CVS: using VLANs to counteract the effect of topology changes in quasi-static mesh access networks

Vincenzo Mancuso
DIEET, Università di Palermo
vincenzo.mancuso@tti.unipa.it

Luigi Monica
DIEET, Università di Palermo
luigi.monica@gmail.com

Abstract

Mesh networks are candidate to play the role of switched Ethernet LANs over extended areas and with a sensibly higher flexibility. Actually, mesh networks can exploit both Ethernet and wireless technologies, e.g. Wi-Fi and/or free-space optical links, to provide a high degree of redundancy in an access network, and to provide users with powerful means to connect with each other. By using such technologies, the mesh topology remains stable for hours, so that the mesh topology can be considered quasi-static. However, large meshes require a heavy overhead in the network control plane as for the management of end-to-end paths, which can change due to mobility of users and due to occasional failure of wireless links.

In this paper we show that a mesh network can be endowed with dynamically-managed VLANs in order to tackle the impairment caused by topology changes. In fact, the availability of multiple VLANs offers redundancy enough to hide network changes to end users. However, a continuous rearrangement of unused (redundant) VLANs is strictly required. We call that approach CVS (Controlled VLANs Switching), since network traffic should be run-time switched to different VLANs, depending on the active network topology.

1. Introduction

Microwave and optical technologies are largely adopted in telecommunication networks, also in the case of local area networks. Wireless devices can be deployed in a LAN, in conjunction with wired network nodes, and permit to easily build complex mesh networks in a short time, and with reduced fixed costs. Wireless is also suitable to merge multiple LAN segments into a switched LAN, both in the case of point-to-point links (e.g., free-space optical links [1]) and point-to-multipoint links (e.g., IEEE 802.11 WDS [2]), and handover systems have become fundamental to allow connection-oriented service for mobile users.

Handovers are critical in mesh networks since they could turn into long-lasting end-to-end path updating procedures, during which several end-nodes cannot be connected. Unluckily, the cause of handover stays in user mobility and physical changes in the connectivity of the access network that generate topology changes in the network. Path changes are critical, since nodes with outdated path information could cause frame loss, whereas nodes that have no knowledge about network routes, simply use to flood the network with their data frames. Thus, wideband applications could cause a congestion if their packets flooded. In such a scenario, traditional spanning-tree-based algorithms [3] are not able to provide robustness to the network with respect to the dynamic behavior of nodes, and cannot avoid unnecessary network flooding. Nonetheless, flooding of backup traffic has been commonly used in optical networks to protect and restore communication paths [4]. Note that flooding and lack of connectivity damage several applications, first of all real-time and interactive applications, and also TCP performance. So flooding should be limited, and reconfigurations should be speed up as much as possible.

The aim of this paper is to show that standard Virtual LANs (VLANs) can be adopted in order to provide multiple paths per connection, so that alternative paths are available when a link or a node fails. Controlled VLANs Switching (CVS), an off-line strategy for configuring of multiple VLANs, can allow the service provider to minimize the impact of the two main detrimental effects due to legacy reconfiguration procedures: i) temporary end-to-end disconnection, and ii) network-wide flooding of data packets.

The paper is structured as follows. In Section 2 we propose to use IEEE standards for bridging and switching in quasi-static meshes, and we review techniques aiming at internal LAN operation. In Section 3 we present an overview of routing proposals for mesh networks, when preservation and restoration of end-to-end paths is explicitly taken into account. Section 4 explains the rational of using CVS and details the CVS implementation. Section 5 discusses CVS performance in mesh networks, and Section 6 concludes the paper with comments and open issues.

