Energy Saving Through Traffic Profiling and Prediction in Self-Optimizing Optical Networks

Domenico Siracusa[1], **Federico Pederzolli**[1,2], **Renato Lo Cigno**[2], **Elio Salvadori**[1]
[1] *CREATE-NET, via alla cascata 56/D, 38123, Trento, Italy*
[2] *DISI - University of Trento, Via Sommarive 5, Trento, Italy*
Author e-mail address: domenico.siracusa@create-net.org

Abstract: A method that automatically learns and predicts the traffic behavior to save energy by adjusting the number of active optical carriers is presented. Simulations prove it provides large savings and ensures low traffic loss probability.

OCIS codes: 060.4250, 060.4256.

1. Introduction

Reducing the energy footprint of the Internet has long been a hot topic in network research, fueled by the economic interest of telecom operators, and recently also by social pressures and ecologic concerns. Literature on the topic, impossible to overview here for lack of space, can be grouped in three broad categories [1]: (*i*) re-engineering/re-design of components and architectures; (*ii*) dynamic adaptation/sleep of the hardware, (*iii*) energy aware routing/traffic grooming techniques for the RWA/RSA and network planning phases.

A broad and still little exploited avenue to improve the energy efficiency of modern networks is to exploit the daily fluctuations in carried traffic: available data confirms that such fluctuations exist also at high aggregation levels [2][3], and furthermore suggests that the daily patterns repeat regularly, in spite of changing habits and applications. Years of research on this subject still leave its potentials largely untapped due to the lack of complete, effective, and autonomic solutions. Authors of [4] propose to establish multiple lightpaths (LPs) for each logical connection in a legacy WDM network, dynamically setting them up or tearing them down to accommodate real traffic requests; the method is unable to predict future increases in traffic and uses a simple heuristic to decide when to change the number of active LPs to serve the current requests. Further methods [5][6] adopt Integer Linear Programming (ILP) techniques to compute the set of nodes, links or interfaces to turn off. These techniques require huge computational resources and time, and can be applied only offline. Moreover, they make use of predictive traffic matrices, without addressing the problem of producing them. Ref. [7] proposes a system to collect the necessary data, but makes use of logical topology reconfigurations that cause traffic losses while IP learns of the changes. According to the data available in [2][3], traffic patterns repeat day after day, but the specific behavior is LP-dependent, as it may present temporary anomalies, or have different trends in time, requiring short-term matrixes to achieve reasonable accuracy. For this reason, a system capable of learning the behavior of the traffic, predict its future behavior, and dynamically adjust the resources assigned to traffic relations may turn early research into usable solutions.

We propose an energy-saving method that *learns* a model of traffic variations for each active LP and builds a traffic profile by *automatically* collecting the necessary data, then exploits it by applying one of the proposed algorithms to *predict* how many carriers are unneeded and can be turned off to save energy. Therefore, this method for self-optimizing networks resorting to cognition can be catalogued as belonging to the dynamic adaptation category (no re-routing of traffic is applied). From a technological point of view, it can operate in a flexible optical network where LPs are served by multiple carriers (including WDM networks that serve some connections with multiple LPs), switching some of them off when the traffic demands are low. It is agnostic to the control plane architecture, which can be either distributed or centralized. We have performed simulations to demonstrate that a network operator adopting the proposed method can save substantial amounts of energy while keeping the probability of overloads below an arbitrary threshold.

2. Energy saving method leveraging on the traffic patterns

In this work we assume that (*i*) there exists a traffic pattern with periodic behavior of approximately one day offered to each LP; (*ii*) any LP lasts, on average, at least a few tens of days. The former assumption is supported by the evidence available on [2][3], which shows a strong similarity between traffic patterns belonging to different days, while the latter is rooted in the way optical networks are operated. Given these assumptions, we divide each day into a number of fixed-length time slots (TS), lasting, for example, 10 minutes, so that the transient needed to power up a transponder (usually less than 30 s, according to data sheets [8]) accounts only for a small fraction of the slot.

The steps of the proposed method are shown in Fig. 1(a): for each active LP and at each TS, an algorithm builds and updates the traffic profile associated to the LP, then another algorithm predicts the future traffic and adjusts it with a safety margin; based on this information, the carriers reserved for the LP are turned on or off to support the predicted amount of traffic with a given probability, by means of whatever control plane protocol the network employs (e.g. the GMPLS suite). In addition, a control algorithm is in charge of dynamically tuning the safety margin parameter in order to achieve the desired probability of mispredicting the real traffic.

