Traffic Demand Modeling for Dynamic Layer 1 VPN Services*

Joachim Scharf, Martin Köhn
University of Stuttgart, Institute of Communication Networks and Computer Engineering (IKR),
Pfaffenwaldring 47, 70569 Stuttgart, Germany
E-mail: {joachim.scharf, martin.koehn}@ikr.uni-stuttgart.de

Abstract

Dynamic Layer 1 Virtual Private Network (L1VPN) services are an emerging and important network service in future transport networks. While on IP layer VPN concepts are already in use, the realization in lower layers is presently a hot topic in research and standardization.

For performance evaluation of a network or network dimensioning a general but concise L1VPN modeling approach is needed. However such a model, which integrates population models as well as different VPN characteristics, is currently not available.

In this paper, we outline the fundamental characteristics of VPN traffic in comparison to normal point-to-point traffic and present a generic framework for multi-point demand modeling. Our framework is open for adaptation to different service scenarios defined by population models, capacitated connections and virtual topologies.

In order to show the behavior of the framework, we apply it to one selected VPN scenario and examine in two case studies the impact on different metrics important for network dimensioning.

1 Introduction

Today, many companies—even middle-sized—are widespread over multiple sites. For many of them simultaneous data exchange between all of the sites is essential, e. g., for mirroring storage data or even email and intranet traffic. But the setup of an own physically separated communication network is very expensive and only in very few cases worthwhile. One feasible solution for this can be the use of existing infrastructure of network providers, e. g., by means of leasing fixed capacity.

Virtual Private Networks (VPN) are a concept that fulfills these requirements. According to [1] VPN is a generic term that covers the use of public or private networks to create groups of users that are logically separated from other network users and that may communicate among them as if they were on a private network. Depending on the VPN functions, a differentiation of VPNs on all lower layers in the ISO/OSI reference model is possible. VPNs on layer 2 (L2VPNs) or layer 3 (L3VPNs) provide connectivity on the respective layer between the endpoints for a certain data client, which also includes switching and routing, respectively. In contrast, for a VPN in the lowest layer, only pipes between the endpoints are provisioned (e.g., dark λ or dark fibre). Such VPNs are called L1VPN.

Fig. 1 depicts two examples for such layer 1 VPNs. In both figures, the edges of the customer's private network (customer edges, CE) as well as the edges of the provider's network (provider edge, PE) are shown. In the left figure, all CEs are interconnected with pipes leading to a full-meshed L1VPN. The right figure shows a so called hub-and-spokes scenario. This is characterized by the connection of the spoke CEs to the hub CE, whereas no direct connection between the spoke CEs exists.

Today, while in IP networks VPNs are already commercially available, the standardization of layer 1 VPN services is still not finished. The International Telecommunication Union (ITU) has defined some recommendations for L1VPNs [2, 3] and the standardization work at the Internet Engineering Task Force (IETF) is in progress [4].

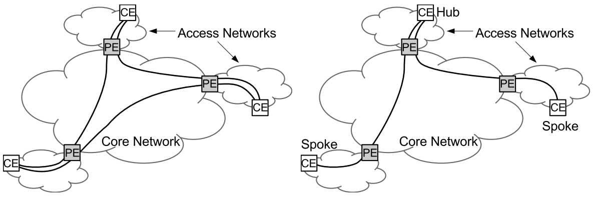

Fig. 1 Full-mesh and hub-and-spokes VPN scenarios

*This work was partly funded within the IST-Project NOBEL by the European Union under contract No. FP6-506760.

Nevertheless, for this work it is sufficient to define a L1VPN as a set of point-to-point connections for which the actual choice of the connection set remains in the hands of the customer. How these connections are established (e.g., the routing through the network) is fully under control of the network provider.

So far, only the topology of the VPN has been discussed. For most purposes, this needs not to be constant over time. In this case, either single links within the VPN are changed or even the entire VPN is only established on demand and torn down as soon as it is not used anymore. The latter case will be referred as dynamic VPN and is further discussed in this paper.