2. Switched LAN technologies

In this section we briefly review some of the most important standard facilities that are available in the Ethernet scenario (i.e., IEEE 802.1 bridging and switching). These are mechanisms that apply to multiple MAC technologies. They can also be extended to new MACs, thanks to the presence of a MAC-independent Logical Link Control (IEEE 802.2, or LLC). Wireless meshes using IEEE 802.11 (e.g. Wi-Fi) are already LLC-compatible. This remains true for free-space optical devices [5], even though they more frequently adopt ATM/SONET [6]. Thus, IEEE bridging is the natural candidate for the management of heterogeneous mesh network comprising Wi-Fi Access Points, free-space optical devices and Ethernet nodes.

Switched networks are characterized by distinct physical and logical topologies. The physical topology of a network allows devices to communicate with each other. The logical topology coordinates the information exchange through the physical topology. When multiple paths between subnets are physically available, it is a matter of logical topology to avoid that packets looping endlessly across the physical network.

The Spanning Tree Protocol (STP, specified in IEEE standard 802.1D) is used by bridges in order to detect and manage redundant links within a network. The management is port-based: devices exchanges information such as MAC addresses and ports of bridges/switches, and link costs. Devices communicate by means of particular frames, namely Bridge Protocol Data Units (BPDU): they elect a root node, figure out the distance from the root and select a port with lowest path cost towards the root direction (designated root port). Segment loops are detected and removed thanks to a flooding exchange of BPDUs. Thus, bridges build up a tree that includes each network node. Traffic will flow through ports labeled as *forwarding.* Note that redundant links can be found on a bridge port, but they are prevented to handle data traffic by setting the port in *blocking* status (see [3] for further details). Devices exchange periodic updates by means of BPDUs, and if a change notification reaches the root bridge, a reconfiguration procedure starts. Ports are switched from *blocking* to *forwarding* (or vice versa) upon a command is originated by the root bridge.

As for the data plane, frame forwarding is optimized by means of an address learning process: bridges retain in a *filtering database* the source address conveyed by a received frame, jointly with the identifier of the port where the frame had come. As soon as a new frame reaches a bridge, it is forwarded through the port specified in the entry of the filtering database matching the destination address. Otherwise, the frame is forwarded to all active ports, except for the arrival port.

An extension of STP is the so called Rapid STP (RSTP, IEEE 802.1w [7]). The main difference is in the negotiation between nodes, which has been made more efficient. RSTP shortens the time to detect changes and to send notifications via BPDUs.

Following the approach proposed in IEEE 802.1Q [8], VLANs are virtual local area network consisting in “a subset of the active topology of a bridged local area network”. Bridge knows a set of VLAN identifiers (or tags, VIDs), that extend the source and destination MAC addresses, by treating frames and addressing information for different VLANs independently. VLANs are supported by a modified version of spanning tree, the Multiple STP (MSTP 802.1s [9]) by introducing one spanning tree for each given group of operative VLANs. MSTP provides a way of allowing a common spanning tree (CIST) to exist across different VLANs within a network, while each VLAN internally uses a particular spanning tree instance.

3. Path (re-)configuration in a mesh

A mesh network can be composed by optical, wireless and wired components (see figure 1). STP-based approaches can be used for routing and loop avoidance (*layer-2 routing*), within the mesh scope. For instance, STP has been implemented and tested over WDS [10]. However a lot of work has been produced in order to find *ad hoc* solutions for particular types of network presenting internal loops. In particular, most of results concerning mesh networks can be found in the specialized literature on optical meshes, aiming at improving network performance in terms of rapid reconfiguration and reliability after a topology change. Solutions are classified into protective and restoration mechanisms, depending on the availability of a backup preplan [4,12]. Considering a group of links that can be affected by a single failure in the physical layer, J.Q. Hu [11] has shown the NP-completeness of the Diverse Routing problem (in which completely disjointed routes have to be computed for backup) and of the Least Coupled Paths problem (where the superposition of alternate paths has to be minimized). [11] also shows that an approach based on Integer Linear Programming allows to compute a new active topology in a few milliseconds over tens of nodes (but estimates do not include deployment).

