Evaluation of End-to-End Quality of Service over VPN Networks through Various Priority Mechanisms

Nasser-Eddine Rikli
Department of Computer Engineering
King Saud University
Riyadh, Saudi Arabia
e-mail: rikli@ksu.edu.sa

Saad Almogari
Department of Computer Engineering
King Saud University
Riyadh, Saudi Arabia
e-mail: smogri@nic.org.sa

***Abstract*— VPN networks running over MPLS have found widespread acceptance as both an efficient and cost effective means to provide connectivity for large organizations and companies. However, providing QoS is still a major challenge that needs to be addressed. Using realistic input traffic, a simulation model is built for a large network where various queueing policies are implemented and evaluated for the provision of certain QoS requirements. After a thorough analysis the merits and shortcomings of each policy are determined, and recommendations are given along with future research directions.**

Index Terms—Virtual private networks; quality of service; multimedia; MPLS; queueing mechanisms.

I. INTRODUCTION

Quality-of-Service (QoS) over Virtual Private Networks (VPN) is prone to many challenges, among which setting policies for a flexible and scalable support of QoS is of primordial importance [1][2]. Any provider of VPN service should be able to offer customers various Classes of Service (CoS) per VPN [3]. Furthermore, depending on the customer choice and selection, the CoS that a particular application would get within one VPN could be different from the CoS that exactly the same application would get within another VPN. Thus, the set of policies to support QoS should allow the decision to be made on a per-VPN basis.

VPN has used two models in providing QoS, namely the *pipe* model and the *hose* model [4]. In the former, a customer is supplied with certain QoS guarantees for the traffic from one Customer Edge (CE) router to another. While in the latter, a customer is supplied with certain guarantees for the traffic that the customer's CE router sends to and receives from other CE routers over the same VPN.

In [5], a programmable framework for CoS Based Resource Allocation (CBRA) in Multi Protocol Label Switching (MPLS) tunneled VPNs is proposed. The resources are partitioned in a way that facilitates the creation of multiple VPNs on a demand basis.

In [6], the QoS over a VPN IP network is presented from a service provider point of view. The study includes the provision of QoS guarantees both at the network level and at the node level.

In [7], a CoS classification with associated QoS parameter set for VPNs over an IP WAN is presented. Various scenarios were studied, and it was determined that by policing the aggregate arrival rates of each class from each VPN access interface into the IP network, the appropriate QoS can be guaranteed for each CoS.

The main purpose of this paper is to propose a simulation model and to study the behavior of a VPN network under various queueing mechanisms and for various types of traffic. A thorough network performance analysis will be carried out for various traffic types with different QoS requirements. A special emphasis will be given to the effects of the bandwidth of last mile link at the main site.

The rest of the paper is organized as follows. In Section II, the architecture of the network to be studied will be presented. Then, in Section III the traffic models and traces to be used in the simulation will be described. In Section IV, the queueing models to be used in the various routers will be introduced. The results will be presented in Section V, along with some network specific data. Finally, in Section VI, conclusions will be summarized.

II. NETWORK ARCHITECTURE MODEL

Based on an existing network, a simulation model for a customer with four sites connected through a VPN service provider (VPN-SP) network was built. The general network architecture is shown in Fig. 1. The network topology of the VPN-SP consists of:

1. Three Provider (P) routers, located at the customer headquarter.
2. One P router and one Provider Edge (PE) router, located at each one of the three satellite locations.
3. Four CE routers: one at site 1, one at site 2, one at site 3, and one at the main site.

The VPN services are assumed to be provided through a

ISBN: 978-1-61208-127-4

hose model, and most traffic is assumed to pass through the router at the main site (whether it is coming from other sites or passing through towards them).