In the following, we present a model for the demands of a layer 1 VPN service. The remainder of this paper is structured as follows. Section 2 reviews a demand model for point-to-point demands. Based on this model, our demand model for VPN services is introduced and the mathematical formulation of the most relevant aspects is presented. In Section 3, numerical case studies show the most important properties of the so modeled demands. Finally, Section 4 concludes the paper and provides an outlook.

2 Traffic Demand Modeling

Network traffic depends on many aspect, e.g., the offered services, the users behavior etc., but also the spatial distribution of the users. In order to dimension a network or conduct a performance evaluation of a network, it is necessary to have a suitable traffic model describing the statistical properties with respect to time as well as an estimation of the demands in the network.

Both kinds of models can be derived from traffic measurement. Also, such data can help to choose appropriate and accurate traffic and demand models. Otherwise, if no such data is available, there is no possibility to check the chosen model against reality. Thus, only plausibility checks are possible in order to ensure the credibility of the model. However, using an approximate model may be sufficient to achieve significant results as long as the impact of the model is well understood.

To the best of our knowledge there is no demand model for dynamic layer 1 VPN services in current work. Since such services are not yet commercially deployed, there is also no measuring data available. Nevertheless, as such services may account for a relevant amount of traffic in future networks, a traffic demand model is required.

In the following, we first review a demand model for point-to-point services and highlight the fundamental ideas. Then, we present a new demand model for multi-point connections, which relies on the same basic ideas. Finally, we give a mathematical formulation of the most important functions.

2.1 Point-to-Point Demand Models

In literature, several traffic models for dynamic point-to-point services are proposed. These models describe, e.g., the distribution of the holding and interarrival time of point-to-point connections or on a rate basis the traffic between two nodes. However, there are only few traffic demand models, which describe the demand between nodes. Most of them differ only marginally from the approach proposed by Dwivedi and Wagner in [5], which is explained below.

Basically, the approach assumes that all individuals living in the area covered by the model want to communicate with each other individual. Further, the amount of communication for each pair of individuals depends only on the spatial distance between them and the type of service. For this, the distribution of individuals in the area has to be reviewed. Depending on its location, every individual is assigned to the next node of the investigated network. This results in a number of individuals for each node. Furthermore, depending on the service, a different population model is used as the originators for different service types are not necessarily identical. In this model it is assumed that all persons contribute to personal telephone communication, whereas for Internet traffic only Internet hosts are considered (cf. [5]).

According to this basic assumption, the demand between any pair of nodes is proportional to the product of the corresponding populations divided by a distance factor. This factor depends on the spatial distance of the two nodes as well as the characteristics of the modeled traffic. While for personal telephone communication the distance factor is proportional to the spatial distance of the nodes, for Internet traffic no distance dependence is assumed. Finally, all demands are scaled such that the total amount of traffic for all nodes is equal to a given total traffic volume.

It should be mentioned that this model does not consider communication between individuals assigned to the same node. This kind of communication is neglected as it does not appear in the investigated network but in the aggregation network.

2.2 Dynamic Layer 1 VPN Services

The above introduced model is only suitable for point-to-point traffic. However, the basic idea can serve as a sound basis for a VPN traffic demand model although there are much more degrees of freedom in case of VPNs.

While for a point-to-point scenario each demand relates exactly two nodes, VPNs can consist of an arbitrary number of nodes. Even more, the number of nodes can change while holding the VPN. This also includes the case of removing one node and replacing it by another.

Further, a VPN can be seen as a set of dependent point-to-point connections. These connections need not to be identical with respect to their capacity. This includes the case that the capacity for particular node pairs is zero, i. e., a connection between all node pairs within a VPN is not needed. Moreover, similar to the point-to-point scenario, the capacity need not to be constant over time.

Concluding, a VPN can be characterized by the set of participating nodes v and the set of connections c which both can optionally change over time. We will assume in the following a VPN to be unchanged during its holding time, i. e., neither the capacity of the connections will change, nor a node participating in the VPN will be added or removed.

