A study of the effect of MPLS on Quality of Service in wireless LANs

Jaco Schutte and Albert Helberg, North-West University, Potchefstroom Campus

Abstract—The increasing deployment and availability of wireless technology presents unique research challenges, especially in the field of Quality of Service (QoS). The IEEE 802.11 wireless LAN (WLAN) standard is the most popular WLAN standard currently in use, but is unable to provide QoS guarantees for the growing number of wireless real-time and multimedia applications. QoS provision is made more difficult in a wireless network due to the lossy physical medium and the limited resources available. Also, wired IP QoS mechanisms such as Differentiated Services (DiffServ) do not always perform well in a wireless environment. In this study, we evaluate the use of Multi-Protocol Label Switching (MPLS) as a means of adding QoS to a wired backbone with limited bandwidth, connected to a WLAN network. We compare the performance of MPLS Traffic Engineering (TE) with that of DiffServ, and show that MPLS-TE is able to effectively share the available resources among all traffic types. The eventual goal is to implement this technique in a pure wireless network, and attempt to improve on the current performance levels of real-time traffic.

Index Terms—IEEE 802.11, Multi-Protocol Label Switching, Quality of Service, Wireless LAN

I. INTRODUCTION

The arrival of wireless communication has had a very significant impact on the world we live in. The cellular phone for example, has evolved from a luxury to a necessity for many of the world’s inhabitants. The ability to communicate without being tethered to a cable continues to attract more customers, as the worldwide increase in mobile subscribers shows. Indeed, the volume of mobile data traffic is fast becoming comparable to that of traditional fixed networks [2]. However, the growing demand for wireless data delivery has brought about new challenges for service providers. Delivering real-time applications such as VoIP, online games, and multimedia entertainment over mobile networks is a much more difficult task than is the case in a wired network. These applications require a certain level of Quality of Service (QoS) support, which is lacking in today’s IP-based networks [1], [15]-[17]. Audio and video in particular are loss-sensitive and unsuited to congested links where packets are simply dropped [6].

This state of affairs has increased the interest in QoS provision, in both fixed and mobile networks. Thus, technologies such as Integrated Services (IntServ) and Differentiated Services (DiffServ) were developed to add QoS and service differentiation to IP-based networks [16], [17]. These solutions work well enough in a wired environment. But experience has shown that the same principles used in a fixed network are not always applicable in a wireless one [1], [3]. This is mainly due to the unique problems inherent to the wireless medium, including high bit error rates, frequent link failures, and limited bandwidth [14].

IEEE 802.11 Wireless LAN is the most popular and widely deployed WLAN standard, and is currently in use in many different areas, including the business and private sector. The 802.11 is standard is a wireless version of the IEEE’s 802.3 Ethernet, and has become very popular since it is robust, flexible, and cost-effective. Unfortunately, 802.11 WLAN is incapable of providing QoS for today’s multimedia applications. [1], [3], [5], [14], [22], [26], [28].

The purpose of this study is to investigate the effect of Multi-Protocol Label Switching (MPLS) on the QoS of an 802.11 WLAN. The eventual goal is enabling 802.11 networks to perform better when running real-time applications such as voice and video.

Section II of this paper defines the concept of QoS as used in this study. Section III is an overview of the IEEE 802.11 WLAN standard and its shortcomings. Section IV looks at current IP QoS technologies. In Section V we outline our experiment, and provide the results obtained in Section VI. Section VII concludes the paper.

II. QUALITY OF SERVICE

There seems to be no single clear-cut definition for QoS. The literature gives various different versions, most of which agree in the main points [1], [6], [9], [18], [30]-[32]. To summarize, QoS is mainly defined as the ability of a network or network element to provide a certain level of assurance to the user, concerning performance and reliable data delivery. Thus, the network must satisfy a set of specific requirements concerning a particular service or data flow it is transporting. These requirements can be described qualitatively, for example “short delay” or “good quality video.” However, they are more often measured quantitatively, using numerical values [30].
The main measurable QoS parameters are bandwidth, delay, jitter and packet loss.

**Bandwidth** – Bandwidth is a measure of the capacity of the transmission link, and thus the amount of data it is capable of transporting [8]. Closely related to bandwidth is the channel throughput, which is the amount of bandwidth available to the user, measured in bits/second. Actual throughput is generally not the same as bandwidth, due to channel loss, errors and physical limitations.