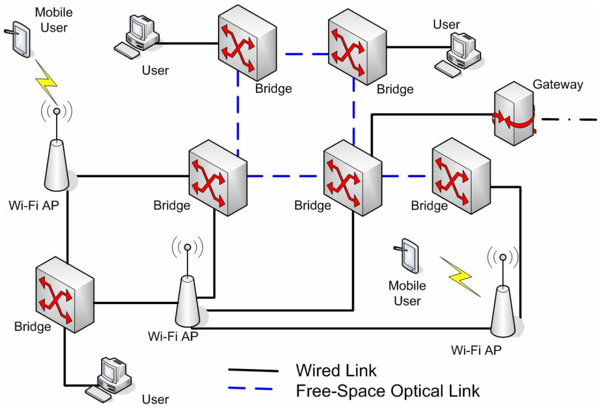

Figure 1 – multiple access technologies in a mesh

Dynamic routing methods have been developed in the frame of WDM meshes, but they can be simply extended to generic mesh networks. For instance, [13] proposes a differentiation of service-specific path quality attributes, and uses periodic transfers of path information messages to find and update multiple feasible paths. A scalable solution based on GMPLS is discussed in [14]. It proposes a distributed path computation and provisioning. It also computes shared restoration paths, and the required input information in terms of availability of link resource and sharing.

For transparent optical network (TON) a rapid recovery method, namely Flooding Based Mesh Restoration (FBMR), is discussed in [15]. This method is based on active flooding of backup traffic over a backup path. Since backup traffic forwarding is performed over pre-computed channels, FBMR does not require any additional signaling and configuration of photonic cross-connections. STREAMS, a method similar to FBMR, exploiting dedicated end-to-end recovery streams, is discussed in [16]. The problem of recovery after a dual link failure is mentioned in [17] which proposes to deconstruct the network into multiple sub-graphs. Eventually, a routing method based on a global knowledge of the network is given in [11]: the Salmon Reservation Protocol, which is an enhancement to GMPLS, provides the network control entity with global wavelength resources information, which are needed to establish traffic-engineered paths.

4. CVS: Multiple VLANs in a mesh

According to the standard [8], a switched LAN can be split up to 4096 different VLANs, which can span different subsets of the overall network. Thus a VLAN can be used in order to differentiate user groups with priorities (IEEE 802.1p/Q), differentiate the access to network paths, create virtual private networks and operate a load balancing of the overall traffic

Furthermore, VLANs are fundamental to our extent, since they can provide path redundancy in the access network. In fact, a VLAN can be thought as a group of *shared mesh restoration paths*: with the exception of trivial networks, the spanning tree is not unique, and all possible VLANs can be separated into *active* VLANs, *backup candidate* VLANs and *unconfigured* VLANs. At any given time, the network traffic is forwarded through active VLANs only, whereas no frames are tagged as belonging to backup candidate VLANs, which are ready to be used. No tags are assigned to unconfigured VLANs. Note that active and backup VLANs have to span all active edge nodes and at least one node connected to the core network, whereas there is no need for a complete spanning of the network.

A VLAN, either an active or a backup candidate, is defined in a limited scope, on a per-class-of-service basis. Furthermore, the possibly limited scope of a VLAN is not a fault; conversely, it permits the segmentation of the access network into subnets that are not simultaneously affected by the same topology change. By using VLANs, the goal of the network designer should be the maximization of the probability that when a VLAN is affected by a topology change, another VLAN is available as backup. If this were the case, the only additional requirement to make the system work properly, were the possibility to switch the frame tagging process at edge nodes. To date, this feature is implemented on many bridges and APs.

In our proposal we envisage the implementation of a network control function (NCF) that surveys the status of links in the access network and configures a limited number of redundant VLANs. The brightness of the NCF stays in the selection of VLANs to be adopted in a bouquet at a given time, and in the capability of switching the traffic from the active VLAN to a candidate, and, finally, in the selection of the optimal candidate within the *bouquet*. These operations, which are out of the scope of this paper, can be as complex as the original routing problem (indeed, this is a NP-problem like the Least Coupled Paths problem [10]), but they can be performed offline, in a predictive way.