With these assumptions, the demand in a point-to-point scenario can be extended to the demand in a VPN scenario. In a point-to-point scenario, the demand between two nodes can be seen as the amount of traffic that will be exchanged between these two nodes. Translating this, the demand in a VPN scenario can be defined as the total amount of traffic that will be exchanged between the participating nodes. According to this definition, the demand of a VPN is equal to the sum of all demands between any two nodes in the VPN.

To determine the traffic demand created by the VPNs that are possible in a certain scenario, we have to start from the ideas of the point-to-point model. As for that model, a number of individuals derived from a population model is assigned to each node. These individuals will be the edge nodes of the VPNs. Furthermore, all individuals are assumed to be equal and thus have the same probability for being connected to a VPN.

At this point two cases must be distinguished. On the one hand, only such VPNs can be considered, in which at most one individual per core node is connected to a VPN. On the other hand, additionally a second type of VPN can be considered, in which more than one individual participating in a VPN is connected to the same core node. It can be easily seen that considering both types of VPNs leads to an exact fulfillment of the equality assumption, whereas the first case is only a approximation of this assumption. But as our target is to model VPNs in core networks without considering the individuals in detail, only the number of edges in the core network is of relevance. Furthermore, the case of connecting at least two individuals within the same area to one VPN can be realized by a separate VPN in this area. Thus, we limit ourselves to the first case and will concentrate in the following on VPNs with the same fixed number m of pairwise disjoint nodes.

Now, after defining the node sets, the connection sets have to be defined. Here several topologies can be used, e. g., a star for modeling a hub-and-spokes scenario or a full-mesh for modeling the connection of different company sites. Also, for each set of nodes a different connection structure can be chosen. However, in this work we do not limit ourselves to a certain structure, but assume the connection structure for all VPNs to be identical. Thus, the VPN can be uniquely identified by its node set v.

2.3 Mathematical Formulation

In this subsection we provide definitions for the most relevant functions. First, the occurrence probability of different node sets will be calculated. Then, we define the properties the connection set must fulfill and derive the demand of each VPN. Finally, we calculate the overall demand using the aforementioned occurrence probabilities and VPN demands.

In $\mathbf{V}_m$, all possible sets of m pairwise disjoint nodes are accumulated. It can be seen that in a network with n nodes there are

$$|\mathbf{V}_m| = \binom{n}{m} \tag{1}$$

of such node sets. The occurrence probability $p_{v,m}$ of a node set $v \in \mathbf{V}_m$ is given by

$$p_{v,m} = c_m \cdot d_{v,m} \cdot \prod_{i \in v} a_i . \tag{2}$$

In this equation, a_i denotes a certain number of individuals assigned to node i. The factor $d_{v,m}$ is called the distance factor and characterizes the impact of the distances between the nodes. The constant c_m normalizes the sum of all probabilities to be equal to 1. Accordingly, c_m is defined as

$$c_m = \frac{1}{\sum_{v \in \mathbf{V}_m} \left(d_{v,m} \cdot \prod_{i \in v} a_i \right)} . \tag{3}$$

With the probabilities $p_{v,m}$, the probability for a core node to participate in a dynamic VPN $p_{N\ i,m}$ can be determined by

$$p_{N\ i,m} = \sum_{v \in \{k \in \mathbf{V}_m | i \in k\}} p_{v,m} . \tag{4}$$

According to the definition above and similar to the calculation of the demand of a node pair, the demand A_v of VPN v can be calculated by

$$A_v = \sum_{i,j \in v} A_{v,ij} \tag{5}$$

with $A_{v,ij}$ being the capacity of the connection between the nodes i and j in the VPN v. Finally, the total demand of all VPNs A_m with m nodes can be calculated by

$$A_m = \sum_{v \in \mathbf{V}_m} p_{v,m} \cdot A_v . \tag{6}$$

This approach can be easily extended to VPNs with different number of nodes by assigning a probability to each set $\mathbf{V}_m$. Similar, for a node set different connection sets can be assumed. However, both extensions are not considered further in this work.

