Virtual Network Topology Reconfiguration based on Big Data Analytics for Traffic Prediction

Fernando Morales, Marc Ruiz, and Luis Velasco*
Optical Communications Group (GCO), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
**e-mail:lvelasco@ac.upc.edu*

Abstract: Big data analytics is applied for IP traffic prediction. When the virtual topology needs to be reconfigured, predicted and current traffic matrices are used to find the optimal topology. Exhaustive simulation results reveal large benefits.

OCIS codes: (060.4250) Networks; (060.4256) Networks, network optimization

1. Introduction

In multilayer IP/MPLS transport networks, virtual network topologies (VNT) are created by connecting IP/MPLS routers through virtual links (*vlinks*) supported by connections in the optical layer (*lightpath*). Static IP network topologies are commonly designed to cope with the traffic forecast. However, the forecast traffic increment entails static VNTs being largely overprovisioned thus, increasing network total cost of ownership (TCO). In view of that, network operators are looking for more efficient architectures, while providing the required grade of service. To that end, VNTs should be dynamically adapted to follow traffic variations to provide the highest quality of service while efficiently using physical resources such as transponders.

To automate VNT adaptability, traffic needs to be monitored in the IP routers. One simple architecture consists in configuring some vlink capacity usage threshold (e.g., 90%) such that the network controller increases the capacity of vlinks by establishing a parallel lightpath when their capacity threshold is exceeded [1].

Notwithstanding the VNT adaptability provided by the threshold-based reconfiguration, the introduction of new types of service requiring large bitrate connectivity (e.g., datacenter interconnection, CDN for live-TV and video distribution, etc.) causes changes in the direction of the traffic along the day. Therefore, in addition to adapt the capacity of existing vlinks, the VNT should be changed by adding and releasing vlinks. Such VNT reconfiguration entails that, instead of monitoring vlink capacity usage, end-to-end traffic should be monitored so as to reconfigure the current VNT based on origin-destination (OD) traffic. That could be even refined by classifying flows into services enabling Deep Packet Inspection monitoring in edge routers.

Because its benefits, VNT reconfiguration has been widely studied in the literature. To follow traffic changes, authors in [2] proposed to add/remove one single lightpath each time the VNT is reconfigured. Another topic is using monitored data to produce estimations. In that regard, authors in [3] proposed a method for reducing errors in traffic estimations, while authors in [4] used estimated traffic to predict pre-defined scenarios.

VNT reconfiguration requires from powerful algorithms to analyze large amounts of traffic monitoring data, so as to anticipate, when possible, to traffic changes. In this paper, we propose using Big Data analytics for traffic prediction that runs periodically (e.g., every hour). In the case of the VNT needs to be reconfigured, predicted traffic is used as input of a VNT optimizer that finds the topology for the next period thus, implementing a decision making process based on the *observe-analyze-act* loop [5].

2. Data analytics -based VNT reconfiguration

As stated above, we assume that traffic monitoring data is collected at the edge IP routers at regular intervals, e.g., every minute. Every edge router collects a set of samples for the traffic to every other destination router, which is stored in a common data repository (Fig. 1). Following a predefined time period, e.g. every hour, a time series is retrieved for each OD pair and pre-processed applying data stream mining *sketches* to conveniently summarize collected data into modelled data representing the OD pair. Modelled data includes, among others, the minimum, maximum, average, and last collected bitrate within the hour.

The set of modelled variables for the current period t is stored in a repository together with variables belonging to previous periods. A prediction module based on machine learning techniques, generates the OD traffic matrix predicted for the next period that is used by a decision maker module to decide whether the current VNT needs to be reconfigured. Specifically, the predictor consists in an Artificial Neural Network (ANN) [6] model, which has been chosen due to its inherent capability of adapting to changes in a non-supervised manner, in contrast to auto-regressive models to fit continuous time series. Assuming an ANN with one hidden layer, the notation p:s:1 indicates an ANN with p inputs, s neurons in the hidden layer, and 1 output. Our model outputs the average traffic in time t for an OD pair given the last p average measures of that pair. A back-propagation algorithm is applied for training the ANN from large initial sets of modelled data from every OD pair.

