, Jul. 2009.

[13] Tucker, R.S.; "Green Optical Communications—Part II: Energy Limitations in Networks," *Selected Topics in Quantum Electronics, IEEE Journal of* , vol.17, no.2, pp.261-274, March-April 2011.

[14] S. Aleksic, “Analysis of Power Consumption in Future High-Capacity Network Nodes” IEEE/OSA J. Opt. Comm. Networking, vol. 1, no. 3, pp 245-258, 2009

[15] R. S. Tucker, “The role of optics and electronics in high-capacity routers”, IEEE/OSA J. Lightw. Technol., vol. 24, pp. 4655–4673, Dec. 2006.

[16] Makoto Murakami and Kazuhiro Oda “Power Consumption Analysis of Optical Cross-Connect Equipment for Future Large Capacity Optical Networks”, in Proceedings of ICTON 2009, Azores Islands, Jul. 2009, pp. 1–4.

[17] Chabarek, J.; Sommers, J.; Barford, P.; Estan, C.; Tsiang, D.; Wright, S.; , "Power Awareness in Network Design and Routing," *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, vol., no., pp.457-465, 13-18 April 2008.

[18] M. Yano, F. Yamagishi and T. Tsuda, “Optical MEMS for Photonic Switching—Compact and Stable Optical Crossconnect Switches for Simple,Fast, and Flexible

Wavelength Applications in Recent Photonic Networks”, IEEE J. Sel. Topics in Quantum Elect., vol. 11, no. 2, Mar./Apr. 2005.

[19] [Online] Available: http://www.garr.it/rete/statistiche/backbone

[20] N. Wang, K. Ho, G. Pavlou, M. Howarth, ”An Overview of Routing Optimization for Internet Traffic Engineering,” IEEE Commun. Surveys Tutorials, vol. 10, no. 1, pp. 36-56, First Quarter 2008.

[21] Van Heddeghem, W.; De Groote, M.; Vereecken, W.; Colle, D.; Pickavet, M.; Demeester, P.; "Energy-efficiency in telecommunications networks: Link-by-link versus end-to-end grooming," *Optical Network Design and Modeling (ONDM), 2010 14th Conference on,* vol., no., pp.1-6, 1-3 Feb. 2010.

[22] Gangxiang Shen; Tucker, R.S.; "Energy-Minimized Design for IP Over WDM Networks," *Optical Communications and Networking, IEEE/OSA Journal of* , vol.1, no.1, pp.176-186, June 2009

[23] Ahmad, A.; Bianco, A.; Bonetto, E.; Cuda, D.; Castillo, G.G.; Neri, F.; , "Power-aware logical topology design heuristics in Wavelength-Routing networks," *Optical Network Design and Modeling (ONDM), 2011 15th International Conference on* , vol., no., pp.1-6, 8-10 Feb. 2011.

[24] Vasic, N., Novakovic, D., Shekhar, S., Bhurat, P., Canini, M. and Kostic, D., Responsive, Energy-Proportional Networks, *EPFL Technical Report*, July 2010.

[25] Chiaraviglio, L.; Ciullo, D.; Mellia, M.; Meo, M.; , "Modeling sleep modes gains with random graphs," *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on* , vol., no., pp.355-360, 10-15 April 2011

[26] K.J. Christensen, C. Gunaratne, B. Nordman, and A.D. George, "The next frontier for communications networks: power management", presented at Computer Communications, 2004, pp.1758-1770.

[27] Idzikowski, F.; Orlowski, S.; Raack, C.; Woesner, H.; Wolisz, A.; , "Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios," *Optical Network Design and Modeling (ONDM), 2010 14th Conference on* , vol., no., pp.1-6, 1-3 Feb. 2010.

[28] L. Chiaraviglio, M. Mellia, and F. Neri, “Reducing Power Consumption in Backbone Networks” IEEE ICC 09, Dresden, pp. 1-6, Jun. 2009.

[29] Cuomo, F.; Abbagnale, A.; Cianfrani, A.; Polverini, M.; , "Keeping the connectivity and saving the energy in the internet," *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on* , vol., no., pp.319-324, 10-15 April 2011.

[30] Amaldi, E.; Capone, A.; Gianoli, L.G.; Mascetti, L.; , "Energy management in IP traffic engineering with Shortest Path routing," *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium on a* , vol., no., pp.1-6, 20-24 June 2011.

[31] Lee S.S.W, Tseng P., Chen A., “Link weight assignment and loop free routing table update for link state routing protocols in energy-aware internet”, *Future Generation Computer Systems*, May 2011.

[32] Cianfrani, A.; Eramo, V.; Listanti, M.; Polverini, M.; , "An OSPF enhancement for energy saving in IP networks," *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on* , vol., no., pp.325-330, 10-15 April 2011.

[33] Mingui Zhang; Cheng Yi; Bin Liu; Beichuan Zhang; , "GreenTE: Power-aware traffic engineering," *Network Protocols (ICNP), 2010 18th IEEE International Conference on* , vol., no., pp.21-30, 5-8 Oct. 2010.

[34] R. S. Tucker, “The role of optics and electronics in high-capacity routers”, IEEE/OSA J. Lightw. Technol., vol. 24, pp. 4655–4673, Dec. 2006.

[35] Eramo, V.; Germoni, A.; Cianfrani, A.; Listanti, M.; Raffaelli, C.; "Evaluation of Power Consumption in Low Spatial Complexity Optical Switching Fabrics," *Selected Topics in Quantum Electronics, IEEE Journal of* , vol.17, no.2, pp.396-405, March-April 2011.

