Leveraging MPLS Backup Paths for Distributed Energy-Aware Traffic Engineering

Frederic Francois, Ning Wang, Klaus Moessner, Stylianos Georgoulas and Ricardo de O. Schmidt

Abstract—Backup paths are usually pre-installed by network operators to protect against single link failures in backbone networks which use Multi-Protocol Label Switching (MPLS). This paper introduces a new scheme called Green Backup Paths (GBP) which intelligently exploits these existing backup paths to perform energy-aware traffic engineering without adversely impacting the primary role of these backup paths of preventing traffic loss upon single link failures. This is in sharp contrast to most existing schemes which tackle energy efficiency and link failure protection separately, resulting in substantially high operational costs. GBP works in an online and distributed fashion where each router periodically monitors its local traffic conditions and cooperatively determines how to reroute traffic so that the highest number of physical links can go to sleep for energy saving. Furthermore, our approach maintains Quality-of-Service by restricting the use of long backup paths for failure protection only and therefore, GBP avoids substantially increased packet delays. GBP was evaluated on the Point-of-Presence representation of two publicly-available network topologies, namely GÉANT and Abilene, and their real traffic matrices. GBP was able to achieve significant energy saving gains which are always within 15% of the theoretical upper bound.

Index Terms—Green networks; MPLS; backup paths; distributed; online; energy efficiency; traffic engineering; failure protection.

I. INTRODUCTION

Network operators have to allocate an increasing amount of their operating budget to electricity bills nowadays, and this is due to the operation of a larger number of network devices in order to meet higher traffic demands and also the increasing price of electricity [1]. European telecom operators currently consume 21.4TWh per year and this is expected to increase to 35.8TWh by 2020 if no green networking technologies are introduced [2]. While backbone networks consume only 10% of the total energy consumption of the global network infrastructure, this is expected to rise to 40% by 2017 if no green actions are taken because more cloud-based applications and services are going to be deployed [3].

Manuscript received September 6, 2013; revised February 15, 2013. The associate editor coordinating the review of this paper and approving it for publication was Dr. Chonggang Wang.

The research leading to these results has been performed within the UniverSelf (www.UniverSelf-project.eu) and Flamingo (www.fp7-flamingo.eu) projects and received funding from the European Community's Seventh Framework Programme under grant agreement no. 257513 and no. 318488 respectively.

Frederic Francois, Ning Wang, Klaus Moessner, and Stylianos Georgoulas are with the Centre for Communication Systems Research, University of Surrey, Guildford GU2 7XH, UK (e-mail: {f.francois, n.wang, k.moessner, s.georgoulas}@surrey.ac.uk). Ricardo de O. Schmidt is with the Design and Analysis of Communication Systems, University of Twente, NL (email: r.schmidt@utwente.nl).

Fig. 1. Basic network topology to illustrate how links are protected in MPLS backbone networks.

Nowadays, it is common for backbone operators to use Multi-Protocol Label Switching (MPLS) to explicitly route traffic between the different Source-Destination (SD) pairs in their networks. These backbone networks are protected against single link failures through the use of pre-installed backup paths [4], [5], [6]. A backup path is used to divert the affected traffic away from the protected link when it fails. The route taken by the backup path is usually the shortest path between the head and tail router of the protected link but without traversing the protected link. In such local protection, the failure recovery is handled only by the head router of the protected link, and none of the other remote routers needs to be aware of the failure if it occurs. The illustrative network topology in Fig. 1 is used to demonstrate how the traffic on a link is protected against the failure of the link by a pre-installed backup path. For example, if the link $B \rightarrow C$ fails, the head router B of the link will divert the flow from A to B onto the pre-installed backup path $BP1$ to avoid traffic loss.

This paper introduces a novel online and fully-distributed Energy-aware Traffic Engineering (ETE) scheme called Green Backup Paths (GBP). GBP improves the power/energy-efficiency of networks by opportunistically diverting traffic away from protected links onto the backup paths in an intelligent manner, so that the protected links can go to sleep. GBP directly uses the existing backup paths for energy savings while not impairing the ability of these paths to protect against single link failures by implementing a novel link failure protection mechanism. Specifically, during normal network operations, some backup paths can be exploited for diverting traffic from their protected links in order to allow them to sleep, while upon the detection of unexpected failure of a working link, its associated backup path will switch back to its original role for traffic recovery. In this scenario, there can be interference between this backup path for failure recovery and other backup paths for energy efficiency purposes, as

these paths may share some common link(s) which cannot accommodate the traffic load upon the post failure traffic diversion. To address this issue, GBP can gracefully disable some active backup paths used for energy efficiency purposes in order to avoid congestion, but at the same time avoid completely sacrificing energy saving gains for the sake of recovery. This novel protection mechanism, as will be shown later, allows GBP to maximize the energy savings during failure-free scenarios by making use of all available backup paths and their capacity while avoiding traffic congestion during single link failures. In addition, GBP also considers the traditional traffic engineering function of resilience of the network against potential traffic upsurges by not causing any link to become overloaded during any of its operations. On the contrary, GBP actively attempts to reduce the traffic on overloaded links so that the peak link utilization decreases in the network. Hence, a second objective of GBP is to increase the resilience of the network against potential traffic upsurges, which is in addition to its other objective of energy savings.

The key novelty of GBP is the exploitation of existing failure-protection backup paths for the dual purpose of energy savings and protection against link failures. This brings the main benefit of achieving energy-efficiency without installing any other paths in addition to the backup paths, which are needed anyway for failure protection. Indeed, GBP differs significantly from most other existing online traffic engineering schemes (e.g. [7], [8], [9]) which target either energy savings or link failure protection but not both. Furthermore, GBP considers Quality-of-Service (QoS) by actively avoiding the use of excessively long backup paths for energy savings so as to avoid substantial packet delays, but allows the use of such paths for handling link failures. An additional advantage of GBP is its fast path manipulation. As it will be shown later, multiple routers can make concurrent conflict-free decisions at the same time thanks to their knowledge of interference relationships with each other. Moreover, GBP uses only a single path to route each SD flow and therefore, avoids packet reordering linked with multi-path routing.

In order to evaluate the performance of GBP, the publicly-available topology and real traffic matrices of two academic backbone networks were used, namely GÉANT and Abilene. It was observed that GBP can achieve significant energy savings which are always within 15% of the theoretical upper bound. This result was achieved without any increase in the peak Maximum Link Utilization (MLU) of the network as a trade-off. In addition, the ability of GBP to reduce the MLU in the network was also evaluated. According to the evaluation results, GBP was able to even significantly reduce the MLU in the case of GÉANT where there is a large diversity of paths and enough spare capacity. Furthermore, single link failures were simulated in the network and it was observed that the use of GBP did not increase the post-failure peak MLU. In addition, the increase in the maximum packet delay due to the use of the backup paths was also found to be minimal and acceptable. Therefore, QoS constraints linked with delay can be met while performing online sleeping reconfigurations through the necessary traffic diversion by GBP.

The rest of this paper is organized as follows: in Section II, the problem formulation for the GBP scheme is described in detail. In Section III, an overview of GBP is first presented and then an extensive description of each component of GBP is provided along with pseudo codes. In Section IV, the results from the evaluation of GBP on the GÉANT and Abilene topologies are presented. In Section V, an overview of other existing ETE schemes is provided. Finally in Section VI, we conclude the paper with our key findings.

II. PROBLEM FORMULATION

Nowadays, a logical link between router pairs in networks is usually made up of a bundle of physical links [10], [11]. Such a strategy reduces the complexity in upgrading network capacities by adding new physical links to the existing bundle. If traffic demands are lower than the capacity of the whole bundle, energy savings can be achieved by putting unused physical links to sleep but without changing the logical network topology. In addition, the line card connected to a physical link can have the opportunity to sleep when the physical link is put to sleep. Sleeping line cards are the major source of energy savings in green networks because they contribute up to 42% of the total energy consumption of a backbone router [7]. Putting part of a logical link to sleep can be viewed as a form of rate adaptation, which is similar to what has been developed for green Ethernet [12].