In case of reconfiguration, the predicted OD traffic matrix and the current set of connections are used as

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement n° 317999 IDEALIST project and from the Spanish MINECO SYNERGY project (TEC2014-59995-R).

978-1-943580-07-1/16/$31.00 ©2016 Optical Society of America

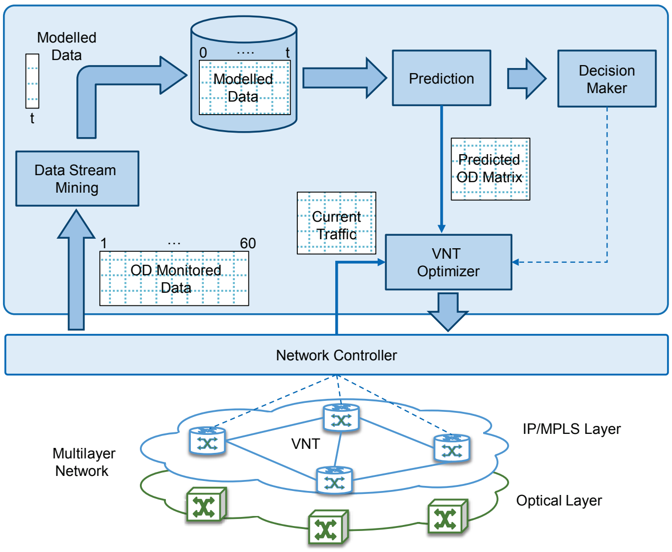

Fig. 1. Applying the Observe-Analyze-Act loop for VNT reconfiguration.

Table 1. VENTURE Heuristic Algorithm

```
INPUT: G_O(V,L), G_V(N,E'), P, D, OD, α, k
OUTPUT: E*, P*
 1:  deallocateDemands(G_V)
 2:  releaseLightpaths(G_O)
 3:  G_V*←∅; G_V*.cost←∞
 4:  OD ← phaseI(G_V,OD,α)
 5:  if OD==∅ then G_V* ← G_V
 6:  else
 7:     OD_aux ← OD; G_Vaux ← G_V
 8:     if |OD|≤k then maxIte==1
 9:     for i in 1..maxIte do
10:        OD ← OD_aux; G_V ← G_Vaux
11:        OD ← phaseII(OD,G_V,k)
12:        if OD≠∅ then
13:           OD ← phaseIII(OD,G_V)
14:           if OD≠∅ then continue
15:        if G_V.cost <= cost(G_V*) then
16:           G_V* ← G_V
17:  if G_V*==∅ then return infeasible
18:  deallocateDemands(G_V*)
19:  routeCurrentDemands(G_V*,D)
20:  return G_V*
```

input of an optimization algorithm that finds a target VNT. Since the threshold-based VNT reconfiguration is able to increase the capacity of existing vlinks, we focused on detecting changes in the traffic direction that might cause changes in the VNT (i.e., adding or releasing vlinks). Thus, we inspect predicted ODs individually to find when the estimated bitrate of those ODs without a direct vlink in the current VNT exceeds a given threshold; in such case, the VNT reconfiguration is triggered.

The VNT reconfiguration based on traffic prediction problem (VENTURE) can then be stated as follows:

Given:

- An optical network represented by graph $G_O(N, L)$, being N the set of OXC nodes and L the set of fiber links.
- The current VNT represented by graph $G_V(V, E')$, being $V \subseteq N$ the subset of IP/MPLS routers and $E' \subseteq E$ the subset of current vlinks. Set E is the set of all possible vlinks among the IP/MPLS nodes.
- The set P of current lightpaths supporting E'.
- The set and the capacity $Q(v)$ of available transponders for each node v in V.
- The set D of demands currently set-up, specifying source, destination, and requested bitrate.
- The predicted OD traffic matrix specifying the predicted bitrate.
- The maximum number k of new vlinks to be added to the VNT.

Find: a reconfigured VNT $G_V*(V, E*)$, where $E* \subseteq E$, and a set $P* \supseteq P$ of lightpaths supporting $E*$.

Objective: Minimize the overall resource utilization to serve both, D and OD sets.