First, the method builds an approximate model, or profile, of the average/peak traffic carried within each slot from a set of measurements of the actual loads historically offered to that LP. Here the idea is to implement on each monitoring agent of an LP an algorithm called profile-building (not shown here due to space constraints) that takes readings from this subsystem to produce an estimate of the offered load during the last time slot, which could be the average, or maximum of multiple averages (given that we are interested in supporting peak traffic), quantizes it into n classes (e.g., each of 1 Gbit/s), encapsulates this information into a message, together with the ID of the LP being monitored and that of the time slot the measurement belongs to, and sends it to the node responsible for managing the considered LP. There, each incoming traffic measure is added to a *Cumulative probability Distribution Function* (CDF), quantized over the set of n traffic classes. The value of n is static, i.e., it does not increase with the number of samples, which ensures that the upper-bound complexity of the prediction algorithms remains fixed.

Once the system has a number of measurements of the traffic carried during a given timeslot, the idea is to derive, for each TS i and LP k, a value C^k_i that represents the predicted traffic for the following days during the same TS. This value should be such that, denoting as x^k_i the real traffic over LP k during TS i of day j, the probability $P(x^k_i(j) < C^k_i)$ would be equal to a user-defined target value. In real operations, with a standard distribution fitting procedure, samples can be treated as drawn from a random variable. By employing maximum-likelihood or Bayesian techniques C^k_i can be estimated with the appropriate reliability. Unfortunately, we could not find in the literature real or measured distributions for x^k_i, thus we resort to heuristic approximations.

The simplest heuristic is the use of the traffic profile given a certain percentile, e.g., 95% or in general the R, to account for outliers that make bad predictors. We named this algorithm **Profiling**. In general profiling on a percentile should guarantee that the estimate is accurate with probability R, however the presence of outlier suggest to use a fairly small R and adjust the prediction with a *safety margin S*, expressed as a fraction of the LP nominal capacity, in order to absorb spikes or unpredicted variations. Of course, large values of S reduce savings, and since the normal fluctuations (not outliers that are statistically hard to predict) are small, small values are sufficient to guarantee good performance.

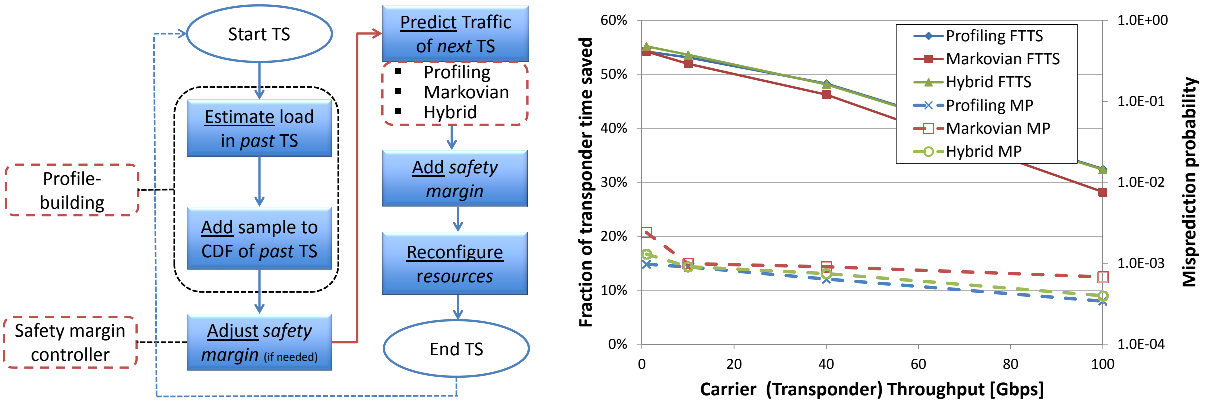
The solution discussed so far relies on a simple learning a model of the traffic, but with a long memory, and then use it to predict when to turn transponders on and off. Another approach, inspired by [4], is to simply consider the last traffic measurement for the LP, and, based on the assumption that the traffic is auto-correlated, just predict that future traffic will be the current one plus the safety margin S, to absorb possible increases. This algorithm, called **Markovian** because it has no memory, has the advantage of working even if no real pattern of traffic behavior exists underneath, as long as the differences between consequent time slots remain relatively low, which, based on the public traffic data available, is an applicable assumption on anything but the smallest, most underutilized LPs.

This simple technique is, however, inefficient during predictable downward gradients at night, were some experience would suggest that the traffic systematically decreases. To overcome this limitation, it is possible to use a variant of the profile-building algorithm that builds a CDF of the differences between subsequent TSs, and uses this information to adjust the predictions of the Markovian solution. We call this algorithm **Hybrid**.