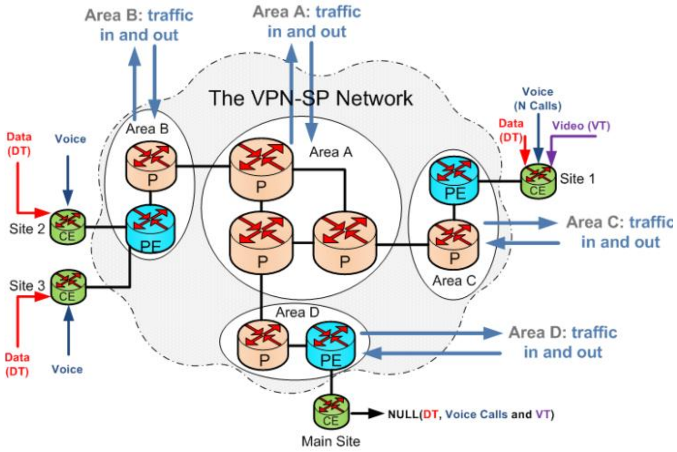

Fig. 1. Network architecture and traffic input locations and types.

The routing protocol used between a CE router and a PE router is the Border Gateway Protocol (BGP). At the PE router, each site connects its customers through an interface that marks all outgoing traffic with a unique VPN label to mark its traffic between PE routers.

Routing table information is exchanged between PE routers using Multiprotocol BGP (MP-BGP). The VPN-SP uses Multiprotocol Label Switching (MPLS) over Open Short Path First (OSPF) network.

III. TRAFFIC MODELS

A. *Types*

The VPN-SP network carries various types of traffic generated by the different customers. We divided the aggregate traffic into three kinds: voice traffic, video traffic, and data traffic.

Voice traffic is assumed to be generated using a G.729 coder. The aggregate traffic model for VoIP was modeled by an ON-OFF source with Exponential durations. During the ON period, packets of fixed size are generated at fixed time intervals [9].

The two other types of traffic, i.e. MPEG-4 video and data, were captured into trace files from the real traffic flows at the various locations of the actual VPN-SP network using a sniffer tool. These files were used as input at their corresponding locations to simulate real traffic from site-to-site of the chosen customer (or inside the VPN-SP network when coming from other customers).

B. *Load Distribution*

The diagram in Fig. 1 illustrates the distribution of the three types of traffic over the various sites. Voice, video, and data traffic were sent from site 1 to the main site, while only voice and data traffic were sent from site 2 to main site, and the same thing from site 3 to the main site. Furthermore, each one of the four Areas (A, B, C, and D) has both external input traffic and output traffic leaving the network. It is assumed that all flows include the three types of traffic.

C. *Requirements*

The QoS traffic requirements are shown in Table I. They were chosen to satisfy both generic requirements of the types of application carried over the network, and the specific requirements of the equipment existing on the premises.

TABLE I. TRAFFIC REQUIREMENTS

Criteria	**Voice**	**Video**	**Data**
packet delay (msecs)	< 200	< 250	-
Jitter (msecs)	< 40	< 40	-
packet loss[1] (%)	< 5	< 10	-
packets resent (%)	-	-	< 10

IV. QUEUEING MODELS

A. *Description*

Various queueing policies may be implemented at the different routers of the considered network. In this study, four types will be considered:

4. *Fair queuing (FQ):* where the traffic is divided into three flows (video, voice, and data) with separate FIFO queues, and served through a round-robin scheduling (each queue sends one byte in every round).
5. *Priority queuing (PQ)*: similarly packets are classified into three queues but served with priority one for voice traffic, priority two for video traffic, and priority three for data traffic. Within each queue packets are served in FIFO. If a newly arriving packet finds the queue full, then it will be dropped.
6. *Custom queuing (CQ)*: it is similar to PQ in that it also supports a certain classification option. The scheduling, however, is completely different. It uses a round-robin service, in which each queue is allowed to forward a certain number of bytes (not packets). The queues are served in a weighted round-robin scheme. Depending of the weight (% of share) the available bandwidth is distributed among queues. Tail dropping is still used with each individual queue. We study two cases of the custom queueing which are commonly used in real networks: (1) 10% voice, 20% video, and 70% data, and (2) 20% voice, 30% video, and 50% data.
7. *low-latency queuing (LLQ)*: it is a combination of PQ

---

[1] In here, packet loss includes both the number of dropped packets and delayed packets.

and CQ policies. The first queue has the highest priority, and is still served first. If the first queue is empty then the second and third queues will be served based on a partition of 40% for the second queue and 60% for the third queue. In this study, the first queue was assigned to voice flow, the second to video flow, and the third to data flow.

