Implementation and Evaluation of a TDMA MAC for WiFi-based Rural Mesh Networks

Ashutosh Dhekne Dept. of CSE, IIT Bombay ashudhekne@gmail.com Nirav Uchat Dept. of CSE, IIT Bombay nirav.uchat@gmail.com Bhaskaran Raman Dept. of CSE, IIT Bombay br@cse.iitb.ac.in

ABSTRACT

WiFi mesh networks with outdoor links have become an attractive option to provide cost-effective broadband connectivity to rural areas, especially in developing regions. It is well understood that a TDMA-based approach is necessary to provide good performance over such networks. While preliminary prototypes of TDMA-based MAC protocols have been developed, there is no implementation-based validation/evaluation in multi-hop settings. In this work, we describe the elements of a multi-hop MAC implementation based on the open-source MADWIFI driver. We also present an evaluation, with a detailed accounting of the various overheads, on a 4-hop (5-node) path. We show that our implementation has no system overheads, achieves good throughput, and low delay/jitter.

1. INTRODUCTION

WiFi-based outdoor mesh networks have been demonstrated to be a viable option for providing cost-effective broadband connectivity to rural regions [1, 2, 3]. It is also well known that a TDMA-based MAC is necessary for good performance in such networks [1, 2, 4], with the default CSMA/CA operation being inefficient.

Although the significance of TDMA is recognized, its practicality has always been in question, especially in multihop settings. Can effective time-synchronization really be achieved? What would be the overheads in practice, with multiple hops? Can an implementation work using low-cost off-the-shelf hardware? Can system overheads be controlled so that the wireless capacity can be maximized? These are significant questions can only be answered through a prototype. However, prototype-based evaluations of multi-hop wireless TDMA systems have been few and far between.

In this paper, we demonstrate a TDMA implementation that can be used over outdoor multi-hop networks using off the shelf inexpensive hardware. Our implementation is based on the open-source MADWIFI driver. It includes multi-hop synchronization, and schedule dissemination from a central node. We carefully account for the various overheads in our implementation, such as the synchronization

Copyright 2009 ACM X-XXXXX-XX-X/XX/XX ...\$10.00.

error, guard time, header overheads, etc. We show through evaluation that these overheads are noticeable but tolerable; the multi-hop throughput is close to what we expect theoretically. And the delay/jitter values are small even over multiple hops and good enough to support real time voice and video-conferencing applications; as noted in [4], providing support for real-time video conferencing is important for applications such as e-learning in rural regions.

2. RELATED WORK

There has been considerable effort in the area of driverlevel radio configuration using open source drivers which facilitates implementing various protocols over inexpensive offthe-shelf WiFi hardware. We present some of these which have demonstrated TDMA implementations.

SoftMAC [5] explores the use of the MADWIFI driver for Atheros-based WiFi radios to experiment with MAC protocols. It disables RTS/CTS, MAC level ACKs, and facilitates custom packet header formats by setting the card in monitor mode. To demonstrate the utility of the platform, [5] has implemented a TDMA system between two nodes. For our work, we borrow from SoftMAC, insights about disabling certain aspects of CSMA.

MadMAC [6] also implements an example TDMA system between two machines with slot sizes of 20ms-60ms and guard bands of 4ms-12ms. However, since we envision a multihop system, increased slot size has a detrimental effect on the delay/jitter. We have used slot sizes and guard bands much smaller than those in MadMAC.

Building over SoftMAC, FreeMAC [7] exposes many more configurable parameters. It also demonstrates a TDMA system, but it uses out-of-band ethernet for synchronization. FreeMAC also implements channel switching in the TDMA system but does not implement multiple hops. FreeMAC uses the hardware beacon timer and indicates that the timer works well under both low load and heavy load conditions. However, we found that the hardware timer is very sloppy with an increased number of RX interrupts as described in Sec. 4.5. FreeMAC's insights into various aspects of Mad-Wifi, including the hardware timer, has served as a starting point for our work.