As a further feature, we have introduced for all three algorithms a *controller* to dynamically adjust the value of S, to keep the probability of losing traffic around a certain target value. Among several possible strategies, we opted for one that immediately enlarges the margin after every misprediction (that implies a traffic loss) by the perceived difference between the requested load and the one that can be carried, then periodically diminishes it, tuning the period between reductions to obtain an a-posteriori desired misprediction probability. This controller is not an optimal solution, as more sophisticated ones may offer better convergence properties or larger savings, but more investigations on this subject are relegated to future works.

3. Performance evaluation

The proposed method has been implemented in an event-driven network simulator based on *omnet++*. Each day is divided into 144 10-minutes slots. In each slot a traffic measure, representing the traffic experienced by a given LP, is drawn from a Gaussian distribution with average equal to the value given by a trace based on [2][3], and a variance related to the LP nominal capacity (e.g., 10% of the average). A first order low pass filter introduces correlation between subsequent slots to emulate traffic continuity. We are well aware that there are better models in literature, but this is enough for our purposes, and simplifies the reproduction of the results. We consider LP



(a) (b)

Fig. 1. Energy-saving method block diagram (a); transponder uptime and misprediction probability against the throughput of a single carrier (b).

requests for {100,200,300,400} Gbit/s and carriers of {1,10,40,100} Gbit/s (reasonable values both for legacy transponders and OFDM sub-carriers). Two parameters are evaluated: (*i*) Fraction of Transponder Time Saved (FTTS), which is the average fraction of time that a transponder is powered down by the proposed method; (*ii*) Misprediction Probability (MP), which is the fraction of cases where the decision of the method result in being unable to carry all of the incoming traffic. For the experiments shown we have set the target misprediction probability of our controller to 10^{-3} and the low pass filter parameter α=0.6.

Fig. 1(b) shows that FTTS decreases faster than linearly for large carriers due to the larger amounts of unrequested active capacity that these carriers produce. This present a strong case in favor of using smaller sub-carriers, which allow for greater transmission ranges at the cost of spectral efficiency, but some savings can still be achieved while maximizing the latter. The absolute fractions of energy saved are related to the traffic model we employ, but show that exploiting the traffic correlation to build profiles (Profiling and Hybrid algorithms) does indeed improve the amount of savings achievable by the network w.r.t. the Markovian algorithm, which nonetheless appears to provide large savings. On the other hand, the Profiling algorithm ensures the lowest misprediction probability, while the Markovian provides the worst. However, all the proposed algorithms are close to our target. We have also tested (not shown here due to space constraints) that our safety margin controller is able to obtain an a-posteriori sample misprediction probability within the desired range in a reasonable time, i.e., less than 20 days with the Profiling and Hybrid algorithms. The proposed algorithms can be improved upon: if we could perfectly predict and fit our capacity to the incoming traffic we should be able, using our model and parameters, to achieve an average energy saving of about 82.5%. We achieve about 55%, which is far from optimal, but a significant improvement over the zero of current deployments.

4. Conclusions

We presented a method for self-optimizing backbone optical networks that exploits the daily traffic load variations to learn a traffic profile and to predict the occupation of the available resources to achieve energy savings, by means of shutting down unneeded carriers in a multi-carrier scenario. We showed, through simulations, that the algorithms proposed to predict the traffic can offer substantial savings without affecting the overall performance of the network.

Aknowledgment: this work is partially supported by CHRON FP7 European Project, grant agreement n. 258644.

References

[1] 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, 13, 223–244 (2011).

[2] AMS-IX public statistics, https://www.ams-ix.net/technical/statistics. Accessed: 06/2013.

[3] GARR public statistics, http://www.garr.it/a/rete/statistiche/46-backbone. Data only available in Italian. Accessed: 06/2013.

[4] A. Gencata and B. Mukherjee, “Virtual-topology adaptation for wdm mesh networks under dynamic traffic,” IEEE/ACM Transactions on Networking, 11, pp. 236-247, 2003.

[5] Y. Zhang, M. Tornatore, P. Chowdhury, and B. Mukherjee, “Energy optimization in ip-over-wdm networks,” Optical Switching and Networking 8, pp. 171-180, 2011.

[6] J. Lopez et al., “Traffic and power-aware protection scheme in elastic optical networks,” NETWORKS 2012.

[7] N. Fernandez et al. “Cognition to design energetically efficient and impairment aware virtual topologies for optical networks,” ONDM 2012.

[8] Pirelli data sheet, http://www.pirellibroadband.com/en_IT/browser/attachments/pdf/Pirelli_Tunable_DWDM_Transponder.pdf.