B. *Placement*

Fig. 2 shows the location of the ports of each router where the proposed queueing mechanisms will be implemented. So, each P router in Area A has three ports, while each P router in the remaining areas (B, C, and D) has only two ports. Also, the PE routers in areas C and D have two ports each, while the PE router in area B has three (since it is connected to two sites). Finally, all CE routers have a single port.

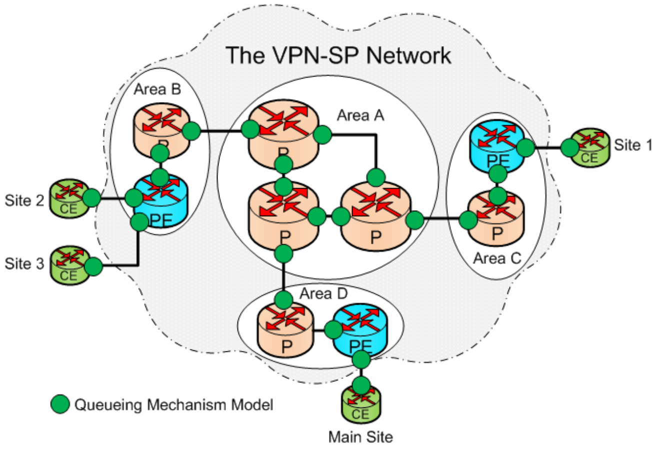

Fig. 2. Points of implementation of the queueing mechanisms.

V. RESULTS

To investigate various aspects of the effects of the queueing policy on the performance of our network, two sets of experiments have been designed. Using different queueing mechanisms, five experiment variations were undertaken in each set.

The simulation experiments were built using NS2, and run for one hour of simulation time. All router queues were assumed of finite buffer sizes and had a total size of 512 KBytes (KB) with 128 KB for the first queue, 128 KB for the second, and 256 KB for the third.

The router capacities were 1 Gbps for the core P routers, 10 Mbps for the area P satellite routers, 1 Gbps for the PE routers, and 1 Mbps for all CE routers except the one at the main site which had a 2 Mbps.

A. *Effects of the Number of Channel Calls*

In the first set, the effects of the voice traffic on the VPN-SP's network was studied by increasing the number of voice calls, initiated from site 1 and going to the main site, from 1 to 7 channel calls. The same experiment was repeated using five different queueing mechanisms. Our focus will be on the traffic flowing from site 1 to the main site, including voice, video, and data.

1) *Effects on Voice Traffic*

Fig. 3 shows the percentage of voice packets dropped due to an excess delay of 200 msecs. The best results were obtained when using the PQ and LLQ mechanisms, which have very similar results. This is due to the fact that voice has the highest priority in both schemes.

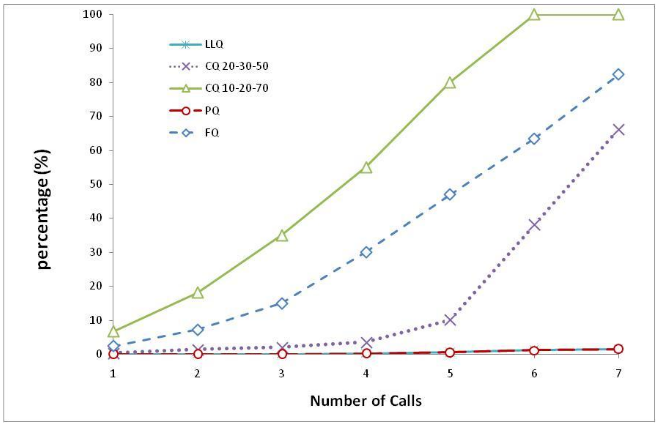

Fig. 3. Percentage of Voice Packets with Delay over 200ms.