Overlay MAC [8] uses the Click router system and implements a configurable module between the MAC layer and the network layer. However, unlike our work it does not have precise control over packet transmission times. And it implements a distributed algorithm for allocating slots, while we have centralized schedule computation and multihop dissemination.

The 2P protocol [1] and its time-synchronized implementation [2] also involve TDMA prototypes. Unlike this paper,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

[1, 2] are specific to directional long-distance links only, and do not apply to generic mesh network links. Furthermore, [1, 2] have only evaluated their prototypes on single-hop (one central node with two other nodes on each side) settings.

The work in [9] implements a multi-hop TDM MAC protocol. However, the focus in [9] is on time synchronization, and most of the measurements in [9] are on a single-hop (Sec. 8.1, 8.2, 8.3.1). Our work includes a multi-hop schedule dissemination mechanism, and we carefully account for various throughput overheads. Another difference is that we have used an off-the-shelf Atheros-based platform, while [9] has used a custom-built platform for implementation.

To summarize, our contribution is the implementation and performance characterization of a multi-hop TDMA-based MAC for WiFi mesh networks. This implementation is set in the context of the FRACTEL architecture [4], where we have a mesh network with both long-distance as well as localaccess (medium-distance) links.

3. BACKGROUND: TDMA APPROACH

We have argued in [4] for a TDMA MAC which is connectionoriented and centralized. We have designed such a MAC protocol. While it is beyond the scope of this paper to present the protocol itself in detail, we now give a brief description to set the context for the implementation (Sec. 4) and performance evaluation (Sec. 5).

Time-slots, frame structure: A time-slot is a unit of resource allocation. In our centralized scheme, a central node decides who should transmit in what slot, to whom, and using which channel. There are three kinds of time-slots. (1) Control slots: used to convey information from the central node toward the other nodes. (2) Contention slots: used to convey information from the central node. (3) Data slots: used for the actual data flows. A frame is a repeating pattern of control, contention, and data slots. The frame structure is shown in Fig. 1.



Figure 1: Components of a Frame

Tree topology: Control information (in the control slots) always goes down along a tree rooted at the central node. Contention slots are always used to convey information from child-to-parent (toward root) in this tree. The tree topology is used only for control and contention slots, and not for data slots.

Control slots: The control slots convey three important pieces of information: (a) time synchronization, (b) the tree topology itself, and (c) the data schedule.

Time synchronization: To handle clock drift, synchronization is done in every control packet using a hardware timestamp. The synchronization propagates down the tree, one hop at a time, to all nodes in the network. Each schedule packet contains its own offset from the global time and the exact global time of the beginning of this slot. Together, these three entities enable the receiving node to synchronize to the current global time and also calculate the next slot time.

The *routing tree* information is just an array of parentchild relationships. Each non-root node must appear at least once as a child node in this array. Figure 2 shows an example



Figure 2: Routing tree info in a control packet

topology and its routing tree. Such centralized routing gives the root node complete control over bandwidth usage and potential QoS guarantees.

The number of control slots in each frame is constant, but they are repeatedly numbered from 0 (shown as R in Fig. 1) to n-1 where n is the number of nodes in the network. This cyclic numbering can span over many frames: in Fig. 1, given three control slots in each frame, and five nodes in the topology, the first three nodes are allocated control slots in the first frame and the remaining two are allocated control slots in the consecutive frame. After the second slot in the second frame, the schedule transmission opportunity rotates again to the root node.

Control packets are constructed in the driver by the root node. It consists of a control header, a (possibly zero) number of scheduling elements which constitute the data schedule, and (possibly null) routing tree information. The control header has synchronization information, the number of data scheduling elements, and the number of routing tree elements contained in the packet. Each *data scheduling element* contains the transmitter, receiver and flowid for a data slot. All scheduling elements together describe the path of all data flows in the network.