Also MPLS techniques could be suitable for such operation, and represent an alternative to the VLAN-based solution, even though traffic engineering in MPLS requires a more complex control plane.

4.1 Issues related to topology changes

When the topology of the access mesh network changes, network control protocols react as previously described. As a consequence, users whose connection path is affected by the reconfiguration procedure, experience a service discontinuity that can last several seconds depending on the adopted protocol. Consider a switched LAN with at most 7 bridges connected in a cascade, and suppose that a topology changes occurs. If STP is in charge of the recovery procedure, it takes almost 30 seconds for STP to converge to a valid topology, and 50 seconds to restore the connectivity. If RSTP is adopted, then the recovery duration reduces to about 10 milliseconds when a preconfigured uplink backup port is available on the bridge. Otherwise, RSTP operates similarly to STP, but with shorter timeouts, and it takes about 18 seconds to converge to the new topology. However, the effectiveness of STP and RSTP decreases linearly with network diameter. Mechanisms thought for optical links can converge as fast as the switching of an optical device (a few milliseconds), but require a great amount of additional resources reserved for backup purposes. Considering additional time intervals due to deploy the updated network topology, the recovery procedure duration can be intolerable for real time streams (even if combined with playout-delay strategies [18]), or connectionless transport services (e.g., UDP), since discontinuities drive to unrecoverable transmission errors.

Following the IEEE 802 approach, during a recovery procedure, some filtering database entries are erased. In any case, when a new node is included in the new active topology, that node floods the network until a complete filtering database will be recovered. This phase extends well beyond the reconfiguration period – i.e. all users have to send at least one frame *after* the link reconfiguration is complete -. The flooding can have a negative impact since it can turn into unnecessary network congestions that slow down TCP performance and degrade the quality of UDP streams. An optimal strategies should consist in minimizing the duration of the recovery phase, and in bounding frame flooding within as less as possible network links.

4.2 Vendors' proprietary solutions

Some vendors produce devices able to rapidly reconfigure the active topology of the network. These are expensive devices, usually managed by proprietary software by which it is possible to handle device's port rules (i.e., load tagging profiles, set port status, etc.), and to create ad hoc rules for traffic flows.

To obtain a faster reconfiguration of a spanning tree of a switched LAN, Cisco System developed a method called Uplink Fast. When multiple ports can provide connections to the root bridge, Uplink Fast guarantees a rapid forwarding of the data traffic without waiting the time requested by standards. In fact, upon a link failure is detected by a node on its root port, the lowest cost port in *blocking* status, able to reach the root bridge, is automatically switched to the *forwarding* status. This dramatically decreases the convergence time of the STP in the event of the failure of an uplink.

Another example is proposed by Colubris, which developed a traffic management system for wireless communications. In particular, Colubris developed a device supporting VLANs and traffic classes over Wi-Fi: the InCharge Network Management System is a standalone WLAN management system that centralize all configuration, monitoring, fault management, and troubleshooting functions for the Colubris Intelligent MultiService System, which extends the services of a wired network infrastructure to wireless users.

4.3 The CVS operation

CVS provides a simple and ready-to-deploy method to cope with the need of multiple pre-configured backup paths, without requiring any additional bandwidth resources or complex signaling exchanges.

CVS adopts a programmable bouquet of backup candidate VLANs, managed by a unique NCF, and, in its basic form, it does not require any modification to existing switching devices. CVS only requires that edge nodes can tag frames with assigned VIDs, and all other nodes support VLANs. CVS is a software that runs on a particular node, namely the NCF , which could either reside in a network switch or in a PC. The NCF can be reactive, when the VLAN switching is performed as a consequence of a failure. Besides, the NCF can be proactive, when VLAN switching can be performed before network failures, due to the possibility to predict topology changes based on the measurement of some link performance parameters, or also based on the knowledge of deterministic node movements (e.g. nodes moving in a regular way: low orbit satellites, HAPS, or even nodes located aboard city trains or buses, tracked with GPS – see [19] for further details).