3 Numerical Case Studies

In this section, results of a numerical evaluation will be presented and discussed. We show general properties in a simple scenario and finally investigate a more realistic reference scenario.

For the first part of the case study, an abstract scenario is used predominantly to reveal characteristics of the model. The network consists of 16 nodes. Each of them is assigned to one of two groups *A* and *B* that differ in the population. 4 nodes belong to group *A*, the remaining 12 to group *B*. The population of each node of group *A* is f times larger than that of a node of group *B*.

The second part of the case study is based on a reference scenario in order to show the behavior of the model in an almost realistic network. The scenario used here relies on the COST Core Network scenario presented in [6]. The network consists of 16 nodes and represents a central European network. The according populations of Internet hosts per node are given in **Tab. 1**.

city	Internet Hosts		city	Internet Hosts	
	in million	**%**		**in million**	**%**
London	16.5	20.5	Zurich	3.4	4.2
Amsterdam	6.8	8.4	Lyon	3.4	4.2
Berlin	6.5	8.1	Paris	2.8	3.5
Frankfurt	6.5	8.1	Strasbourg	2.8	3.5
Hamburg	6.5	8.1	Brussels	2.7	3.3
Munich	6.5	8.1	Vienna	2.7	3.3
Milan	5.5	6.8	Prague	2.2	2.7
Rome	5.5	6.8	Zagreb	0.3	0.4

Tab. 1: Internet hosts per node (absolute and relative) in European reference network scenario

The nodes in this reference network scenario can also be classified with respect to their population. We separate them into four groups. The groups with the largest and the smallest nodes consist of only a single node which is London and Zagreb, respectively. All other nodes can be divided into the two remaining groups: one group with nodes with a population of about 6 million Internet hosts and the other group with nodes with a population around 3 million.

With respect to the topology of the VPNs, we assume them to be full-meshed and the capacity of all connections is equal. In order to remove the impact of the network topology, the distance factor $d_{v,m}$ is set to 1, i. e., the distance between the nodes of a VPN has no effect. It can be easily shown that for these assumptions the actual population has no influence and only the relation between the populations is of importance.

One relevant metric to show the characteristics of the network is the fraction of the total demand $w_{ij,m}$ between any node pair i, j. This allows to compare the demands of the VPN model to those of a point-to-point service. For the scenario considered here, the demand can be calculated for a pair of nodes by

$$w_{ij,m} = \frac{1}{m(m-1)} \cdot \sum_{v \in \{k \in \mathbf{V}_m | i,j \in k\}} p_{v,m} \tag{7}$$

For other scenarios an analogous procedure is possible, but will result in another formula.

3.1 Abstract Scenario

First, the basic characteristics of the model shall be investigated. In **Fig. 2**, the node participation probability $p_{N\,i,m}$ for this abstract scenario is shown versus the number of nodes per VPN for different f. As expected, it can be seen that if the population in all nodes is equal, i. e., $f = 1$, the node participation probability is also equal for all nodes and increases linearly.

The behavior of the other extreme, namely $f \rightarrow \infty$, is also very intuitive. Due to the large population of nodes of group *A*, VPNs with 4 or less nodes are build up only by nodes of group *A*. Accordingly, the probability of group *B* nodes converges to 0. For VPNs consisting of more than 4 nodes, nodes belonging to group *B* must be part of the VPN. This explains the increase of the participation probability of group *B* nodes. However, all group *A* nodes are involved in each VPN and thus their participation probability remains 1.

For all other values of f, the curves lie in the area specified by the two cases. It can be shown that the node