Contention slots: These have two important functions. (a) New nodes join the network by first listening to the control slots, thus getting synchronized with the network, and then sending a node-join request toward the root node using contention slots. (b) Similarly, a node which wants to start a new data flow conveys this request to the root using the contention slots.

Data packets: These are attached a header that help in routing the packets. In addition to the next hop and end-toend source and destination fields, it also has the flowid field, to enable per-flow scheduling.

4. TDMA MAC IMPLEMENTATION

We have implemented many of the features of the above TDMA MAC; we describe our implementation in this section. We have used Mikrotik [10] single board computers with Ubiquity SR2 cards with Atheros AR5212 chipset. The boards run OpenWRT [11] Kamekazi 8.09 and we have modified the madwifi [12] driver extensively to match our requirements. We set the madwifi driver to run in monitor mode, which allows us to receive all packets seen on air. It automatically disables MAC level ACKs and RTS/CTS [5]. It also allows sending RAW packets without attaching the 802.11 header. We now describe the implementation of a number of other features required by our MAC protocol.

4.1 Generating RAW packets

We do not use the standard 802.11 [13] frame structure while sending control or data packets. Data packets arrive to the MAC layer from the network layer and are attached with a custom data header. Control packets, on the other hand, are generated in the madwifi driver itself. We have written a function similar to ieee80211_send_qosnulldata() that uses ieee80211_getmgtframe to allocate skb, fills in the schedule data-structures and then sends the packet on air by calling the ath_startraw() function. Both the control header and the data header contain 0xFF as the first byte so that all receivers can distinguish valid IEEE802.11 packets from ours.

4.2 NAV and Sequence Number Fields

The 802.11 frame structure contains a NAV field that causes other nodes to backoff while the current transmission is under way. We do not need this field since we have assigned slots for each node's packet transmissions. However, the value of the NAV field is used by the receiving devices to perform virtual carrier sensing and backoff their own transmissions in hardware. Since our headers replace the standard 802.11 header, the value present in this field must be zeroed to prevent other devices from unnecessary backoff. Also, the hardware stamps a sequence number at byte 22 and 23: the driver does not have control over what value will be written in this field. Since our custom header exceeds 23 bytes, stamping of the sequence number corrupts the header. As a work-around, we set the RETRY flag to suppress the stamping of the sequence number field by the hardware, as suggested in [5].

4.3 Synchronization

Each control packet has a hardware timestamp, for hopby-hop synchronization of all nodes. This timestamping, if done in software, is inaccurate, because we can not be sure when the packet will leave the hardware. In regular 802.11, a similar requirement is present for beacon frames: they are timestamped with a 64bit microsecond granularity value by the hardware at bytes 24 to 31. The Atheros hardware can be made to timestamp any packet by setting the type flag to HAL_PKT_TYPE_BEACON in the call to the ath_setup_txdesc() function. All control packets are sent with this flag set and the hardware timestamp is used by the receiving node to accurately synchronize with the global time. In addition to the hardware timestamp (tx_hw_ts), the schedule also contains the sender's offset (tx_offset) with the global time and the global time when the current control slot ideally started (slot_start). Using this information along with the receive timestamp (rx_ts) and the slot interval (slot_interval) each node calculates its offset from the global time and the time of the next slot as per equation 1 and 2 respectively.

$$rx_offset = rx_ts - (tx_hw_ts - tx_offset)$$
(1)

 $next_slot_time = slot_start + slot_interval + rx_offset$ (2)

4.4 TDMA queueing at MAC layer

All packets arriving at the madwifi driver enter through the ath_hardstart() function. Depending on whether the device is in the monitor mode or not, the ath_hardstart() function sends the packet to the ath_tx_startraw() or ath_tx_st function respectively. Since we need precise control over packet transmission times, we buffer all incoming packets in a software queue instead of allowing them to flow through the ath_hardstart() function. During the TDMA transmission opportunity, packets are dequeued and handed over to the ath_hardstart() function which attaches the data header and sends the packet on air. The number of packets sent during a transmission opportunity is equal to the lesser

of the number of packets that can be transmitted in the slot interval and the number of packets present in the buffer.