The actual online optimization within each periodical GBP operation cycle can be expressed as:

$$\text{minimize } f - \frac{\alpha}{100} c_l \quad \forall f_l > \frac{\alpha}{100} c_l, \text{ with } \alpha = [0, 100] \quad (1)$$

$$\text{maximize } \sum_{l=1}^{|L|} (y_l \times p_l) \quad (2)$$

subject to

$$\sum_{j=1}^{|R|} b_{ij}^{sd} - \sum_{j=1}^{|R|} b_{ji}^{sd} = \begin{cases} 1 & \forall s, d, i = s \\ -1 & \forall s, d, i = d \\ 0 & \forall s, d, i \neq s, d \end{cases} \quad (3)$$

$$\sum_{j=1}^{|R|} b_{ij}^{sd} f_{ij}^{sd} - \sum_{j=1}^{|R|} b_{ji}^{sd} f_{ji}^{sd} = \begin{cases} t^{sd} & \forall s, d, i = s \\ -t^{sd} & \forall s, d, i = d \\ 0 & \forall s, d, i \neq s, d \end{cases} \quad (4)$$

$$f_l < \frac{\alpha}{100} c_l \quad \forall l \text{ with } \alpha = [0, 100] \quad (5)$$

The symbols in all equations above are explained in Table I. Eq. (1) is the first objective of GBP which is to minimize the Maximum Logical Link Utilization (MLLU) in the network so that the network is more resilient to traffic upsurges because of the more balanced load. The MLLU was referred to as the MLU before the introduction of the concept of bundle links in this section. Eq. (2) represents the second objective of GBP which is to maximize the total amount of energy saved in network operations. This is represented by the sum (over all logical links of the network) of the product of the number of sleeping physical links in a logical link and the energy consumed by the physical links if they were left active. Eq. (3)

TABLE I
DEFINITION OF SYMBOLS

Variable	Description
$G(R, L)$	Directed graph with R being set of routers and L being set of logical links.
y_l	Number of physical links in logical link l which are in sleep mode.
p_l	Power consumed by an active physical link in logical link l .
c_l	Capacity of logical link l .
t^{sd}	Traffic demand from router s to d .
b_{ij}^{sd}	Specifies if a logical link from router i to j is used to route traffic from router s to d . A value of "1" means the logical link is used, otherwise it is "0".
f_{ij}^{sd}	Traffic demand from router s to d that traverses logical link from router i to j .
f_l	Total traffic demand on logical link l .
α	Maximum allowable utilization of logical link.

is the constraint which enforces a single path to be taken by all traffic which has the same source and destination. Equation (4) is the conventional flow conservation constraint. Eq. (5) prevents a logical link from being loaded above the threshold, α , due to the operation of GBP. Moreover, GBP does not use backup paths for energy saving if their path length (delay) is too long but they will still be used for link failure protection.

The simpler problem of maximizing the number of physical links which can go to sleep while respecting the above constraints has already been proven to be NP-hard in [10]. Therefore, we present a computationally-efficient heuristic called *Green Backup Paths* (GBP) which can be applied in a network in an online and distributed fashion without requiring significant modifications to existing network protocols.

III. GREEN BACKUP PATHS

In this section, an overview of GBP is first presented followed by an in-depth description of all its different components. Table II acts as a point-of-reference for the name and description of all the notions used during the description of the operation of GBP.

A. Scheme Overview

The proposed GBP scheme consists of two distinct operational components, namely the *offline* and the *online* components. The offline component identifies the eligible backup paths for GBP operations. This is done based on the delay (length) characteristics of the paths. Network operators can obtain the delay of a path by using conventional end-to-end network measurement techniques. It is worth noting that in GBP, the primary paths are configured by using the shortest path algorithm with the link delay being the routing metric. In the same style, the backup paths are configured on per link basis where a backup path for a protected link follows the shortest route (according to delay as well) between the head and tail router of that link but without passing through it.

The offline component also identifies an *Interference-Risk Links List* (IRLL) for each logical link in the network. The IRLL for a logical link contains all the logical links that

Fig. 2. Timeline for the online operation of GBP.

can be potentially affected by this link if it is offloaded. The IRLLs are essential for GBP to be able to concurrently and independently offload multiple logical links without any conflict. The details of how the IRLLs are obtained and used will be described in Section III.B.2 and III.C.3. If the alternative path for a logical link introduces substantially longer delay, the network operator may choose not to offload such a logical link for energy savings and MLLU reduction purposes. Such logical links are also identified in the offline phase. The IRLL and eligibility for GBP operations for each logical link are distributed to the routers only once since this information is static as long as the network topology is not changed.

The second component of GBP performs an online optimization by using a heuristic to periodically divert traffic away from logical links by activating/deactivating backup paths in order to optimize both the energy savings and the MLLU within the network. The periodicity of the online operation can be determined by the network operator as a trade-off between the overhead of monitoring the network and the need to detect any significant changes in the network traffic condition.

Fig. 2 shows an illustration of the online GBP operation cycle. When a new GBP cycle starts, each router collects information about the traffic conditions by receiving the *Traffic Engineering-Link State Advertisements* (TE-LSAs) [13] that are broadcasted by every router in the network. Based on this traffic information, each router can then calculate whether any of its directly attached logical links can successfully offload part of their traffic onto alternative paths to save energy and/or reduce the MLLU. If sufficient traffic offloading is achieved, one or more physical links in the concerned logical link can go to the sleep mode. In GBP, the head routers are responsible for determining how many active physical links in each of their logical links should be put to sleep, so that only the minimum number of physical links is active without causing any traffic congestion. This design choice is made so as to maximize the energy savings in the network but without compromising on post-failure traffic loss and the ability of the logical links to handle sudden traffic surges.

At the beginning of a GBP optimization cycle, each router needs to check if there are any logical links which have become overloaded, i.e. do not comply with the constraint in Eq. (5) because of the sleeping reconfigurations in previous GBP optimization cycles. This may happen due to the increased volume of incoming traffic since the last traffic monitoring observation. In case such logical links are identified, the associated router will wake up some sleeping physical links in the relevant protected logical links and restore the currently diverted traffic back to the protected logical link(s). As a result,

TABLE II
DESCRIPTION OF ALL NOTIONS

Name	Description
<i>IRLL</i>	Each logical link has an Interference-Risk Links List (IRLL) to store the logical links whose spare capacity must not be modified when the logical link is undergoing offloading by GBP.
<i>TH_link</i>	A logical link which has been selected by GBP for part of its traffic to be offloaded to an alternate route.
<i>TH_router</i>	A router which is the head router of a TH link and therefore, it is the network device which is responsible to attempt the traffic reroute away from the TH link.
<i>priority_links_list</i>	List which contains all the logical links that have their utilization above a pre-defined threshold α .
<i>normal_links_list</i>	All logical links not in <i>priority_links_list</i> .
<i>conflict_links_set</i>	List of all logical links which are already in use by TH routers in the current Multiple TH Links Selection iteration.
<i>p_flows_list</i>	List of all SD flows that normally use the logical link.
<i>b_flows_list</i>	List of all SD flows that were diverted onto the logical link by GBP and the logical link has the same head router as the protected link.
<i>s_flows_list</i>	List of all SD flows on the logical link which are not in the <i>p_flows_list</i> and <i>b_flows_list</i> .

the previously active backup path is deactivated and traffic is no longer diverted on its links. After each router deactivates these overloaded backup paths, they wait for a settling period and then broadcast a new TE-LSA to notify all other routers about the new state of their logical links.

Each router then continues the online decision process by collecting the new TE-LSAs and updating the list of logical links. The list of logical links is always sorted at each router such that all routers have an identically ordered list. Each router goes through the list and selects the Token Holding (TH) links that can be concurrently offloaded without interfering with each other. A TH link is a logical link selected by GBP for part of its current traffic to be rerouted so that its overall traffic load is reduced, potentially reducing its energy consumption by putting a subset of its physical links to sleep. Each router is aware of the interference-free TH links due to the pre-calculated IRLLs (see Section III.B.2).

If the router is the head router of an interference-free Token Holding (TH) link, it becomes a TH router and is responsible for locally offloading that TH link. Since multiple non-interfering TH links can be concurrently selected, GBP can converge quicker compared to other ETE schemes [10], [11], [14] which are based on purely sequential operations. A non-TH router does not offload any logical link unless it becomes the head router of a selected TH link during forthcoming selection rounds. Moreover, a logical link can be selected to become a TH link only once per GBP cycle.

TH routers broadcast an *operation-completed* message when they have finished operating on all their TH links. Routers only broadcast the new TE-LSAs upon receiving *operation-completed* messages from all current TH routers (all routers in the network know which routers are TH routers because

they all compute the links which are TH links). Each router then repeats the process of selecting the TH links. The GBP cycle stops after the list of logical links is exhausted, i.e. all logical links have become TH links.

GBP always achieves loop-freedom in traffic diversion because for any diverted traffic away from the primary path, the termination node of the backup path is guaranteed to be in the downstream location of the head node from where the backup path is branched out. In addition, the traffic rerouted on a backup path is not allowed to be rerouted onto another backup path by GBP and hence, it is not possible for traffic to be diverted in a recursive manner further away from the backup paths.

Routing oscillations cannot happen within one GBP cycle. The reason is when a link is offloaded in GBP; this is done so that the link goes down in energy level and/or link utilization. Therefore, it is not possible for GBP to reroute any traffic back onto the same link within the same GBP cycle because this will cause the link to break the objectives of GBP which is to lower the energy consumption and/or reduced the link utilization below the set threshold α .

B. Offline Component

The offline component of GBP consists of two stages. The first stage is responsible for the identification of the backup paths which are eligible for participation in GBP operations and this is done by filtering out the backup paths with excessive end-to-end delay. The second stage is the generation of the IRLL for each logical link of the network. The IRLLs will allow several logical links to concurrently have their traffic diverted without any conflict that could lead to adverse effects on the network such as traffic congestion.