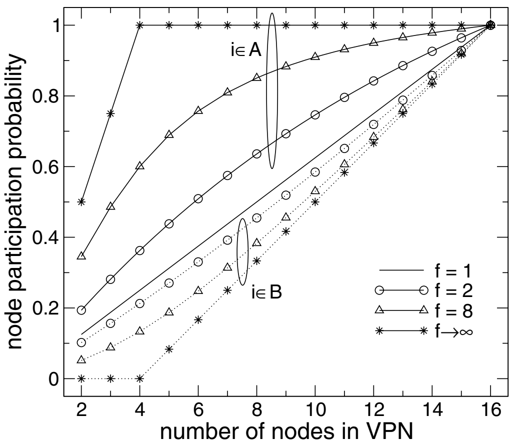

Fig. 2 Node participation probability in abstract scenario

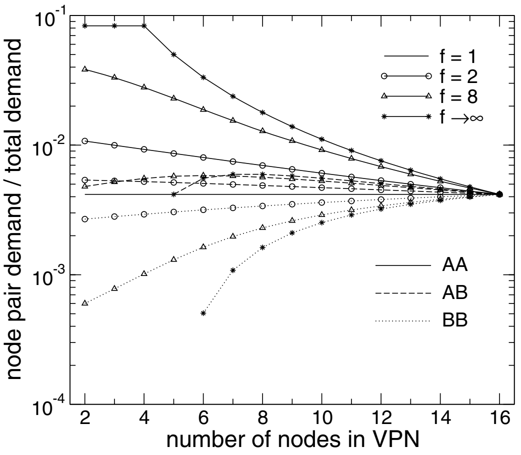

Fig. 3 Share of demand between node pairs in abstract scenario

participation probability for this scenario is monotonic increasing. However, the proof for this behavior is omitted here.

With increasing number of nodes per VPN m, the differences between nodes of group *A* and *B* diminish and are vanished if m reaches the number of nodes in the network n. This effect does always occur despite differences in population. The reason can be found in the assumption taken above that every node can appear in every VPN at most once.

In **Fig. 3**, the relative demand per node pair $w_{ij,m}$ is plotted versus the number of participating nodes in the VPN for different f. Obviously, with two groups of nodes, three types of node pairs are possible (*AA*, *AB* and *BB*). It can be seen that for $f = 1$ the relative demand per node pair is independent of the node pair. Even more, this holds also independent of the number of nodes in the VPN. As in this case the probability for all node pairs being part of the VPN is equal, the share of demand for each node pair must also be equal. With $16 \cdot 15 = 240$ unidirectional communication relationships between node pairs, the share of demand is $1/240$ of the total demand. This also reasons the fact that for $m = 16$, i. e., all nodes participate in all VPNs, all shares are equal independent of the relation of the populations.

In the special case of $f \to \infty$, the fraction of demand between nodes *AA* remains constant at first. In this case, all demand is split up among the *AA* node pairs until at least one node of group *B* appears in a VPN. This is the case for $m = 5$. Then, the fraction of demand decreases.

The share of demand between nodes of group *B* is low for a small number of nodes in the VPN and a large f and increases with the number of nodes. The occurrence of such demand necessitates at least two nodes of group *B* in the VPN, i. e., for f going to infinity this is not the case for $m < 6$.

The fraction of demand between nodes of both groups behaves nonuniform but is always (except $m = 16$) below that of *AA* and above *BB* node pairs.

Finally, both figures show the general behavior that the bigger the difference in population is, the bigger are the differences in the share of demand and the node participation probability, respectively.

3.2 European Reference Network

After discussing the principle effects, this case study for a reference scenario will show the behavior of the model in an almost realistic network. **Fig. 4** shows the node participation probability for the European reference network scenario versus the number of nodes in the VPNs. The prominent position of London and Zagreb has great influence on this metric. It should be highlighted that London is represented nearly in every second 3-node VPN. For Zagreb, 15 nodes per VPN are necessary to reach that point. The two remaining groups can be seen clearly, but do not show any surprising behavior.

In contrast to the previous abstract scenario, the node populations are now constant and in particular finite. This is the reason, why all nodes have always a participation probability greater than 0. However, the general behavior of monotonic increasing node probability meeting at 1 for $m = n$ still holds.