Link status sensing functions are needed not only in order to detect failures, but, mostly important, to estimate failures and link degradations in advance. In practice, wired links exhibit an ON/OFF behavior, so that failures cannot be predicted; conversely, wireless devices suffer for slow and fast fading effects, and an estimate of link performance can be obtained by monitoring the smoothed SNR and/or BER temporal behavior. This is particularly relevant for free-space optical devices used in the a mesh access network, and for Wi-Fi WDS systems covering large areas. In free-space optics, sensing functions are already adopted for automatic alignment of power beams [1]. Similarly, Wi-Fi devices are provided with a spectrum analyzer and a BER measuring functions, and the signal strength can be monitored by means of several freely-available tools, like AEROSOL (for Windows OS), KISMET (Linux), MACSTUMBLER (MacOS), and

NETCHASER (PalmOS). Thus, proactive CVS operations can be performed by commercial devices.

When sensing functions are available, CVS runs in proactive mode, and so *completely* eliminates reconfiguration periods, whereas the flooding is limited to a single active VLAN. Otherwise, without predictive tools, CVS operates in reactive mode, and reconfiguration periods are reduced to the time interval needed to detect the topology change and to communicate the new VID to be used at edge nodes.

5. Performance evaluation

A preliminary evaluation of proactive CVS performance has been carried out by means of the OPNET simulator. We assume to profit from a reliable sensing function for network internal links. A simple mesh network has been simulated, as shown in figure 2. We used common bridges endowed with STP, RSTP, VLAN facilities and IEEE 802.11g APs. Users are connected through 100 Mbps Ethernet links and 54 Mbps Wi-Fi cards. Bridge links have a capacity of 100 Mbps, but B4-B5, whose capacity is 10 Mbps.

In the first set of simulations, only fixed users are considered, and the effects of failure of link B1-B6 are evaluated. Users 1 and 2 take part to a video conference at 800 Kbps, with 1000-bytes packets transmitted with exponentially distributed inter-times. At time t=427s, the link B1-B6 fails: figures 3 to 5 depict the following flooding, normalized to the number of flooded links. By using STP (figure 3), a long lasting flooding arises after a 15 seconds delay. During that interval, the bridge B1 drops packets addressed to User 2. The analysis of network traffic (not reported here for sake of space) shows that the overall STP reconfiguration lasts almost 30 seconds, whereas RSTP slightly reduces the flooding duration, even thought the flooding phase occurs in advance. In fact, RSTP skips the dropping phase occurring in STP. Figure 4 reports results obtained with CVS, which allows to switch B1-B6 traffic (VLAN #2) to path B1-B2-B6 (VLAN #1) and it proofs that the flooding is negligible (1% of generate traffic). It is also remarkable that flooding in STP/RSTP affects all network nodes, whereas CVS flooding is bounded to VLAN #1.

In figures 5 and 6 we represent TCP performance experienced by Users 3 and 4, which are respectively an FTP client and server. FTP traffic saturates the 10 Mbps link between B4 and B5. In this case, Users 1 and 2 simulate 1.6 Mbps video conference, and a failure occurs at t =423s in link B1-B6. Flooding of UDP packets turns in a 10% reduced TCP throughput if RSTP is adopted (figure 5), whereas proactive CVS reduces the flooding and does not affect the FTP transfer (figure 6). As to UDP traffic between Users 1 and 2 (not shown here) CVS eliminates discontinuities.

The CVS solution could be used to reduce the flooding due to the Access Point handover of mobile user. Our results show that using STP, RSTP or CVS, does not change the duration of flooding, and the number of flooded packets only depends on the statistic of traffic generated by mobile users. However, CVS bounds flooding traffic in a single VLAN scope.