1) *Identification of Eligible Backup Paths*: In this stage the offline component of GBP identifies the backup paths that are eligible to participate in GBP operations based on the end-to-end delay of the backup path compared to its protected logical link. Each backup path entry in the MPLS label table of a router has an associated binary bit, *delay_ok*, which is set to 1 if the path meets the constraint on the maximum path length as determined by the network operator or 0 otherwise. If *delay_ok* is 0, the backup path is only used for link failure protection. For simplicity, from now on we only consider the identified eligible backup paths.

2) *Generation of IRLLs*: The second stage of the offline component calculates the IRLL for each logical link in the network. The interference-risk links of a TH link are defined as *all logical links which must not be used by other TH links to divert traffic to when that specific TH link is undergoing offloading*. Taking the network topology in Fig. 3 as example, when logical link $B \rightarrow C$ becomes a TH link, no other TH links are allowed to divert traffic onto logical link $B \rightarrow C$ and the logical links of the backup path $BP1$ which consists of logical links $B \rightarrow E$ and $E \rightarrow C$. Therefore, logical links $B \rightarrow C$, $B \rightarrow E$ and $E \rightarrow C$ are in the IRLL of logical link $B \rightarrow C$. In addition, a TH link may have some Source-Destination (SD) flows which are currently diverted on it. GBP allows a TH link to divert these flows back to their original

Fig. 3. Basic network topology to illustrate how IRLs are generated.

respective protected logical links if the head router of the SD flows is also the head router of the TH link. Therefore, logical link $B \rightarrow F$ is also added to the IRL of TH link $B \rightarrow C$ in order to allow $B \rightarrow C$ to deactivate the backup path BP2 so that the diverted SD flow BF is re-routed back to the original protected logical link $B \rightarrow F$. Hence, TH link $B \rightarrow C$ has an IRL consisting of logical links $B \rightarrow C$ (itself), $B \rightarrow E$, $E \rightarrow C$ and $B \rightarrow F$. The description of how IRLs are used to avoid interference when multiple TH links are concurrently offloaded will be given in Section III.C.3.

C. Online Component

Fig. 4 shows a top-level view of the online component of GBP. This component consists of four different stages. At the start of each GBP optimization cycle, each router in the network needs to collect link state information from other routers in the network so as to get an updated and consistent view of the state of the network. Following this gathering of information, the second stage of GBP is performed where routers may need to deactivate some already activated backup paths because the diversion of traffic on them has led to the logical links constituting these paths to become overloaded.

After the deactivation of the overloaded backup paths, GBP continues the optimization process by choosing logical links which have not been offloaded in this optimization cycle and do not conflict with each other according to the IRLs. Attempts are then made to divert traffic from the selected logical links so that their energy consumption and/or utilization go down while not overloading any paths.

The next step is for all routers to broadcast the state of their logical links so that the new state of the network is captured by all routers. In the same manner as before, a new set of unselected logical links is then selected to have their traffic diverted. This iterative process of selecting logical links to have their traffic diverted is continued until all logical links in the network have been considered in the current GBP optimization cycle. In the remaining part of this section, all the four different stages of the online component of GBP are described in more detail.

1) *Stage 1: Gathering the State of the Network:* At the start of each GBP optimization cycle, all routers need to concurrently collect the broadcasted information about the state of all logical links in the network. This procedure can leverage the TE-LSAs, which are already specified in the suit of traffic engineering protocols such as OSPF-TE [15].

Fig. 4. Flow chart showing the different operations in one GBP optimization cycle.

GBP requires two types of information about the logical links from the TE-LSAs, namely the current load and value of the TH_status_flag . Each logical link has a binary bit called TH_status_flag which is set to 0 at the beginning of each GBP optimization cycle and to 1 after its associated logical link has become TH link in the current GBP optimization cycle. This prevents a logical link from becoming a TH link again in that particular cycle.

2) *Stage 2: Deactivation of Overloaded Backup Paths:* After routers have collected information from TE-LSAs, they verify that the logical links in activated backup paths are not overloaded. This situation may be caused by an increase in traffic volume. If such overloaded backup paths are identified, they are deactivated to relieve their overloaded logical links. The pseudo code¹ to perform this operation is given in Alg. 1 and the algorithmic complexity of Stage 2 is $O(|L|^2)$.

3) *Stage 3: Selection of Multiple TH Links:* The main purpose of Stage 3 is to calculate which TH links can be selected at the same time without any interference. After the deactivation of the overloaded backup paths, routers broadcast new TE-LSAs so that others are aware of the new state of the network. On receiving the new TE-LSAs, each router forms a list of logical links that excludes all logical links which have their TH_status_flag equal to 1. Initially all logical links will be included since their respective TH_status_flag is set to 0 at the start of each GBP optimization cycle. Given that all routers have the same view of the state of the network through the new

¹In all pseudo codes in this paper, $x.y$ means y is a property/variable of x . In addition, $x[y]$ means x is a list where y is a position in the list.

Algorithm 1: Deactivate_Overloaded_Backup_Paths

```

1 begin
2   every router  $r$  in  $R$ 
3     foreach backup_path of  $r$  do
4       if backup_path.activated == true then
5         foreach link  $l$  of backup_Path do
6           if  $l.utilization > \alpha$  then
7             backup_path.activated = false
8             break
9
10
11
12

```

broadcasted TE-LSAs, they will therefore form an identical list.

The list of logical links is partitioned into two disjoint sub-lists, namely *priority_links_list* and *normal_links_list*. The *priority_links_list* contains all the logical links that have their utilization above a predefined threshold α and therefore, violate the constraint in Eq. (5). These have priority to become TH links because if they are successfully offloaded, the resilience of the network against traffic upsurges will improve. The *priority_links_list* is then sorted in descending order based on the excess load of the logical links. This excess load, x_l , is defined by Eq. (6) where z_l is the bandwidth capacity of one physical link of the logical link l .

$$x_l = \max \left(\text{mod} \left(\frac{f_l}{z_l} \right), f_l - \frac{\alpha}{100} c_l \right) \quad (6)$$

The first term of the maximum function in Eq. (6) represents the excess load on the TH link that prevents the TH link from going to the next lower power level by putting an additional physical link to sleep. The second term is used to calculate the excess load on the TH link that prevents its utilization from dropping below $\alpha\%$ of its total capacity.

The impact of the offloading of highly-utilized logical links on other logical links is minimized by setting a limit on the maximum spare capacity of the other logical links in the network, so that their utilization does not exceed the predefined threshold α . The design choice of prioritizing the offload of highly-utilized links can be seen as a way of improving the resilience of the network against traffic upsurges.

After all logical links from the *priority_links_list* have become TH links in the current GBP optimization cycle; it is the turn of the ones in the *normal_links_list*. Unlike the previous list, this one is sorted in ascending order according to the excess load, which is calculated by using Eq. (6). This is because it is easier to offload small excess loads to alternative paths and therefore, achieve greater energy savings. The overall process of generating the two sub-lists is described in Alg. 2.

The second part of Stage 3 involves the selection of multiple logical links that can concurrently become TH links. The pseudo code for this part is given in Alg. 3. Each router has an initially empty set called *conflict_links_set*, which is also emptied after each iteration of Multiple TH Links Selection

Algorithm 2: Generate_Links_Lists

```

1 begin
2   foreach link  $l$  in  $L$  do
3     if  $l.TH\_status\_flag == 0$  then
4       Add  $l$  to links_list
5   foreach link  $l$  in links_list do
6     if  $l.utilization > \alpha$  then
7       Add  $l$  to priority_links_list
8     else
9       Add  $l$  to normal_links_list
10  DescendingSort(priority_links_list)
11  AscendingSort(normal_links_list)

```

Algorithm 3: Select_Multiple_TH_Links

```

1 begin
2   conflict_links_set = 0
3   X = 0
4   if priority_links_list.size > 0 then
5     while  $X < \text{priority\_links\_list.size}$  do
6       if all links in priority_links_list[X].IRLL not in
7         conflict_links_set then
8         Add priority_links_list[X] to TH_links_list
9         priority_links_list[X].TH_status_flag = 1
10        Add priority_links_list[X].IRLL to
11        conflict_links_set
12      X++
13   else
14     while  $X < \text{normal\_links\_list.size}$  do
15       if all links in normal_links_list[X].IRLL not in
16         conflict_links_set then
17         Add normal_links_list[X] to TH_links_list
18         normal_links_list[X].TH_status_flag = 1
19         Add normal_links_list[X].IRLL to
20         conflict_links_set
21      X++