Table 1 presents the proposed heuristic algorithm for the VENTURE problem. After releasing the current demands and lightpaths supporting the current VNT (lines 1-2), the heuristic goes for three sequential phases. Each phase allocates a subset of OD pairs and modifies the topology and the capacity of the VNT. Phase I (line 4) is a deterministic procedure where large OD pairs that can be routed through direct vlinks are processed. If not all OD pairs were processed, Phase II and Phase III are executed for a given number of iterations in the hope of finding the minimum cost VNT (lines 9-16). In Phase II (line 11), a set of at most k new vlinks to serve a OD pair in one single hop are created. Phase III (line 13) extends the connectivity of current VNT before solving the minimum cost routing for each remaining OD. If a feasible VNT is obtained, the current set of demands is allocated back to this new VNT (lines 18-19). Recall that ODs used for the new VNT might be different than current demands. Thus, a randomized algorithm is applied to allocate current demands fitting OD pairs.

3. Results

For evaluation purposes, we implemented a simulator in OMNeT++ containing the modules described in Fig. 1. As stated in the introduction, traffic changes can be as a result of changes in the volume and the direction. To measure the effect of those changes, we implemented randomized generators that inject traffic following pre-defined traffic profiles (see average daily evolution in Fig. 2). Two experiments were carried out on the 30-node Telefonica's national network: first, the traffic between any OD pair was generated as a mix of all three traffic profiles and second, the set nodes was divided into two subsets to generate changes in traffic direction; one subset mixed *Business* and *CDN* profiles, while the other used the *DC2DC* profile only. A limited number of 100Gb/s transponders were installed.

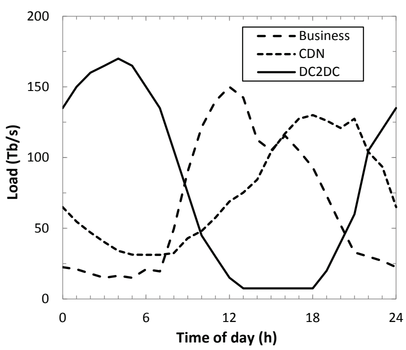

Fig. 2. Considered traffic profiles.

With such configuration, we firstly generated modelled data to train the ANN. By configuring p=24 and s=5,

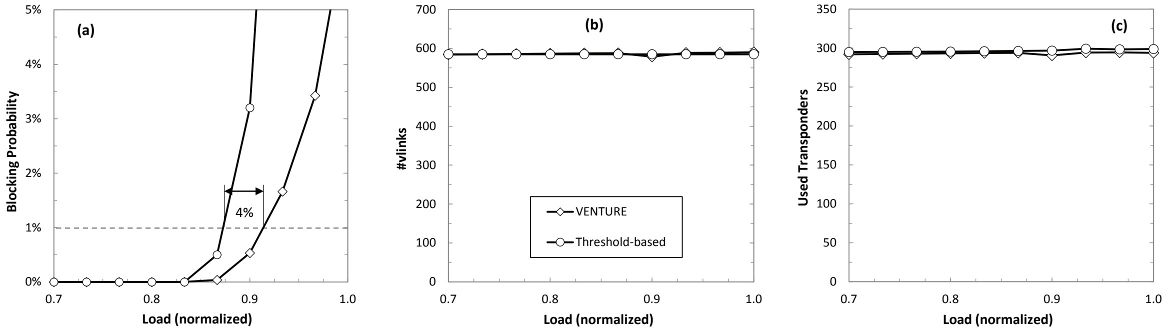

Fig. 3. Blocking (a), amount of vlinks (b), and maximum transponders usage (c) vs. load for the non-directional traffic variation.