**Delay** – Network delay, usually measured as one-way delay or end-to-end delay, is the time elapsed in milliseconds between a packet being transmitted, and its reception at the destination. Delay is made up of several components, including propagation, scheduling and switching delays [8].

**Jitter** – Jitter, also known as delay jitter, is the variation in delay between packets. It is also measured in milliseconds and is usually computed by measuring the difference in delay between consecutive packets [8], [23].

**Packet loss** – The presence of too many packets in a certain area of the network will cause some of them to be dropped; this is known as packet loss [9].

Different traffic types are impacted in different ways by these parameters. Non-real-time data, for example telnet and ftp applications, are much less affected by delay and jitter but are sensitive to packet loss [6]. Real-time applications, including voice and video, are severely affected by delay and jitter. For example, to satisfy high-quality Voice-over-IP (VoIP) QoS requirements, a network must have less than 0.25 % loss, a maximum jitter of 5 ms and no more than 150 ms delay time [8], [23].

In the next section, we discuss the IEEE 802.11 WLAN standard and show why it is unsuitable for QoS, especially when transmitting real-time data.

### III. IEEE 802.11 WLAN

#### A. The IEEE 802.11 standard

The IEEE started the development of the 802.11 WLAN standard in the early 1990’s. The initial version was passed in June 1997, and operated in the 2.4 GHz band using Frequency Hopping Spread Spectrum (FHSS) or Direct Sequence Spread Spectrum (DSSS) physical layer technology. Speeds of 1 and 2 Mbps were possible. In 1999, the IEEE made amendments to the standard, adding 802.11b which is capable of 11 Mbps using DSSS at 2.4 GHz [6], [22], [29]. Further additions made since then include 802.11a and 802.11g, both with data rates of 54 Mbps using Orthogonal Frequency Division Multiplexing (OFDM). 802.11a operates in the 5 GHz band [1], [6].

According to the Open System Interconnect (OSI) reference model, the 802.11 standard is placed at the physical layer and the Medium Access Control (MAC) sub-layer [1], [19], [22].

(For a detailed description of 802.11 MAC operation, see Refs [1], [3], [5], [19], [21]). There are two modes of operation available in 802.11 WLAN. **Ad hoc mode** is distributed, and the wireless stations (STAs) communicate on a peer-to-peer basis. **Infrastructure mode** consists of a number of STAs connected to an Access Point (AP), which provides centralized control. This is known as a Basic Service Set (BSS). The AP can also be used to connect the BSS to a larger network backbone.

Two different coordination functions are used to control access to the wireless medium, namely the Distributed Coordination Function (DCF) and the Point Coordination Function (PCF). DCF is the default mechanism and is based on Carrier Sense Multiple Access (CSMA), as used in Ethernet. All the STAs contend equally for the medium, and no STA is allowed to transmit while the medium is busy. If however two STAs do transmit simultaneously, a collision occurs. Unlike in a wired network, collisions cannot be detected and thus a collision avoidance (CA) scheme is included, forcing each STA to wait for a random back-off time before transmitting, even if the medium is free. After sending a frame, the STA waits for an acknowledgement (ACK) packet from the receiver. If no ACK is returned within a given time, a collision is assumed and the frame is resent.

PCF is only available in infrastructure mode, and adds a contention-free mode of access. The AP polls each STA on its polling list, allowing it to access the medium without competing with other STAs. PCF was introduced to support real-time services, and can differentiate between different traffic types. Accordingly, certain STAs can be granted priority access to the medium.

Unfortunately, even with the use of PCF, 802.11 WLAN is incapable of providing adequate QoS support for time and delay-sensitive applications, as we will now discuss.

#### B. QoS shortcomings of 802.11 WLAN

As mentioned, guaranteeing QoS in a wireless environment is a very challenging task. For several reasons, the legacy 802.11 standard cannot provide QoS guarantees for multimedia and real-time applications. The main problems are the following:

1) **Lossy and error-prone physical layer:** Wireless links suffer from interference, have a high packet loss rate, low available bandwidth, large packet delays, and a high bit error rate, as much as three orders of magnitude greater than a wired link. These factors greatly reduce the ability of the network to give consistent and reliable data delivery [1], [4], [6], [18], [25], [30].