The CQ 20-30-50 mechanism was able to handle up to four voice calls dropping rate less than 5%), while the CQ 10-20-70 mechanism barely handled one call. However, in both cases, the results were worst than the ones achieved with PQ and LLQ. This is because not all voice traffic has the highest priority, with an advantage of the 20% scheme over the 10% scheme since a higher share of its traffic was privileged.

Lastly, the FQ mechanism was not able to handle even one call, since there is no priority mechanism implemented. Also, we notice that the performance trend is almost constant with PQ and LLQ mechanisms, while with all other mechanisms it deteriorates rapidly after a certain number of calls.

2) *Effects on Video Traffic*

Fig. 4 shows the dropping rate for video traffic exceeding 250 msecs as the voice traffic is increased. The best performance was achieved through the CQ 20-30-50 and FQ mechanisms, with a slight advantage of the latter. As the voice traffic increased, the video performance was kept very close to the required bound.

In the case of the LLQ and PQ mechanisms, the video traffic performance was kept acceptable up to four calls, and then it deteriorated very quickly.

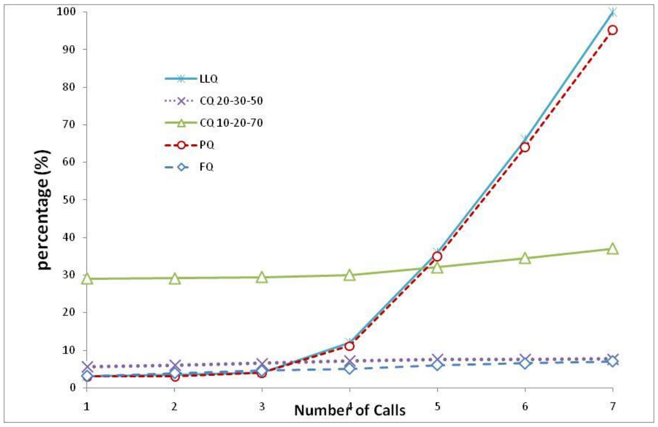

Fig. 4. Percentage of Video Packets with Delay over 250ms.

Lastly, for the CQ 10-20-70 mechanism, although the performance was kept almost constant, it was very far from the required limit.

These results are in concordance with the fact that video traffic has the second priority in the LLQ and PQ mechanisms, where the performance was the best when the first priority traffic (i.e., voice) was comparatively low (< 5 sources). As the first priority traffic was increased, all lower priority traffic suffered. In the case of the other mechanisms, the share of the video traffic was not affected by the increase in voice traffic.

Here also, the CQ mechanisms have better performance than the FQ mechanism, since they use some sort of priority for video. Furthermore, the 30% CQ case performed better than the 20% one, although the voice share also was decreased from 20% to 10%.

3) *Effects on Data Traffic*

Fig. 5 shows the retransmission rate of data traffic as the voice traffic was increased. The best performance was achieved through the CQ 10-20-70, CQ 20-30-50, and FQ, with the former being the best and the latter the worst. In the three cases, the results were kept almost constant, in accordance with non-prioritized mechanism or partially prioritized ones. Here also, the mechanism that allowed 70% of the data traffic to be served as a third priority performed better than the one allowing only 50%.

In the case of LLQ and PQ mechanisms, the performance was kept constant up to three calls, and then increased rapidly. However, while the LLQ performance was acceptable before the three calls knee, the PQ's was unacceptable in all cases. This is similar to the video traffic results, but with a much larger gap in favor of LLQ.

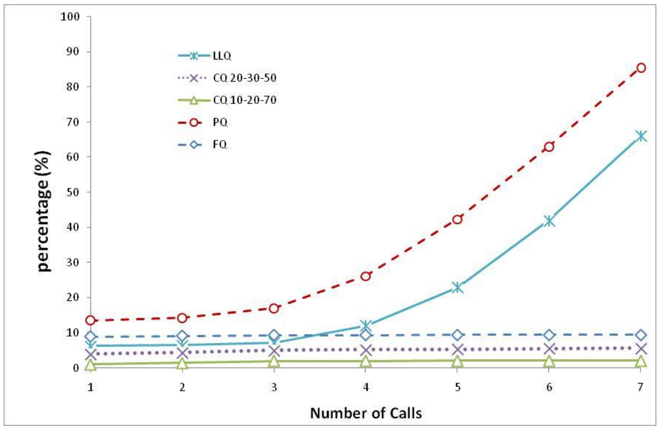

Fig. 5. Percentage of data packets being resent.