6. Conclusions

Mesh access networks with heterogeneous devices can be merged in a switched LAN by means of already available IEEE bridging and switching facilities, such as LLC, STP, RSTP and VLANs. In particular, VLANs seems to be able to provide shared backup paths, semantically equivalent to wavelength-path designed for optical devices, or MPLS virtual paths, without requiring additional bandwidth assignment.

We explained how CVS, by means of VLANs, allows to reduce the impact of flooding and dramatically steps down the reconfiguration time, which becomes comparable with fast switching optical devices, and avoids service disruptions.

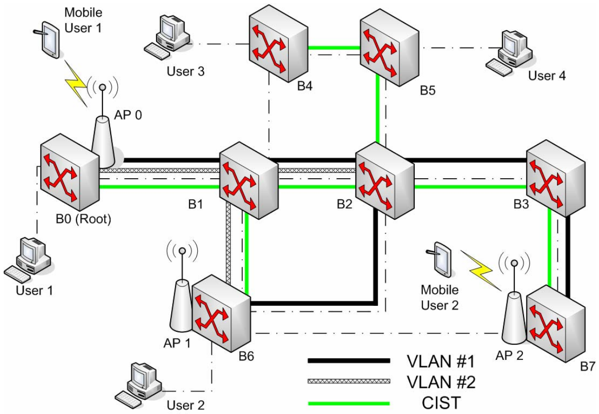

Figure 2 – Simulated network

References

[1] H. Willebrand, B.S.Ghuman, “Free Space Optics,” Sams Pubs, 1st edition, December 15, 2001.

[2] IEEE Std. 802.11 -1999: Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications

[3] ANSI/IEEE 802.1D, 1998 Edition, ANSI/IEEE Std 802.1D, 1998 Edition, “Part 3: Media Access Control (MAC) Bridges”

[4] S. Ramamurthy, L. Sahasrabuddhe, and B. Mukherjee, “Survivable wdm mesh networks”. IEEE/OSA JLT, 21(4):870–83, April 2003

[5] Y.-S. Hurh, K.-W. Shin, S.-H. Lee, J.-S. Lee, “Weather-Insensitive Optical Free-Space Communication Using Gain-Saturated Optical Fiber Amplifiers”, Journal of Lightwave Technology, Vol.23, Issue 12, Dec. 2005, Pag :4022 – 4025

[6] D. M. Britz, AT&T Labs, Research, “New Local Business Access Opportunities for AT&T using Free Space Optical Technology that Complements Existing Millimeter-Wave Radio Technology”

[7] IEEE Std 802.1w-2001, “IEEE Std for Local and metropolitan area networks - Common specifications Part 3: Media Access Control (MAC) Bridges - Amendment 2: Rapid Reconfiguration”

[8] IEEE Std 802.1q™, 2003 Edition, “802.1Q - IEEE Std for Local and metropolitan area networks - Virtual Bridged Local Area Networks”

[9] IEEE Std 802.1s™-2002 Edition, “802.1s - IEEE Std for Local and metropolitan area networks Virtual Bridged Local Area Networks - Amendment 3: Multiple Spanning Trees”

[10] M.Portoles, J.L.Valenzuela, D. Perez, O. Sallent, “Link recovery in IEEE 802.11 WLAN using WDS” IEEE VTC2004-Spring, May 2004 Pag: 2239 - 2242 Vol.4

[11] Jian Qiang Hu, “Diverse Routing in Optical Mesh Networks”, IEEE Trans. on Comm., Vol.51, No.3, Mar. 2003

[12] A. Desai, S. Milner, “Autonomous Reconfiguration in Free-Space Optical Sensor Networks”, IEEE Journal on Selected Areas in Communications, Volume 23, Issue 8, Aug. 2005 Page(s):1556 – 1563