[36] Tucker, R.S.; "Scalability and Energy Consumption of Optical and Electronic Packet Switching," *Lightwave Technology, Journal of* , vol.29, no.16, pp.2410-2421, Aug.15, 2011.

[37] Yong Wu; Chiaraviglio, L.; Mellia, M.; Neri, F.; "Power-Aware Routing and Wavelength Assignment in optical networks," *Optical Communication, 2009. ECOC '09. 35th European Conference on* , vol., no., pp.1-2, 20-24 Sept. 2009.

[38] Wiatr, P.; Monti, P.; Wosinska, L.; "Green lightpath provisioning in transparent WDM networks: Pros and cons," *Advanced Networks and Telecommunication Systems (ANTS), 2010 IEEE 4th International Symposium on* , vol., no., pp.10-12, 16-18 Dec. 2010.

[39] [Online]. Available: http://www.dowslakemicro.com/products/Protocol-Transparent-Optical-Amplification.php.

[40] E.Q.V. Martins, and M.M.B. Pascoal, "A New Implementation of Yen's Ranking Loopless Paths Algorithm", 4OR, vol. 1, no. 2, 2003.

[41] K. Zhu and B. Mukherjee, "Traffic grooming in an optical WDM mesh network," IEEE J. Sel. Areas Commun., vol. 20, no. 1, pp. 122–133, Jan. 2002.

[42] C. Xin, "Blocking analysis of dynamic traffic grooming in mesh WDM opticla networks," IEEE Trans. Networking, vol. 15, no. 3, pp. 721–733, June 2002.

[43] K. Kompella, Y. Rekhter, L. Berger, "Link Bundling in MPLS Traffic Engineering", *IETF RFC 4201*, Oct. 2005, http://www.ietf.org/rfc/rfc4201.txt.

[44] E. Rosen, A. Viswanathan and R. Callon, "Multiprotocol Label Switching Architecture", IETF RFC 3031, Jan. 2001.

[45] ITU Telecommunication Standardization Sector, "Architecture for the automatic switched transport network (ASTN)," *ITU-T Rec. G.807/Y.1302,* Nov. 2001.

[46] J. T. Pin-Han Ho and H. T. Mouftah, "On achieving optimal survivable routing for shared protection in survivable next-generation internet", IEEE Trans. Reliability, vol. 53, no. 2, pp. 216–225, Nov. 2004.

[47] H. Zang, J. Jue, and B. Mukherjee, "A Review of Routing and Wavelength Assignment Approaches for Wavelength-Routed Optical WDM Networks", Opt. Netw. Mag., vol. 14, no. 1, pp. 47-60, Jan. 2000,

[48] Z. Ali et Al. “Graceful Shutdown in MPLS and Generalized MPLS Traffic Engineering Networks”, IETF RFC 5817, 2010.

[49] K. Chan and T. P. Yum, ”Analysis of Least Congested Path Routing in WDM Lightwave Networks,” Proc., IEEE INFOCOM ’94,Toronto, Canada, vol. 2, pp. 962-969, April 1994.

[50] H. Harai, M. Murata, and H. Miyahara, ”Performance of Alternate Routing Methods in All-Optical Switching Networks,” Proc., IEEE INFOCOM ’97, Kobe, Japan, vol. 2, pp. 516-524, April 1997.

[51] S. Ramamurthy and B. Mukherjee, ”Fixed-Alternate Routing and Wavelength Conversion in Wavelength-Routed Optical Networks,” Proc., IEEE GLOBECOM’98, vol. 4, pp. 2295-2302, Nov. 1998.

[52] S. Subramaniam and R. A. Barry, ”Wavelength Assignment in Fixed Routing WDM Networks, Proc., ICC ’97, Montreal, Canada, vol. 1, pp. 406-410, June 1997.

[53] X. Zhang and C. Qiao, “Wavelength Assignment for Dynamic Traffic in Multi-fibre WDM Networks,” Proc., 7th International Conference on Computer Communications and Networks, Lafayette, LA, pp. 479-485, Oct. 1998.

[54] A. Coiro, M. Listanti, A. Valenti and F. Matera, “Power-Aware Routing and Wavelength Assignment in Multi-Fiber Optical Networks” IEEE Journal of Optical Communications and Networking, Vol. 3, Issue 9, pp.816-829, July 2011.

[55] A. Coiro, A. Valenti, F. Matera and M. Settembre, “Reducing Power Consuption in Core WDM Optical Networks”, Fiber and Integrated Optics, Vol. 30, Issue 3, pp. 166-177, June 2011.

[56] A. Coiro, M. Listanti, A. Valenti and F. Matera, “Reducing Power Consumption in Wavelength Routed Networks by Selective Switch Off of Optical Links”, IEEE Journal of Selected Topics in Quantum Electronic, vol.17, no.2, pp.428-436, March-April 2011.

[57] A. Silvestri, A. Valenti, S. Pompei, F. Matera, A. Cianfrani and A. Coiro, "Energy saving in optical transport networks exploiting transmission properties and

wavelength path optimization", Elsevier Optical Switching and Networking, Volume 7, Issue 3, July 2010.

[58] A. Coiro, F. Iervini and M. Listanti, “Distributed and Adaptice Interface Switch Off for Internet Energy Saving”, International Conference on Computer Communication Networks 2011, Maui, August 2011.

[59] A. Coiro, M. Listanti, A. Valenti, “Dynamic Power-Aware Routing and Wavelength Assignment for Green WDM Optical Networks” International Conference on Communication 2011, Kyoto, Jun 2011.