```

in a single GBP optimization cycle. The router goes through the *priority_links_list* and if the logical link does not have any logical link of its IRLL in the *conflict_links_set*, the router makes the logical link become a TH link and adds its IRLL links to the *conflict_links_set*. The example topology in Fig. 4 can be used to illustrate this process. If the logical link $B \rightarrow C$ is the first selected TH link in a Multiple TH Links Selection iteration, its IRLL links (i.e., links $B \rightarrow C$, $B \rightarrow E$, $E \rightarrow C$ and $B \rightarrow F$) are added to the *conflict_links_set*. Any subsequent selected TH links in this iteration must have none of their IRLL links in the *conflict_links_set*. For example, logical link $B \rightarrow F$ cannot become TH link in this iteration because some of its IRLL links (i.e. $B \rightarrow C$ and $B \rightarrow F$) are already in the *conflict_links_set*. Since $B \rightarrow C$ cannot become TH link again during the current GBP optimization cycle, $B \rightarrow F$ will have the opportunity to become TH link during the next iteration of the Multiple TH Links Selection. When a logical link becomes a TH link, its *TH_status_flag* is

Algorithm 4: Offload_TH_Link

```

1 begin
2    $x_l = \text{Calculate\_Excess\_Load}()$ 
3    $\text{Offload\_b\_Flows}()$ 
4   if  $x_l > 0$  then
5      $\text{Offload\_p\_Flows}()$ 
6   if  $x_l \leq 0$  or  $\text{TH\_link.utilization} > \alpha$  then
7     foreach flow  $w$  of  $\text{flows\_to\_reroute\_list}$  do
8       if  $w.\text{backup\_path.activated} == \text{true}$  then
9          $w.\text{backup\_path.activated} = \text{false}$ 
10      else
11         $w.\text{backup\_path.activated} = \text{true}$ 

```

set to 1. After going through the whole *priority_links_list*, the router performs the same selection procedure for links in the *normal_links_list*.

When a router has finished calculating the TH links and under the condition that it is the head router of at least one TH link, it becomes a TH router. A TH router will attempt to offload its TH links through the process in Stage 4, described next, and then broadcast an *operation-completed* message to all other routers in the network upon concluding the whole operation. Routers in the network will broadcast a new TE-LSA immediately after they have received the *operation-completed* message from all the current TH routers. The next iterations for the Multiple TH Links Selection in the current GBP optimization cycle can begin after routers receive all the new TE-LSAs. The overall algorithmic complexity of Stage 3 is $O(|L|^2)$.

4) *Stage 4: Offloading of the Token Holding Link*: The overall pseudo code for Stage 4 is given in Alg. 4. Each TH router has three lists of SD flows for each of its logical links and they are used to classify all the SD flows on the logical link. The first list is the *p_flows_list* which is a list of all SD flows that normally use the link, i.e. the flows were not diverted onto the logical link by GBP. The second list is the *b_flows_list* which is a list of all SD flows that were diverted onto the logical link by GBP and the logical link has the same head router as the protected link. The third list is the *s_flows_list* which contains all the remaining SD flows on the logical link. It is worth mentioning that the traffic load on a path basis does not need to be distributed at any point to other routers through TE-LSAs; this per path traffic monitoring process is local to each router with only the aggregate traffic load on a link basis and the *TH_status_flag* values only needing to be distributed through TE-LSAs, following the process described in Section III. A.

The TH router has direct control over the SD flows in the *p_flows_list* and *b_flows_list* because the TH router acts as the head router for these flows and therefore, it can decide whether to route these SD flows on either their original protected link or backup path. The selection between routing either on the protected link or backup path is based on where the SD flow is currently routed and whether there is enough spare capacity on the alternate route to support the SD flow without either increasing the energy consumption of that route or overloading the logical links of the alternate route. The spare capacity, h_l ,

of a logical link l is given by

$$h_l = \min \left(z_l - \text{mod} \left(\frac{f_l}{z_l} \right), \frac{\alpha}{100} c_l - f_l \right) \quad (7)$$

where the first term of the minimum function is the amount of traffic that can be added to a logical link without this link going to the next higher power level by waking up an additional physical link. The second term restricts the amount of traffic that can be added to a logical link so that its overall utilization percentage does not go above the predefined threshold α . If Eq. (7) results in less than zero, then h_l is 0 meaning there is no spare capacity.

In the first step of Stage 4, each TH router selects one of its TH links and calculates the excess load on the link using Eq. (6). Since the SD flows in the *p_flows_list* and *b_flows_list* of the TH link are under the direct control of the TH router because it is the head router of these flows, they are the only SD flows targeted for removal from the TH link. The flows in the *b_flows_list* of the TH link are first targeted. By rerouting them to their respective original protected links, the TH link will be offloaded and additionally, both delay and wastage of bandwidth will be reduced because of the shorter path taken by the SD flow. The *b_flows_list* is sorted in descending order according to load so that the least number of SD flows are moved back to their respective protected link when the excess load on a TH link is removed. Hence, allowing the offloading of the TH link to be quicker because the smallest number of reconfigurations is done.

The decision whether a flow in the *b_flows_list* can be moved back to the protected logical link depends on the spare capacity of the respective protected logical link. The spare capacity is calculated by using Eq. (7). If the spare capacity of the protected link is larger than the size of the diverted flow, then the flow is added to the list of flows, *flows_to_reroute_list*, to be rerouted. The information contained about the logical links in the TH router needs to be updated to reflect that the traffic in a flow is to be rerouted. As the TH router knows all the logical links involved in the backup path rerouting, it can decrease the load of these logical links and therefore increase their spare capacity. The protected logical link of the backup path will have its load increased and consequently its spare capacity decreased because some of its previously-diverted SD flows have been rerouted back on it. The amount of excess load on the TH link is reduced by the total size of the rerouted SD flows. If the excess load is still above zero, the next flow in the *b_flows_list* is selected for rerouting. The algorithmic complexity of rerouting diverted SD flows back to their protected link is $O(|R|^2(|L| + \log |R|^2))$ and the pseudo code for this step is presented in Alg. 5.

In the third step of Stage 4 if the excess load is still greater than zero and all the flows in the *b_flows_list* have become candidate for rerouting, the flows in the *p_flows_list* are then considered for rerouting. The process is similar to the one for flows in the *b_flows_list* and its pseudo code is presented in Alg. 6. First, the list is sorted in descending order according to the size of the flows. An SD flow can be diverted to the backup path of the protected TH link if the *delay_ok* bit of the backup path is equal to 1. This enables the backup path to accept SD

Algorithm 5: Offload_b_Flows

```

1 begin
2   b_flows_list = Sort_b_flows_DescendingOrder()
3   X = 0
4   while X < b_flows_list.size and xl > 0 do
5     if hl of protected_link ≥ b_flows_list[X].load then
6       xl = xl - b_flows_list[X].load
7       hl of protected_link = hl of protected_link -
8         b_flows_list[X].load
9       hl of all links of p_flows_list[X].backup_path = hl
10      of all links of p_flows_list[X].backup_path -
11        b_flows_list[X].load
12      Add b_flows_list[X] to flows_to_reroute_list
13    X++

```

Algorithm 6: Offload_p_Flows

```

1 begin
2   p_flows_list = Sort_p_flows_DescendingOrder()
3   X = 0
4   while X < p_flows_list.size and xl > 0 do
5     if p_flows_list[X].backup_path.delay_ok == 0 and hl
6       of p_flows_list[X].backup_path ≥ p_flows_list[X].load
7       then
8         xl = xl - p_flows_list[X].load
9         hl of all links of p_flows_list[X].backup_path = hl
10        of all links of p_flows_list[X].backup_path +
11          p_flows_list[X].load
12        hl of p_flows_list[X].protected_link = hl of
13        p_flows_list[X].protected_link -
14          p_flows_list[X].load
15        Add p_flows_list[X] to flows_to_reroute_list
16      X++

```

flows to be diverted on it. The spare capacity of the backup path for this SD flow on the TH link is calculated by taking the minimum spare capacity of all the logical links involved in the backup path with the spare capacity of a logical link being calculated using Eq. (7). If the spare capacity of the backup path is greater than the size of the SD flow to be rerouted, the SD flow is added to the *flows_to_reroute_list*. The spare capacity and load of the logical links, involved in this flow rerouting, are updated.

This process continues until either the excess load is equal/less than zero or all the SD flows in the *p_flows_list* have become candidate for rerouting. The algorithmic complexity of offloading the flows is $O(|R|^2(|L| + \log |R|^2))$. The final step of Stage 4 is to implement all the SD flow reroutes if either the excess load is less/equal to zero or the utilization threshold of the TH link is above the predefined threshold α . The second criterion is used in the case where the excess load is still greater than zero, meaning that not enough SD flows have been successfully rerouted but it is still desirable to implement all the successful SD flow reroutes because this will decrease the load of an overloaded TH link and make it more

Fig. 5. (a) illustrative topology to demonstrate the conventional failure-protection mechanism in an MPLS-enabled network and (b) illustrative topology to demonstrate GBP enhanced failure-protection mechanism.

resilient to traffic upsurges. Due to the way the spare capacity of a logical link is calculated, it is not possible for GBP to overload a logical link above a set threshold α while offloading other logical links. The algorithmic complexity for this step is $O(|R|^2)$. The overall algorithmic complexity of Stage 4 and of a whole GBP optimization cycle are $O(|R|^2(|L| + \log |R|^2))$ and $O(|L|^2 + |R|^2(|L| + \log |R|^2))$ respectively

D. Handling of Logical Link Failures when GBP is active

When GBP is active in a MPLS-enabled backbone network, single logical link failures are handled by two mechanisms: the conventional failure protection mechanism and a GBP enhanced failure-protection mechanism. The conventional failure-protection mechanism is applied regardless of whether GBP is active or not in the network.