B. *Effects of Last-Mile Bandwidth*

In the second part of experiments we want to study the effects of the last-mile bandwidth. It is the channel capacity of link connecting the CE router to the PE router at the main site, and it is expected to be the bottleneck for the customer's traffic behavior in the VPN-SP's network.

Its effects will be studied by increasing the capacity of the link from 128 Kbps to 8 Mbps. Here also, the five different queueing mechanisms will be tested, and the performance of the voice, video, and data traffic from site 1 to the main site will be monitored.

1) *Effects on Voice Traffic*

Fig. 6 shows the dropping rate for voice traffic that exceeds a 200 msecs delay as a function of the last-mile bandwidth and for the various queueing mechanisms. In all cases the dropping rate decreases as more bandwidth is made available at the bottleneck link. The same relative performances were obtained as in Fig. 3.

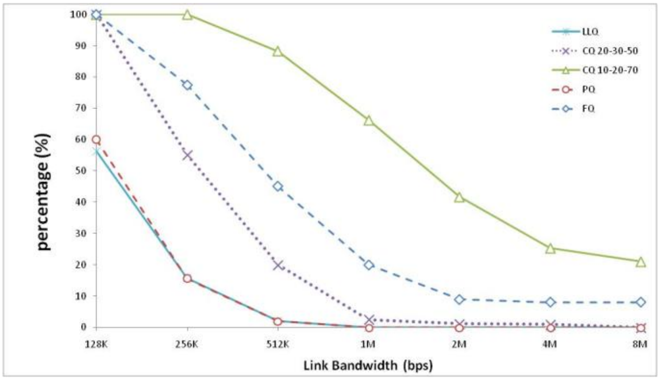

Fig. 6. Percentage of voice packets with delay over 200ms.

The LLQ and PQ mechanisms achieved acceptable performance for bandwidths larger than 512 Kbps, the CQ 20-30-50 mechanism required at least 1 Mbps, while FQ and CQ 10-20-70 failed for all bandwidths.

2) *Effects on Video Traffic*

Fig. 7 shows the dropping rate of video traffic as a function of the last-mile bandwidth. The CQ 10-20-70 mechanism had a poor performance for all bandwidth values, while the remaining mechanisms had very close performance, with a bandwidth requirement of at least 2 Mbps. The PQ mechanism achieved the best performance for all bandwidths.

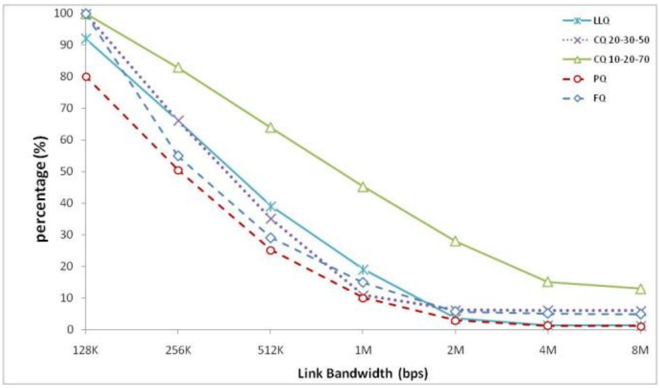

Fig. 7. Percentage of video packets with delay over 250ms.

3) *Effects on Data Traffic*

Fig. 8 shows the retransmission rate of data traffic as a function of the last-mile bandwidth. The minimum required bandwidth for acceptable data traffic performance were summarized in Table 2. The two CQ mechanisms achieved the best performance, with a noticeable advantage of CQ 10-20-70, which had a larger fraction reserved for data (70%), and this was true for all bandwidth values.