2) **Mobility:** While node mobility is the greatest asset of a wireless connection, it unfortunately also causes difficulties. The link characteristics constantly change as users move about, altering the network topology. This may cause path breaks, and further increase delays while the connection is re-established. These unpredictable delays in the link layer add to the problem of jitter [1], [18], [25], [30].

3) **Lack of service differentiation:** The basic DCF access method is a best-effort service and provides no QoS guarantees. All STAs are given the same priority when competing for access to the radio medium. Even though PCF was intended to provide for real-time services, the polling scheme adds complexity, and reduces network performance under high load. Bandwidth is also wasted when using PCF, since all communications have to go through the AP [1], [6].
IV. PROVIDING IP QoS

IP networks currently provide only a best-effort service, even though the protocol can support different service classes. This means that IP networks have high delay, jitter, and unreliable packet delivery, with the Internet being the prime example [7]. To remedy this, the IETF (Internet Engineering Task Force) developed and standardized several techniques for QoS provision in IP networks.

A. IntServ

IntServ emerged as a QoS approach in 1994. It was designed to provide fine-grained QoS on a per-flow basis, by using resource reservation [1], [7], [16], [31]. A node that has data to transmit sends a reservation request to the receiver to determine if the required bandwidth is available. The Resource Reservation Protocol (RSVP) is then used to reserve resources at each node along the intended route [15]. Two service classes were introduced, namely Guaranteed Services (GS) for real-time traffic requiring assured bandwidth, and Controlled Load Services (CLS) for applications capable of tolerating some loss. All other traffic is treated as Best Effort. Unfortunately, resource reservation at each node adds a lot of information and complexity to the network. This has a negative effect on scalability and has severely limited the deployment of IntServ [1], [15], [32].

B. DiffServ

DiffServ was introduced in 1998 in an attempt to address the shortcomings of IntServ. DiffServ offers coarser QoS, but is simpler and thus more scalable than IntServ [1], [31]. The routers at the edge of a DiffServ domain classify all the incoming traffic. Packets requiring the same level of QoS form a behaviour aggregate (BA). Each packet is then marked with a DiffServ Code Point (DSCP) and forwarded to the next node. The DSCP is a reference to a Per-Hop Behaviour (PHB) which determines the queuing and scheduling of the packet at every hop. The core routers in the network then only need to check the DSCP value to determine how each individual packet should be treated. Currently, DiffServ has two PHB classes besides Best Effort [7], [8], [12], [15], [17]. Expedited Forwarding (EF) is used for applications requiring low delay, low jitter, and low packet loss. Assured Forwarding (AF) is for traffic with higher-than-best-effort requirements, but doesn’t provide service guarantees. While DiffServ does add service differentiation and class-based treatment, it does not consider the forwarding route and as such cannot guarantee bandwidth [12], [32].

C. MPLS

The IETF developed MPLS as an advanced forwarding technique. MPLS evolved from Cisco’s Tag Switching and was originally intended to improve router performance, but is now being used for traffic engineering (TE), and as a QoS-enabling technology. TE is the process of controlling traffic flow through a network. It can be used for example to route packets away from congested areas to avoid delay and packet loss [10], [15]. MPLS supports both the switching and routing functions of layer 2 and 3.

At the ingress point of the MPLS domain, the edge router (known as a Label Edge Router or LER) attaches a short fixed-length label to each packet, ahead of the IP header. The label assigns the packet to a particular Forward Equivalence Class (FEC). An FEC is a packet grouping that is based in information contained in the packet, for example the packet destination. All packets belonging to the same FEC follow the same path through the network. Forwarding is then done by using a label-swapping algorithm to determine the path of the packet through the network. As the packet is forwarded, each intermediate MPLS-enabled router (Label Switch Router or LSR) checks the incoming label to determine the next hop and also the value of the new label. The network-layer header is ignored, which reduces both overhead and the size of the routing table and improves scalability. The label is then discarded at the egress point from the MPLS network [2], [8]-[13], [15]-[17], [32].