When a logical link fails in an MPLS-enabled backbone network, the head router of the failed logical link will divert the SD flows in the failed logical link to its backup path. This conventional failure protection mechanism can be illustrated with the simple example topology in Fig. 5a where all the logical links have a capacity of 100Mbps. If the logical link $B \rightarrow C$ fails, its traffic is diverted by the failure protection mechanism onto its backup path BP1.

For example, if $B \rightarrow C$ was initially carrying 50Mbps of traffic, upon its failure the 50Mbps traffic will be diverted on BP1 which consists of logical links $B \rightarrow E$ and $E \rightarrow C$. It should be noted that the logical links $B \rightarrow E$ and $E \rightarrow C$ may be carrying their own traffic (as they can be involved in other default and backup paths), and the diverted traffic from the failed logical link $B \rightarrow C$ will add to this demand. If the utilization of a logical link is greater than 100% of its capacity, the excess traffic on that link will be lost due to congestion. For example, if $B \rightarrow E$ and $E \rightarrow C$ were initially carrying 50Mbps and 60Mbps before the failure of $B \rightarrow C$, the utilization of link $B \rightarrow E$ and $E \rightarrow C$ will become 100% and 110% after the failure of $B \rightarrow C$. Hence, link $E \rightarrow C$ will suffer from traffic loss because its utilization is greater than 100%.

Moreover, GBP incorporates an enhanced failure-protection mechanism which allows it to minimize the probability that any logical link will become over-utilized after single logical link failures. This enhanced failure-protection mechanism has two objectives; the first one is the reduction of the traffic that is diverted from the failed logical link onto its backup path. The rationale behind this objective is to avoid the logical links of the backup path of the failed logical link to become over-

utilized due to the traffic diversion. The second objective is the increase of the spare capacity of the backup paths because this will allow the backup paths to accommodate diverted traffic without becoming over-utilized.

In order to support this GBP enhanced failure-protection mechanism, the head router of the failed logical link needs to broadcast a failure notification to all routers in the network when the logical link fails. Upon receipt of the failure notification, routers will check and *deactivate* any of their activated backup paths which use any IRLLL logical links of the failed logical link. This is done to achieve the two objectives of the GBP enhanced failure-protection mechanism. The deactivation of affected backup paths is done by diverting the traffic on them back onto their protected logical link and as a result, one or more physical links contained by that protected logical link may need to wake up to carry the reverted-back traffic.

In order to illustrate the GBP enhanced failure-protection mechanism, the topology in Fig. 5a is extended into Fig. 5b with the same traffic demands still being used and all logical links having capacity of 100Mbps. In this topology, it can be seen that the logical link $B \rightarrow D$ has part of its traffic diverted onto its backup path $BP2$ (which uses the failed logical link $B \rightarrow C$) when GBP is active so that additional physical links can go to sleep in $B \rightarrow D$. Therefore, when $B \rightarrow C$ fails, the traffic to be diverted is greater compared to the scenario where GBP is not active since $B \rightarrow C$ is carrying traffic from another protected logical link through the activated $BP2$. In this case, the GBP enhanced failure-protection mechanism will deactivate $BP2$ so as to reduce the traffic to be diverted due to the failure of $B \rightarrow C$. For example, if $BP2$ is diverting 5Mbps on $B \rightarrow C$ from $B \rightarrow D$, the total traffic diverted by $B \rightarrow C$ onto $BP1$ when it fails will reduce from 60Mbps to 55Mbps due to the deactivation of $BP2$. In order to enable the deactivation of $BP2$, it is necessary to deactivate all backup paths that are originally using the protected logical link of $BP2$, i.e. $B \rightarrow D$. This is done so that $B \rightarrow D$ has enough spare capacity to accommodate the increased traffic due to the deactivation of $BP2$. As mentioned in Section III.B.2, $B \rightarrow D$ is part of the IRLLL of $B \rightarrow C$ and according to the GBP enhanced failure-protection mechanism; any backup paths which use a link in the IRLLL of a failed link need to be deactivated.

Moreover, it may happen that the backup path of a failed logical link has reduced spare capacity because its logical links are part of the backup paths of other logical links. For example in Fig. 5b, if the logical link $E \rightarrow F$ has part of its traffic diverted onto its backup path $BP3$ for the energy saving operations of GBP, this traffic diversion by GBP will reduce the spare capacity of backup path $BP1$ of the failed logical link $B \rightarrow C$. Therefore, $BP1$ may become congested when $B \rightarrow C$ fails. In order to alleviate this problem, it is necessary to deactivate any backup paths that are using any logical links of the backup path of the failed link. For example, if $BP3$ was initially sending 5Mbps and is deactivated when $B \rightarrow C$ fails, then the traffic on $B \rightarrow E$ will fall by 5Mbps to 55Mbps and will have 45Mbps of spare capacity. Since the traffic on $B \rightarrow C$ is now reduced to 45Mbps due to the

previous deactivation of $BP2$, $BP1$ can now support all the traffic diverted by $B \rightarrow C$ when it fails.

When a logical link fails, the two failure protection mechanisms operate at the same time and independently from one another. If GBP has not previously activated any backup paths which use the IRLLL links of the failed logical link, only the conventional failure-protection mechanism will have an effect on the traffic distribution in the network

IV. GBP PERFORMANCE EVALUATION

In this section, the results of the evaluation of the proposed GBP scheme are presented and discussed. The performance evaluation is done using two operational academic network topologies, as described in Section IV.A. More specifically, the following network parameters are measured and discussed: 1) power and energy consumption; 2) MLLU; 3) increase in maximum packet delay; and 4) effect of single logical link failures on the post-failure peak MLLU and energy savings.

A. Network Scenarios

GBP was evaluated by using two academic network topologies, namely GÉANT and Abilene, and their real traffic matrices [16].

The GÉANT topology, summarized in Table III, consists of 23 Points-of-Presence (PoPs) and 74 unidirectional links with different capacities. In Table III, $|L|$ represents the number of logical links of a specific capacity c that have q number of physical links that individually transmit at λ optical carrier speed and consume p amount of power. The power consumption for each physical link was obtained from the maximum power consumption of Cisco line cards [17]. For a physical link of capacity of OC-48, a one-port line card uses 140W and for OC-3, there is no one-port line card but rather a four-port line card with total power consumption of 196W. For simplicity, an OC-3 physical link is assumed to consume $196/4 = 49W$. The Abilene topology consists of 12 PoPs and 30 unidirectional links of varying capacity, as shown in Table IV (which have the same notation as in Table III). During the evaluation of GBP with the two network scenarios, only the backup paths which did not have a delay greater than 25ms compared to their protected link were considered as eligible for GBP operations to ensure that traffic diversion does not lead to excessive increase in delay. Certainly this threshold can be flexibly configured by operators according to their own policies in practice.

For the traffic demands in the GÉANT and Abilene network, 480 consecutive traffic matrices that were measured at 15-minute intervals were considered [16]. Consistently, the 15-minute interval was also adopted as the period of the GBP optimization cycle. That is, the application of each traffic matrix on the network corresponds to the starting point in time of a new optimization cycle of GBP.

B. Power and Energy Saving Gains

The power saved by GBP was calculated by using Eq. (8) below. In order to evaluate the power saving gains of GBP, it

TABLE III
POWER MODEL OF GÉANT NETWORK TOPOLOGY.

Logical links				Physical links			
$ L $	c (Mbps)	q	$ L \times q$	z (Mbps)	λ	p (W)	$ L \times q \times p$ (W)
32	9953	4	128	2488	OC-48	140	17920
2	4876	2	4	2488	OC-48	140	560
32	2488	1	32	2488	OC-48	140	4480
8	155.2	1	8	155.2	OC-3	49	392
Σ	74		172				23352

TABLE IV
POWER MODEL OF ABILENE NETWORK TOPOLOGY.