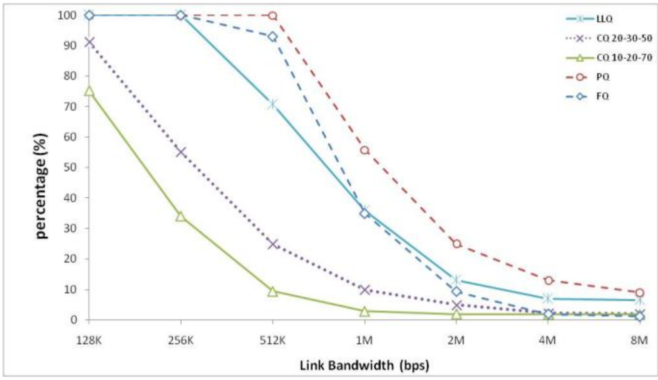

Fig. 8. Percentage of data packets being resent.

TABLE II. MINIMUM BANDWIDTH FOR ACCEPTABLE DATA PACKETS RESENT.

Mechanism	PQ	FQ	LLQ	CQ20-30-50	CQ10-20-70
***BW_{min}* (Mbps)**	8	4	4	1	0.512

The PQ mechanism, which gives data traffic the least priority, achieved the worst performance. With high bandwidths, the FQ mechanism reaches the same level of performance as the CQ mechanisms.

VI. CONCLUSION

In this paper, we have considered a large VPN-SP network providing service to a customer with four remote sites. A simulation model was built with real traffic input, and run under various service policies with the QoS performance being observed.

Four queuing mechanisms were considered, namely: FQ, PQ, CQ (two versions), and LLQ. Criteria for acceptable performance was set for each carried traffic type which was assumed to be carried over the network.

As a result, an estimation of the impact of a new voice call on the performance of the other traffic types being carried over the network was quantified. Consequently, it was possible to determine the limitation on the number of calls in each customer's sites.

Finally, we varied the bandwidth of the last-mile link located at the customer's main site, given that it was considered as the main bottleneck to the traffic being carried. Consequently, it was possible to advise the service provider whether to increase the bandwidth of the last-mile link at the main site if the need for accepting more customers of certain type may arise.

REFERENCES

[1] M. Rahimi, H. Hashim, and R.A. Rahman , "Implementation of Quality of Service (QoS) in Multi Protocol Label Switching (MPLS) networks," 5th International Colloquium on Signal Processing and Its Applications, pp. 98-103, March 2009.

[2] M. El Hachimi, M.-A Breton, and M. Bennani, "Efficient QoS Implementation for MPLS VPN," International Conference on Advanced Information Networking and Applications, pp. 259-263, March 2008.

[3] F. Luyuan, N. Bita, J.-L. Le Roux, and J. Miles, "Interprovider IP-MPLS services: requirements, implementations, and challenges," IEEE Communications Magazine, vol.43, no.6, pp. 119-128, June 2005.

[4] N.G. Duffield et al., "Resource management with hoses: point-to-cloud services for virtual private networks," IEEE/ACM Transactions on Networking, vol.10, no.5, pp. 679- 692, Oct 2002.

[5] P. Kumar, N. Dhanakoti, S. Gopalan, and V. Sridhar, "CoS Based Resource Allocation (CBRA) in VPNs over MPLS", IEEE Workshop on IP Operations and Management, pp. 140-145, 2004.

[6] J. Zeng and N. Ansari, "Toward IP Virtual Private Network Quality of Service: A Service Provider Perspective", IEEE Communications Magazine, vol.41, no.4, pp. 113-119, April 2003.

[7] M. Girish, J. Yu, and T. Soon, "A QoS Specification Proposal for IP Virtual Private Networks", IEEE Workshop on IP Operations and Management, pp. 85-90, Oct. 2003.

[8] C. Metz, "The latest in VPNs: part II," IEEE Internet Computing, vol.8, no.3, pp. 60- 65, May-Jun 2004.

[9] H. Hassan, J.M. Garcia, and C. Bockstal, "Aggregate Traffic Models for VoIP Applications," IEEE international conference on Digital Telecommunications, pp. 70, Aug. 2006.