The path a packet follows through the MPLS domain is called a Label Switched Path (LSP). LSPs are set up using IP links state routing protocols such as OSPF or IS-IS to obtain information about available resources in the network. Routing decisions can then be based on the availability of resources along a path, and in this way MPLS can provide guaranteed bandwidth for aggregated traffic flows. MPLS does not, however, differentiate between traffic classes [11], [32].

D. Combining MPLS and DiffServ

As seen above, MPLS and DiffServ are both suitable candidates for providing IP QoS, but both have shortcomings. DiffServ can provide class-based QoS treatment, but cannot influence the route of a packet through the network. MPLS is able to set up paths with guaranteed resources, but does not differentiate between flows.

The solution is a combination of DiffServ and MPLS [2], [8], [10], [11], [17], [32]. By mapping the DiffServ IP header to the MPLS header, the DiffServ Classes of Service (CoS) can be used in the MPLS domain. This enables MPLS routers to give certain packets preferential treatment in terms of scheduling and routing.

V. METHOD

The previous sections have shown the need for adding QoS to IP networks. We have also seen that this is a particularly difficult prospect in a wireless network, where network resources are limited by the radio medium.

IEEE 802.11b was chosen as the subject of this study, since its rapid commercial acceptance has made it one of the definitive and most widely-deployed wireless LAN technologies [33]. Many schemes have been proposed to improve the QoS performance of 802.11 WLAN, especially at the MAC layer [1], [3], [14], [18]. For the sake of brevity, these will not be discussed here. Instead, our approach is to investigate the suitability of current IP network QoS mechanisms for use in the WLAN environment. We pay particular attention to MPLS, as its label-based forwarding reduces overhead, allowing optimal use of the limited resources in a wireless network [2], [11].
In an ad hoc environment, each wireless node would also double as a router, forwarding data to the other nodes [18], [19]. For this study, we adopt the infrastructure approach, with dedicated routers forming a backbone between two wireless networks.

Fig 1 shows the simulation setup, as implemented in the OPNET Modeler network simulation package [34]. The wireless component of the network consists of two 802.11b Basic Service Sets, labelled BSS 0 and BSS 1. BSS 2 contains only one node, a WLAN server running file transfer protocol (ftp), to act as background traffic. Each BSS is connected to the wired backbone via a wireless router that also serves as an access point. Each AP connects to the backbone via a 10 Mbps 10BaseT Ethernet link, providing approximately the same amount of raw bandwidth as an 802.11b link. The five backbone routers are connected by DS1 duplex point-to-point links, each providing only 1.5 Mbps. Thus the available resources in the core area are severely limited, as would be the case in a pure wireless network. There is no mobility in the network.

Of the 11 wireless stations in each of BSS 0 and 1, four run the background applications provided by the server. Average ftp file size is approximately 100 kilobytes. Four stations in each BSS transmit GSM-quality voice to each other through the backbone on a peer-to-peer basis, giving a total voice throughput of about 26000 bytes/second. The remaining three stations have a peer-to-peer video conferencing connection, also through the backbone. Total video throughput is 273000 bytes/second. All transmissions start at 100 seconds, and the total simulation time is 300 seconds. The following scenarios are simulated:

1) Baseline, no QoS: This scenario provides the basis for comparison. No QoS methods are implemented, and the routers handle all traffic on a First In, First Out (FIFO) basis.
2) DiffServ QoS: DiffServ is set up in the backbone routers. The Video traffic receives EF treatment, the voice traffic AF, and the background traffic only Best Effort.
3) MPLS TE: LSPs are created to give the real-time traffic a dedicated route. Voice and video traffic are each mapped to their own FEC, which associates them with an LSP through the backbone, while the background applications are still routed on a Best-Effort basis.

For each scenario, the delay, throughput, and jitter (where applicable) of each traffic type is recorded. In the next section we will examine the results.

VI. RESULTS AND DISCUSSION

Let us look at how the traffic types fared in each scenario:

1) Background traffic: Fig 2 shows the average throughput achieved by each scenario for the background applications. The baseline and MPLS networks show similar throughputs, while the DiffServ-enabled simulation has a dramatically lower figure. Looking at the download response times, the MPLS simulation performs well again, while the baseline network has very long download delays, averaging in excess of 60 seconds. The DiffServ network also has low delays, but this is due to the much-reduced traffic volume as compared to the other two.