[13] Admela Jukan, “Path Selection Methods With Multiple Constraints in Service-Guaranteed WDM Networks”, IEEE/ACM Trans. on Networking, Vol. 12, No. 1, Feb. 2004

[14] H. Liu, “Distributed Route Computation and Provisioning in Shared Mesh Optical Networks“, IEEE Journal On Selected Areas in Communications, Vol.22, No.9, Nov. 2004

[15] Sun-il Kim, “Restoration of All-Optical Mesh Networks With Path-Based Flooding”, Journal of Lightwave Technology, Vol. 21, No. 11, Nov. 2003

[16] Sun-il Kim, Steven S. Lumetta, “Capacity-Efficient Protection with Fast Recovery in Optically Transparent Mesh Networks“ Proceedings of BROADNETS’04

[17] M.T. Frederick, P. Datta, A.K. Somani, “Evaluating Dual-Failure Restorability in Mesh-Restorable WDM Optical Networks”, Proceedings of ICCCN 2004

[18] V. Mancuso, G. Bianchi, “Streaming for vehicular users via elastic proxy buffer management”, IEEE Communication Magazine, Volume: 42 , Issue: 11 , pp. 144-152, Nov. 2004

[19] V. Mancuso, G. Bianchi, N. Blefari Melazzi, U. Birnbacher, “Switched Ethernet Networking over LEO Satellite”, proc. of IEEE IWSSC’05, Siena, Italy, Sept. 2005

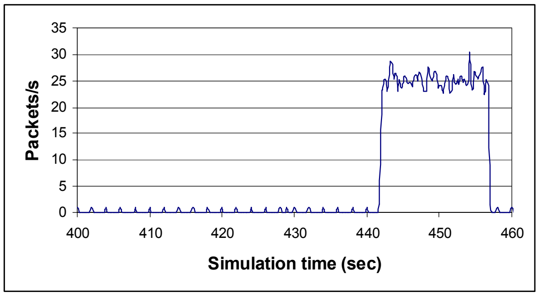

Figure 3 – Link failure: normalized flooding using STP

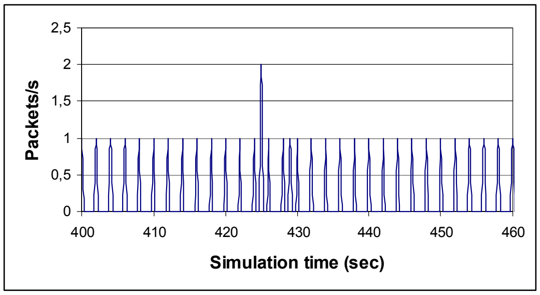

Figure 4 – Link failure: normalized flooding with CVS

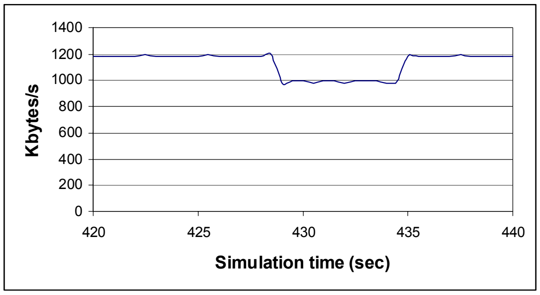

Figure 5 – Flooding effects on TCP throughput with RSTP

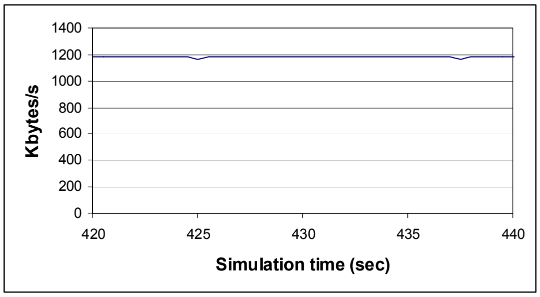

Figure 6 – Flooding effects on TCP throughput with CVS