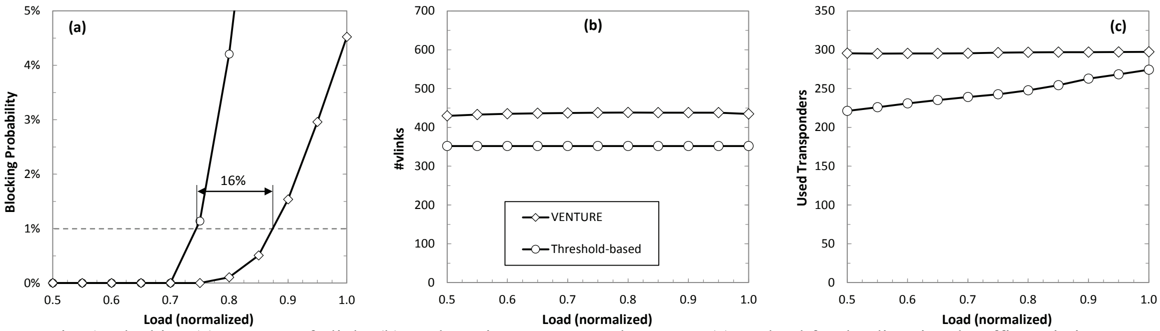

Fig. 4. Blocking (a), amount of vlinks (b), and maximum transponders usage (c) vs. load for the directional traffic variation.

we obtained predictions with a relative error of ±5% for more than 95% of cases for every OD pair.

Fig. 3 presents the obtained results for the non-directional traffic variations scenario, where the load has been normalized to the highest value with a probability close to 5%. Since the threshold-based scheme does not add/release vlinks, it is very dependent of the actual VNT configured. For that very reason and for the sake of a fair comparison, the VNT topology was selected from one of those found by the VENTURE scheme. In Fig. 3a, the daily average blocking probability against the load is presented. The plot for the threshold-based shows how that scheme is able to increase the capacity of the vlinks until no more transponders are available. Notwithstanding the performance of the VENTURE scheme show a gradual blocking increment, little improvement is achieved. Fig. 3b plots the average cardinality of the VNT (number of unidirectional vlinks) as a function of the load, where the cardinality of the VNTs implemented by the VENTURE scheme is constant and similar to that of the threshold-based scheme. Fig. 3c plots the maximum number of transponders used in the network, which is also constant and close to the maximum number of transponders available under both approaches. This is as a consequence of the daily traffic variation where traffic peaks follow traffic flows.

Fig. 4 focuses on the directional traffic variation scenario. In this case, the benefits of the VENTURE scheme are clear as a result of the variation in the direction of the traffic between night and day. In Fig. 4a, gain around 16% in terms of traffic for 1% of daily blocking probability is observed. Regarding vlinks, the cardinality of the VNTs that is implemented by the VENTURE scheme is 24% larger than those of the threshold-based. This is as a result of the threshold-based scheme was initialized with one VNT with an optimal topology for day periods, while the optimal topology for night periods is very different. In fact, the amount of transponders used by the VENTURE scheme remains stable even for high loads (although VNT reconfiguration allows to tear-down lightpaths and set-up new ones to cope with traffic directionality). In contrast, the increasing number observed for the threshold-based points out the fact that capacity increments follow increments of traffic during day periods. Finally, note that the threshold-based scheme is unable to fully exploit available transponders. As an illustrative example, for the normalized load 0.75 (about 1% blocking) circa 75 transponders are unused (25% of total). In contrast, VENTURE takes advantage of full capacity resources, thus leading to 0% blocking for the same load.

4. Conclusions

Big data analytics is applied for periodical IP traffic prediction, which is used as input of a VNT optimization algorithm named as VENTURE; its performance was compared to that of an effective threshold-based VNT reconfiguration for IP traffic variation in volume and directionality. VENTURE showed large benefits in terms of performance and resources utilization under directionality variation. Moreover, the threshold-based scheme might be also applied in case of unpredicted inter-period traffic increments.

References

[1] A. Aguado et al., "Dynamic VN Reconfiguration over SDN Orchestrated Multi-Technology Optical Domains," in Proc. ECOC 2015.
[2] A. Gençata et al., "Virtual Topology Adapt. for WDM Mesh Netws Under Dynamic Traffic," IEEE Transac. on Netw, vol. 11, 2003.
[3] Y. Ohsita et al., "Gradually Reconfiguring VNT Based on Estimated Traffic Matrices," IEEE Transac. on Netw., vol. 18, 2010.
[4] N. Fernández et al., "Virtual Topology Reconfig. in Opt Netws by Means of Cognition," IEEE J. of Opt. Comms. Netw., vol. 7, 2015.
[5] J. Boyd, "Destruction and Creation," U.S. Army Command and General Staff College, 1976.
[6] R. Adhikari, "A neural network based linear ensemble framework for time series forecasting," Neurocomputing, vol. 157, 2015.