Logical links				Physical links			
$ L $	c (Mbps)	q	$ L \times q$	z (Mbps)	λ	p (W)	$ L \times q \times p$ (W)
28	9920	4	112	2480	OC-48	140	15680
2	2480	1	2	2480	OC-48	140	280
Σ	30		114				15960

was compared with a Theoretical Upper Bound (TUB) scheme. TUB was obtained with IBM CPLEX [18] by adding the concept of link sleeping to the conventional non-integer Multi-Commodity Flow problem. Therefore, the restriction of using only the predefined protected links and backup paths to route traffic demands is not applied in TUB. GBP was simulated with different values of α for both GÉANT and Abilene.

$$\text{Power Saved} = \frac{\sum_{l=1}^{|L|} y_l \times p_l}{\sum_{l=1}^{|L|} q_l \times p_l} \times 100\% \quad (8)$$

Fig. 6 and 7 show that for all traffic matrices, GBP was able to save a significant amount of power for both simulated networks. Understandably, there is a gap between GBP and TUB in terms of power saving performance because GBP uses a single path for each SD flow while TUB uses a large number of paths. Of course, it should be noted that the path configuration given by TUB cannot be implemented in practice because of the large number of paths between each SD pair in the network that would be required.

Fig. 6 shows that, for the GÉANT scenario, power savings do not change significantly when α is reduced from 90 to 50. We also calculate the total energy saved by GBP as a proportion of the theoretical optimal energy that could be saved, Ψ , by using Eq. (9). The value of Ψ does not change much when α is reduced in Table V, which correlates with the observation made for Fig. 6. The curves for α equal to 80, 70 and 60 % are not shown in Fig. 6 for clarity of the figure but the curves follow a similar path as the curves for α equal to 90 and 50 %. The observations made in Fig. 6 and Table V can be explained by the fact that even when α is high, GBP does not have a high degree of freedom to divert a significant amount of traffic onto a backup path because this will make the logical links constituting that backup path to consume a larger amount of energy. That is, the spare capacity of most backup paths in the network remains mostly constant when α is varied from 90 to 50 because the first term in the spare capacity equation, Eq. (7), is the dominating one for most logical links in the

Fig. 6. Power saved for the GÉANT topology.

Fig. 7. Power saved for the Abilene topology.

TABLE V
TOTAL ENERGY SAVED, Ψ .

GÉANT		Abilene	
α	Ψ (%)	α	Ψ (%)
90	86.6	50	89.2
80	86.5	40	89.1
70	86.2	30	88.8
60	86.2	20	88.1
50	86.4		

network.

$$\Psi = \frac{\text{Energy saved by GBP}}{\text{Theoretical optimal energy saved}} \times 100\% \quad (9)$$

Fig. 7 and Table V show that the power and energy saved for the Abilene network also does not change much with the variation of α . It is interesting to see from the performance curves in Fig. 7 that GBP reacts to sudden changes in traffic conditions as TUB does even though the paths available to “absorb” these changes are limited for GBP.

C. Maximum Logical Link Utilization

The dynamicity of the Maximum Logical Link Utilization (MLLU) resulting from the GBP operations across the evaluation period is shown in Fig. 8 and 9 for GÉANT and Abilene respectively. The original MLLU values that were measured in the network are also included in the figures. For GÉANT, GBP was able to offload highly-utilized logical links and the maximum MLLU became close to the value of α when this

Fig. 8. Variation of MLLU for Original and GBP for GÉANT topology.

Fig. 9. Variation of MLLU for Original and GBP for Abilene topology.

was varied from 90 to 70. This shows that GBP is able to successfully enforce the constraint in Eq. (5) by offloading over-utilized logical links while not overloading under-utilized ones as a result of its operations.

When α is further reduced from 70 to 50, GBP is unable to reduce the peak MLLU to meet the value of α because there are not enough logical links with sufficiently large spare capacity to carry the traffic from the logical links whose utilization is above α . During its operations, GBP does not divert traffic on any logical link whose utilization is greater than α and therefore, it does not make the peak MLLU becomes worse compared to when GBP is not operated. Moreover, GBP will not divert an excessive amount of traffic to any logical links because this may result in the utilization of the logical links to go above α . GBP can enforce these two restrictions on the diversion of traffic by calculating the spare capacity of a logical link through Eq. (7).

Fig. 9 shows that the MLLU experiences substantial and frequent fluctuations during the original operation of Abilene where GBP was not activated. This is also reflected in the change in MLLU during the operation of GBP. The peak MLLU remains the same as the original one when α is varied from 50 to 20 during the operation of GBP. This happens when there is not enough space capacity on the backup paths to accept diverted traffic because they are already carrying a high volume of traffic. Hence, GBP cannot reduce the peak MLLU for Abilene as it did for GÉANT due to a lack of spare capacity in the Abilene network. For some traffic matrices in Fig. 8 and 9, it can be observed that the MLLU values of GBP can go up when compared to the original ones because

TABLE VI
INCREASE IN MAXIMUM PACKET DELAY (MS).

GÉANT				Abilene			
α	Max.	Avg.	Min.	α	Max.	Avg.	Min.
90	17.5	5.72	0	50	6.96	1.44	0
80	20.8	6.17	0	40	6.96	1.58	0
70	13.8	6.38	0	30	6.96	1.53	0
60	13.8	6.39	0	20	6.96	1.52	0
50	21.9	5.26	0				

GBP wants to concentrate traffic on the minimum number of logical links possible so as to save the maximum amount of power. This concentration of traffic on a minimum number of logical links is always done with the constraint that the MLLU should not go above the value of α because of the rerouting actions of GBP. For the traffic matrices for which the MLLU is above α for GBP, the original MLLU is also above α even though GBP is not being operated. This shows that GBP is not responsible for breaking the constraint α , but it is just the original high volume of traffic that causes the MLLU to be above α .

D. Increase in Maximum Packet Delay

Table VI shows the increase in the maximum packet delay when GBP is operated for GÉANT and Abilene. For the GÉANT network topology, the average increase in the maximum packet delay is small, at most 6.39ms. For the maximum increase in the maximum packet delay, the increase was not higher than 21.9ms. For the Abilene network topology, the average and maximum increase in maximum packet delay were quite small at around 1.58ms and 6.96ms respectively. There was no change in the observed minimum maximum packet delay for both network scenarios. Such extra delay introduced is also comparable to what is presented in the recent work based on hop-by-hop routing [19].

The main conclusion from these delay results is that it can be assumed that GBP does not significantly affect the packet delay in a well-connected network, which is the case of GÉANT and Abilene. This is despite the fact that GBP reroutes some traffic on longer backup paths to offload logical links.

E. Number of Concurrent TH links Offloading and Running Time Analysis

In terms of the number of links concurrently undergoing offloading at the same time, GBP was able to achieve on average 6 and 4 concurrent links offloading for GÉANT and Abilene respectively. The value for GÉANT is higher because of its larger topology size with more path diversity and therefore, the probability of backup paths sharing the same links is lower. An additional observation from these numbers is that GBP can scale with network size, that is as the network size becomes larger, the number of current TH links will increase and compensate for the larger number of TH links that must be offloaded.

A non-optimized single thread version of GBP was run on a 2.13Ghz dual-core laptop with 8 GB memory and the

maximum, average and minimum time to complete one GBP cycle for GÉANT was 0.110, 0.081 and 0.058 s respectively and for Abilene, it was 0.094, 0.058 and 0.056 s respectively. In an operational network, a GBP cycle is expected to have a much lower running time due to the routers being able to concurrently offload TH links. The low running times show that GBP is a practical solution for even large network such as GÉANT in terms of running time.

F. Single Logical Link Failure Analysis

A whole logical link can fail due to the cut of all its bundled physical fiber links, e.g. due to earthquakes or other physical damages. Since GBP makes use of backup paths for energy savings in addition to their primary purpose of preventing traffic loss upon single logical link failures, it is important to investigate the effect of GBP on the post-failure peak MLLU [20] and the energy savings upon any single logical link failure. As mentioned in Section III.D, upon the failure of an active logical link, the failure protection mechanism will divert traffic from that failed logical link to its corresponding backup path. This failure protection mechanism of logical links is applied in an MPLS-enabled network regardless of whether GBP is in place or not. In addition, GBP has an enhanced failure-protection mechanism which enables the amount of traffic to be diverted from the failed logical link to be reduced and also for the spare capacity of the backup path of the failed link to be increased so that the backup path can support the traffic diverted from the failed logical link.

1) *Post-failure Peak MLLU*: Single logical link failures were simulated for all considered traffic matrices of both GÉANT and Abilene scenarios. The aim was to examine the effect of the dual use of the resources of the backup paths for both power savings and logical link failure protection. The effect of GBP on logical link failure protection was quantified by computing the peak MLLU after any logical link has failed in the network. Fig. 10 and 11 show the post-failure peak MLLU for normal energy-agnostic operation (i.e. GBP is not operated) and when GBP is active with different values of α for GÉANT and Abilene respectively. The post-failure MLLU curves for when α is 80, 70 and 60 % for GÉANT are not shown in Fig. 10 for clarity since those curves follow a similar path as the curves when α is 90 and 50 %. It is the same case for the curves when α is 40 and 30 % for Abilene in Fig. 11. It can be observed that during single logical link failures, there was no increase in post-failure peak MLLU when GBP is active in most cases. When the traffic matrices for which the post-failure peak MLLU value is higher than 100% were further analyzed in detail, it could be observed that the peak MLLU values is the same irrespective of whether GBP is active or not. This is an important observation since it is only when the peak MLLU is higher than 100% that traffic is actually lost in the network and therefore, the same peak MLLU suggests that the network suffers from the same degree of over-utilization for these traffic matrices for both the scenarios where GBP is activated or not.