Fig. 5 depicts the share of demand per node pair versus the number of nodes per VPN. Here, two types of connections shall be highlighted. First, the connections between London and any other node experience a

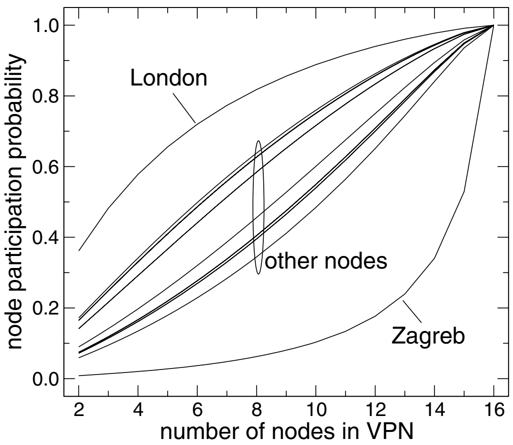

Fig. 4 Node participation probability in European reference network scenario

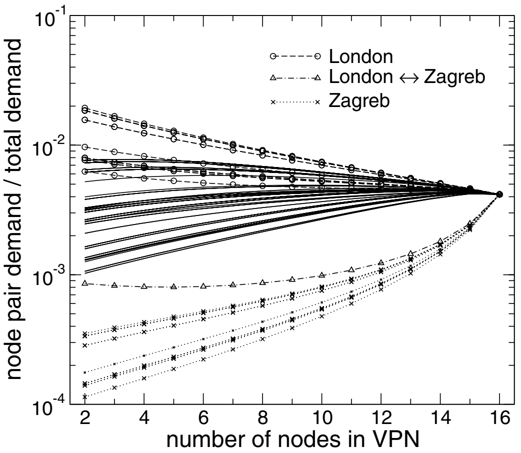

Fig. 5 Share of demand between node pairs in European reference network scenario

(partly tremendous) decline in the fraction of demand. The only pair with finally increasing demand is London-Zagreb. Second, all node pairs with Zagreb have only a small amount of demand, which increases considerably not until nearly all nodes occur in the VPN.

The remaining node pairs show a nonuniform behavior. The values are increasing with growing number of nodes for some while there is a decline after an increase for others.

In analogy to the result of the abstract scenario, for $m = 16$ the share of demand of each node pair is $1/240$ of the total demand. As already mentioned, the node populations are finite in this scenario. Thus there is at least a small demand between every pair of nodes for an arbitrary number of nodes in VPN.

4 Conclusions and Outlook

In this paper, we presented a new generic demand model for L1VPNs, which facilitates network dimensioning and performance evaluation. It is generic with respect to the population model, the virtual topology of the VPNs as well as different traffic types.

Two case studies have been performed for one specific type of VPN, namely full-meshed VPNs with uniform connection capacity. The results show, that changing the VPN size can lead to tremendous differences in demand between nodes. Those effects may have an important influence on network dimensioning.

As currently no measuring data of VPN traffic is available, an evaluation of the proposed traffic demand model remains for future work. Also, future work could use this model for evaluation of network dimensioning algorithms as well as VPN traffic engineering concepts. Beyond that, extensions for VPNs with changing topology or node set as well as for L2VPNs or L3VPNs are feasible.

Acknowledgments

The authors would like to thank Christoph Gauger and Sebastian Gunreben for invaluable discussions.

References

[1] L. Andersson, T. Madsen: “Provider Provisioned Virtual Private Network (VPN) Terminology.” *RFC4026 (Informational)*, March 2005

[2] “Layer 1 Virtual Private Network generic requirements and architecture elements.” *ITU-T Recommendation Y.1312,* September 2003

[3] “Layer 1 Virtual Private Network service and network architectures.” *ITU-T Recommendation Y.1313,* July 2004

[4] “Layer 1 Virtual Private Networks (l1vpn).” *IETF Working Group*, http://www.ietf.org/html.charters/l1vpn-charter.html

[5] A. Dwivedi, R. E. Wagner: “Traffic model for USA long-distance optical network.” *In Proceedings of Optical Fiber Communication 2000*, Baltimore, 2000, pp. 156-158

[6] S. De Maesschalck, D. Colle et al.: “Pan-European Optical Transport Networks: An Availability-based Comparison.” *Photonic Network Communication*, Vol. 5, No. 3, 2005, pp. 203-225