2) Voice traffic: As seen in Fig 3, the MPLS network achieves the best throughput of voice traffic. DiffServ performs even worse than the baseline network, even though the voice traffic was marked for Assured Forwarding. The DiffServ network also has the highest jitter value of the three, with MPLS performing the best in this department. The end-to-end delay values for all three scenarios are comparable at around 1 second. This clearly shows that DiffServ cannot always provide QoS guarantees through queuing and scheduling alone. MPLS shows better performance by providing a dedicated path for the voice traffic, thus avoiding congestion and unwanted delays.

3) Video traffic: The throughputs of video traffic for the respective scenarios are shown in Fig 4. Clearly, DiffServ easily outperforms the other mechanisms. It also has the lowest delay and jitter by some margin, although this comes at the expense of the other traffic types, as we have seen. The MPLS network has the lowest throughput and the highest delay variance, but has slightly lower delays than the baseline network. It does however, provide more fairly for the other traffic types, even at the cost of lower video performance.
The mean results from all three scenarios as calculated by the OPNET simulation are summarised in Table 1. We can see from the results that each approach has its advantages and disadvantages:

The baseline network gives all traffic equal importance, and thus there are no guarantees, resulting in unreliable service in a congested network. Video traffic, with the highest network load of the three, is especially poorly served. All traffic types are subject to the same queuing delays, resulting in some real-time frames being dropped while ftp traffic is served. This forces retransmissions, which further increases network delay and congestion, as is evident from the extremely long delay times experienced by ftp traffic. The baseline scenario also has higher jitter values than any of the others, as a result of unreliable packet delivery.

DiffServ provides very good video performance with Expedited Forwarding treatment, but this has a hugely negative impact on the voice and background traffic. While marking video traffic as EF guarantees low delays and reliable delivery, the performance of the other two traffic types suffers unacceptable degrading as a result. Voice throughput is roughly one-fifth that of the other scenarios, even though voice traffic receives Assured Forwarding treatment from the routers. The lack of performance of the best-effort ftp traffic is even more pronounced, with ftp throughput continually declining as the simulation progresses. In summary, we see that using DiffServ service classes in a limited-resource environment results in the high-priority traffic using almost all available bandwidth, leaving insufficient capacity to service other traffic types.

MPLS provides a greater degree of fairness by using dedicated paths and routing traffic away from congested areas. The only negative aspect is the lack of video performance compared to the other scenarios. Even though the video traffic was given a guaranteed route, the available resources were clearly not enough to guarantee QoS along that path. The lack of service differentiation in the MPLS domain compounded this problem. In contrast, however, the less resource-hungry applications performed much better than in the DiffServ or baseline scenarios. The LSP provided for voice traffic had sufficient resources to guarantee reliable voice delivery. Also, routing the real-time traffic away from certain routers freed more bandwidth for ftp traffic. As a result the MPLS scenario shows the best ftp performance in terms of throughput and download delay times.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we have attempted to show the suitability of MPLS for use in networks with limited resources. Our simulation results show that MPLS is able to provide QoS more fairly to different traffic types than DiffServ. DiffServ tends to give high priority data too much bandwidth and then neglect the other traffic. MPLS traffic engineering enables us to route different types of traffic onto dedicated paths to improve QoS.
A further development to the current simulation setup would be to combine MPLS and DiffServ, as discussed in Section IV. This would ensure guaranteed QoS by using traffic engineering and dedicated paths together with different queuing and scheduling mechanisms for each type of traffic.

We have not yet been able to test this technique in a purely wireless network due to limitations in the simulation software. However, our approach shows promise when applied to a network with limited available bandwidth, compared to current wired network QoS mechanisms such as DiffServ.

The next important step is then to implement this technique and evaluate its performance in a pure wireless environment as soon as the software allows. In this way, we will attempt to achieve the stated objective of improved real-time performance by maximising the effective usage of the resources available in a wireless environment.

REFERENCES


Jaco Schutte obtained his B. Eng degree (cum laude) in computer and electronic engineering from the North-West University in 2004. He is currently enrolled for his M.Eng degree at the same university.