For the peak MLLU values which are less than 100%, it can be observed in Fig. 10 that the post-failure peak MLLU

Fig. 10. Post-failure peak MLLU between no-GBP and GBP operations for the GÉANT topology.

Fig. 11. Post-failure peak MLLU between no-GBP and GBP operations for the Abilene topology.

can sometimes be reduced when GBP is active for GÉANT. This can happen because GBP can reduce the MLLU in the network as shown in Fig. 8 for GÉANT. Therefore, there is a probability that the post-failure peak MLLU will be lower when GBP is activated because there is more spare capacity on the backup paths to accommodate the diverted traffic during single logical link failures. For Abilene in Fig. 11, there are a few traffic matrices where the post-failure peak MLLU is greater under GBP when the peak MLLU values are less than 100%. This is because GBP concentrates traffic on the lowest number of logical links as possible so that the physical links in the other logical links can go to sleep. Therefore, there may not be enough spare capacity on some protected logical links in the network when the GBP enhanced failure-protection mechanism diverts traffic to them, when it deactivates their respective backup paths because these backup paths use any IRL links of the failed logical link.

The main overall conclusion from Fig. 10 and 11 is that backup paths can be used for greater energy savings without reducing the ability of the backup paths to prevent traffic loss during single logical link failures. Therefore, there is no need to provision additional paths which are dedicated for energy savings because non-conflicting use of the backup paths for energy savings and prevention of traffic loss during single link failures is feasible.

2) *Impact on Energy Saving Gains during Single Logical Link Failures*: The energy saving gains by GBP can be affected by three different factors when a logical link fails

in the network. The first factor is that a failed logical link will not consume any energy, i.e. all the physical links of the failed logical links are considered to be “sleeping” and not consuming energy.

On the other hand, it may be necessary for logical links involved in the activated backup path of a failed logical link to wake up additional physical links in order to provide the extra spare capacity required to accommodate the diverted traffic from the failed logical link without causing any post-failure traffic congestion. This is the second factor which can affect the energy saving gains. In addition, some already activated backup paths (for energy savings and/or reduction of utilization at their protected logical links) need to be deactivated so as to reduce the amount of traffic to be diverted from the failed logical links, and also to increase the spare capacity of the backup path of the failed logical links. Therefore, the third factor is the possible reduction in the energy consumption of the logical links involved in the deactivated backup paths while the protected logical links of the deactivated paths may consume more energy because of the increased traffic on them. In GBP, the head routers are responsible for determining how many sleeping physical links in each of their logical links should wake up, so that only the minimum number of physical links is active without causing any traffic congestion. This design choice is made so as to maximize the energy savings in the network but without compromising on post-failure traffic loss.

Table VII shows the change in total energy saved during single logical link failures compared to a failure-free scenario for GÉANT and Abilene. In this table, a positive number means logical link failures increase the energy saved in the network compared to a failure-free scenario while a negative number means the opposite. On average, the energy saved decreases during single logical link failures. This is because several physical links will have to wake up in the logical links, constituting the backup path of the failed logical link, if there is not enough spare capacity in the currently active physical links of the backup path. The energy consumption of the protected logical links of the deactivated backup paths also increases due to increased traffic on them. However, the maximum reduction in energy saved during single logical link failures is not significant.

Interestingly, the Table VII indicates some unexpected observations where the energy saving gains can further increase upon single link failures. This can be explained by the fact that logical link failures are similar to putting links to sleep but with the key difference that the MLLU constraint in Eq. (5) needs not be respected during single logical link failures. In other words, GBP cannot always achieve the same level of energy savings compared to when some single logical links fail because GBP would break the MLLU constraint defined in Eq. (5). However, single logical link failures do not have this restriction because during these events, logical links can be loaded with as much traffic as the failure protection mechanism can handle.

TABLE VII
CHANGE IN TOTAL ENERGY SAVED, $\Delta\Psi$.

GÉANT				Abilene			
α	Max.	Avg.	Min.	α	Max.	Avg.	Min.
90	-0.0000847	-3.91	-12.0	50	0.0970	-1.37	-2.61
80	-0.0752	-3.89	-11.9	40	0.153	-1.33	-2.65
70	-0.345	-3.59	-8.20	30	0.476	-1.19	-2.70
60	-0.344	-3.58	-8.20	20	0.987	-0.834	-3.13
50	0.0365	-3.61	-8.66				

V. RELATED WORK

The research area of Energy-aware Traffic Engineering (ETE) was pioneered by [21] in 2003. Henceforth, there have been a number of schemes to tackle energy saving techniques in the context of backbone networks [22], [23]. ETE schemes can be classified broadly as either offline or online solutions. Offline schemes rely on historical traffic matrices to pre-compute a (long-term) network configuration, which will be applied during the live operation of the network. Examples of offline ETE schemes were proposed in [7], [10], [24], [25], [26], [27], [19], [28]. The search for the best solution can be enabled by some global optimization scheme. This is only possible because a holistic view on the global network conditions is available. It is also worth mentioning that such a strategy is suitable for the network scenarios where traffic patterns are relatively regular. Online ETE schemes make on the fly decisions according to the obtained information about the state of the network during the live operation of the network. Examples of these schemes were proposed in [8], [9], [21], [29], [30], [31]. While online ETE schemes are more efficient in reacting to unexpected traffic behaviors, a major technical hurdle for these schemes is how to efficiently coordinate between different decision-making entities in order to avoid conflicting decisions that may lead to severe and unpredictable consequences on the network performance.

In [8], the authors proposed a complex online ETE mechanism to dynamically distribute traffic load to multiple paths for each SD pair. This way, a subset of the links can have the opportunity to operate in a lower transmission rate and, hence, consume less energy. However, the proposed scheme is highly complex due to synchronization operations with the goal of avoiding conflicting decisions between decision-making routers. Moreover, it uses multi-path routing for each SD pair. The same authors later proposed a simpler ETE scheme in [9]. This simpler scheme uses historical traffic matrices to identify different sets of paths that are suitable for different traffic patterns in the network. Their scheme introduces high complexity in the processing of traffic matrices. In addition, the scheme may perform badly if infrequent traffic behaviors have not been accounted for during the processing of traffic matrices.

As far as fully-distributed ETE is concerned, in [31] the authors proposed that each router independently takes a decision on which of their directly attached links to put to sleep. The authors, however, did not explicitly address how to deal with conflicts in forwarding behaviors due to frequent recalculation of the forwarding tables of routers according to the changed

topology of the network. Moreover, the decisions made by routers are not evaluated before being implemented in the network. Therefore, there may be short periods of time where there is traffic loss in the network because it will take some time for the routers to actually get feedback from the network that the decisions made are causing traffic congestion. The fully-distributed ETE scheme in [32] only considers putting whole routers to sleep and not individual links.

A number of ETE schemes which use bundle links have been proposed in [10], [11], [14]. In these ETE schemes, decisions are made by either a centralized controller or only one decision entity at any one time irrespective of the size of the network. In contrast, GBP is a fully-distributed scheme where the number of routers which can concurrently make decisions increases as the size of the network grows. This is because there is a smaller chance of conflict in a large network where the backup paths are less likely to have logical links in common.

VI. CONCLUSIONS

The energy consumption of backbone networks is rapidly rising because of the exponential increase of the number of network devices. Such problem is aggravated by the fact that improvements on energy efficiency of the network components have not been enough to offset the increased energy consumption. In this paper we proposed a new Energy-Aware Traffic Engineering (ETE) scheme called Green Backup Paths (GBP).

GBP is able to leverage existing MPLS backup paths that are made available for protecting the network in case of link failures. GBP is a fully distributed online ETE algorithm which applies intelligent rerouting of traffic in order to allow bundled physical links within logical links to go to sleep by lowering the amount of traffic being routed on those logical links.

GBP was validated through simulations of two academic networks, namely GÉANT and Abilene, and their respective real traffic matrices. Results showed that GBP achieves significant energy savings while also decreasing the Maximum Link Utilization in the network. Energy savings achieved by GBP are within 15% of the Theoretical Upper Bound (TUB), which indicates the efficiency of the proposed algorithm. GBP only considers traffic reroute for energy saving in failure-free network conditions.

In the future, alternative ways of calculating the backup paths will be investigated to examine if the performance of GBP can be improved further. In addition, a limited number of additional paths can be added to the network to investigate if this can significantly improve the performance of GBP while keeping the processing requirement on the routers still at a low level.

REFERENCES

- [1] J. G. Koomey, "Estimating total power consumption by servers in the U.S. and the world," Tech. Rep., Lawrence Berkley National Lab., 2007.
- [2] R. Bolla, F. Davoli, R. Bruschi, K. Christensen, F. Cucchietti, and S. Singh, "The potential impact of green technologies in next-generation wireline networks: Is there room for energy saving optimization?" *IEEE Com. Mag.*, vol. 49, no. 8, pp. 80–86, 2011.
- [3] C. Lange, D. Kosiankowski, R. Weidmann, and A. Gladisch, "Energy consumption of telecommunication networks and related improvement options," vol. 17, no. 2, March 2011, pp. 285–295.
- [4] A. Raj and O. C. Ibe, "A survey of ip and multiprotocol label switching fast reroute schemes," *Elsevier Computer Networks*, vol. 51, no. 8, pp. 1882–1907, 2007.
- [5] Q. Li, M. Xu, J. Wu, P. Lee, X. Shi, D.-M. Chiu, and Y. Yang, "A unified approach to routing protection in IP networks," *IEEE Transactions on Network and Service Management*, vol. 9, no. 3, pp. 306–319, 2012.
- [6] T. Cicic, A. Hansen, A. Kvalbein, M. Hartmann, R. Martin, M. Menth, S. Gjessing, and O. Lysne, "Relaxed multiple routing configurations: IP fast reroute for single and correlated failures," *IEEE Transactions on Network and Service Management*, vol. 6, no. 1, pp. 1–14, 2009.
- [7] M. Zhang, C. Yi, B. Liu, and B. Zhang, "GreenTE: Power-Aware Traffic Engineering," in *Proc. of IEEE ICNP*, 2010.
- [8] N. Vasic and D. Kostic, "Energy-aware traffic engineering," in *Proc. of International Conference on Energy-Efficient Computing and Networking*, ser. e-Energy, 2010, pp. 169–178.
- [9] N. Vasić, P. Bhurat, D. Novaković, M. Canini, S. Shekhar, and D. Kostić, "Identifying and using energy-critical paths," in *Proc. of 7th CoNEXT*, 2011, pp. 18:1–18:12.
- [10] W. Fisher, M. Suchara, and J. Rexford, "Greening backbone networks: reducing energy consumption by shutting off cables in bundled links," in *Proc. of the ACM SIGCOMM workshop on Green networking*, 2010.
- [11] B. Mumey, J. Tang, , and S. Hashimoto, "Enabling green networking with a power down approach," in *Proc. of the IEEE International Conference on Communications*, ser. ICC, 2012, pp. 2867–2871.
- [12] C. Gunaratne, K. Christensen, B. Nordman, and S. Suen, "Reducing the energy consumption of ethernet with adaptive link rate (alr)," *IEEE Transactions on Computers*, no. 57, pp. 448–461, 2008.
- [13] H. Alnuweiri, L. Wong, and T. Al-Khasib, "Performance of new link state advertisement mechanisms in routing protocols with traffic engineering extensions," *IEEE Com. Mag.*, vol. 42, no. 5, pp. 151–62, 2004.
- [14] M. Charalambides, D. Tuncer, L. Mamatas, and G. Pavlou, "Energy-Aware Adaptive Network Resource Management," in *Proc. of the IFIP/IEEE Integrated Management Symposium*, ser. IM, 2013.
- [15] K. Kompella and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)," RFC 4203, 2005.
- [16] TOTEM Project, "The totem traffic engineering toolbox," <http://totem.run.montefiore.ulg.ac.be>, online. Accessed Jan. 2014.
- [17] Cisco Systems Inc., "Cisco 12000 Series Packet Over SONET/SDH (POS) Line Cards," http://www.cisco.com/en/US/prod/collateral/routers/ps6342/product_data_sheet0900aecd803fd7b9.pdf, 2006, online. Accessed Jan. 2014.
- [18] IBM, "IBM ILOG CPLEX Optimizer," <http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/index.html>, online. Accessed Jan. 2014.
- [19] Y. Yang, D. Wang, M. Xu, and S. Li, "Hop-by-hop Computing for Green Internet Routing," in *Proc. of IEEE ICNP*, 2013.
- [20] K.-H. Ho, N. Wang, G. Pavlou, and C. Botsiaris, "Optimizing post-failure network performance for ip fast reroute using tunnels," in *Proc. of 5th QShine*, 2008, pp. 44:1–44:7.
- [21] M. Gupta and S. Singh, "Greening of the internet," in *Proc. of SIGCOMM*, ser. SIGCOMM, 2003, pp. 19–26.
- [22] A. Bianzino, C. Chaudet, D. Rossi, and J. Rougier, "A survey of green networking research," *IEEE Commun. Surveys Tutorials*, no. 99, pp. 1–18, 2010.
- [23] 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 Commun. Surveys Tutorials*, no. 99, pp. 1–22, 2010.
- [24] L. Chiaraviglio, M. Mellia, and F. Neri, "Energy-aware backbone networks: A case study," in *Proc. of the IEEE International Conference on Communications*, ser. ICC, 2009, pp. 1–5.
- [25] L. Chiaraviglio, M. Mellia, and F. Neri, "Reducing power consumption in backbone networks," in *Proc. of the IEEE International Conference on Communications*, ser. ICC, 2009, pp. 2298–2303.
- [26] A. Cianfrani, V. Eramo, M. Listanti, M. Polverini, and A. Vasilakos, "An ospf-integrated routing strategy for qos-aware energy saving in ip backbone networks," *IEEE Trans. Netw. Service Manag.*, no. 99, pp. 1–14, 2012.
- [27] F. Francois, N. Wang, K. Moessner, and S. Georgoulas, "Optimizing Link Sleeping Reconfigurations in ISP Networks with Off-Peak Time Failure Protection," *IEEE Transactions on Network and Service Management*, vol. 10, no. 2, pp. 176–188, 2008.
- [28] B. Addis, A. Capone, G. Carello, L. Gianoli, and B. Sanso, "A robust optimization approach for energy-aware routing in mpls networks," in *Proc. of ICNC*, Jan 2013, pp. 567–572.

- [29] M. Shen, H. Liu, K. Xu, N. Wang, and Y. Zhong, "Routing on demand: toward the energy-aware traffic engineering with OSPF," in *Proc. of IFIP Conference on Networking*, ser. Networking, 2012, pp. 232–246.
- [30] C. Panarello, A. Lombardo, G. Schembra, L. Chiaraviglio, and M. Mellia, "Energy saving and network performance: a trade-off approach," in *Proc. of International Conference on Energy-Efficient Computing and Networking*, ser. e-Energy, 2010, pp. 41–50.
- [31] A. Bianzino, L. Chiaraviglio, and M. Mellia, "Grida: A green distributed algorithm for backbone networks," in *Proc. of IEEE GreenCom*, ser. GreenCom, 2011, pp. 113–119.
- [32] K. H. Ho and C. C. Cheung, "Green distributed routing protocol for sleep coordination in wired core networks," in *Proc. of ICNC*, 2010.

Frederic Francois is a research assistant at the University of Bristol, UK. He was awarded his Ph.D. on "Policy-driven Traffic Engineering in Energy-aware ISP Networks" at the University of Surrey in October 2013. He received his M.Eng. in Electronic and Communications Engineering at The University of Sheffield in 2010. His research interests include: traffic engineering, packet and optical network virtualization, software-defined networking, test-bed federation and network function virtualization.

Klaus Moessner is a Professor for Cognitive Networks and Services, in the Centre for Communication Systems Research at the University of Surrey, UK. Klaus earned his Dipl-Ing (FH) at the University of Applied Sciences in Offenburg, Germany, an MSc from Brunel University and PhD from the University of Surrey (UK). His research interests include dynamic spectrum allocation, cognitive networks, reconfiguration management, service platforms and adaptability of multimodal user interfaces.

Stylianos Georgoulas received his Diploma in Electrical and Computer Engineering from University of Patras in Greece and his Ph.D. degree in Computer Networks from University of Surrey, UK. He is currently a Research Fellow with the Centre for Communication Systems Research, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, U.K., where he has worked for several EU and EPSRC research projects. His research interests are in the areas of autonomic/dynamic service management using machine learning techniques, energy efficient traffic engineering and protocol design, and formal verification for stability control of protocols and networking processes.

Ning Wang is a Lecturer at the Centre for Communication Systems Research (CCSR), University of Surrey in UK. He received his B.Eng (Honours) degree from the Changchun University of Science and Technology, P.R. China in 1996, his M.Eng degree from Nanyang University, Singapore in 2000, and his Ph.D. degree from the University of Surrey in 2004 respectively. His research interests mainly include energy-efficient networks, network resource management, Quality of Service and resilience control mechanisms.

Ricardo de Oliveira Schmidt is a Ph.D. candidate at the Design and Analysis of Communication Systems Group (DACS) of the University of Twente, The Netherlands. He received a M.Sc. degree in Computer Science from the Federal University of Pernambuco (UFPE), Brazil, and a B.Sc. degree in Computer Science from the University of Passo Fundo (UPF), Brazil. His research interests include network management, traffic measurements, routing and addressing.