On the receiver side, an arriving packet will have one of the following three destinies. It may either be intended for consumption by the receiving node, or may be required to be forwarded to another node (this node is a relay) or may have nothing to do with this node, in which case, it must be dropped. Specifically, a packet is thought to be destined to a node if its ID appears in the end_destination field or if both the next_hop_dest and the end_destination fields have broadcast address. In such a case, the packet is sent to the network layer for its consumption. If the node's ID appears in the next_hop_dest field, but not in the end_destination field, the packet must be forwarded to another node. Such packets are enqueued in the TDMA buffer. All other packets are dropped.

4.5 Timers

Timers are an integral part of our TDMA implementation since they maintain the slotting structure at each node. The Linux kernel provides *software timers* that allow a minimum granularity of 1ms (we are not using a real-time kernel). The Atheros chipset AR5212 uses a *one-shot hardware timer* and another *periodic hardware timer* for sending out periodic beacons [14]. The one-shot timer has a granularity of $128\mu s$ and the periodic timer has a granularity of 1ms.

Our decision of whether to use the hardware or the software timer depends on which timer provides us the best precision and accuracy at both low and high load conditions. In order to characterize the timers, we subjected a node to lowload conditions and then to high-load (load in terms of the number of RX interrupts). Fig. 3 plots the CDF of the measured timer accuracy. The x-axis shows difference between when the timer should ideally trigger and when it actually triggered; i.e., the *accuracy* of the timer. The closer this value is to zero, the better. We see that the hardware timer is very precise with RX interrupts disabled, but performs poorly when RX interrupts are enabled.

The software timer on the other hand performs well in both high load and low load cases. The variation of the software timer is under 5μ s. Also, the accuracy of these timers is better than the hardware timer. Hence we chose the 1ms software timer for our implementation, and not the hardware timer.

In our implementation, we do not require a node to cause an interrupt at each slot. The interrupt is necessary only when it is the node's transmission slot. This optimization radically reduces the load on the timer system and improves the robustness of our implementation.

4.6 Clock Drift

Another important aspect in time synchronization is clock drift. We have observed clock drift to vary between different pairs of hardware from a few microseconds to about 25μ s per second as shown in Table 1. The clock drift must be accommodated in the guard band in addition to the timer inaccuracies.

5. EXPERIMENTATION

Through various experiments, we seek to answer the following questions. (1) What is the impact of changing the number of hops on UDP and TCP throughput? (2) What is the impact of changing the slot size on UDP and TCP throughput? (3) What is the impact of the number of hops and the slot size on the delay/jitter of packets?



Figure 3: Performance of software and hardware timer.

Table 1: Clock drift for different pairs of cards in μ s/s. Negative values indicate that the difference between the offsets keeps decreasing.

Card Pair	1-5	1-6	1-3	6-3	5-3
Average	-2.78	9.72	4.07	-1.41	13.24
Std Dev	0.42	0.47	1.03	0.5	1.51

All the experiments done here are in an interference free 802.11a frequency (channel 160). These experiments were conducted in an indoor setting. We believe that the results would extend to interference free outdoor links too, although only actual experiments can confirm this. For long distance links, the large propagation delay (e.g. $\sim 83\mu$ s in a 25km link) also has to be accounted for in the guard time; but we have not included this overhead in our current experimentation.

5.1 Experimental Setup

In order to answer the above questions, we conducted a number of experiments on a linear topology consisting of up to five wireless nodes as shown in Fig. 4. Admittedly, this is a simplified setup as compared to a full fledged mesh. But this setup helps us understand various performance aspects and gain confidence in the TDMA implementation.



Figure 4: Linear topology used in our experiments.

One node is designated as the root node and generates control packets. The routing tree contains information about the linear topology. Each node is numbered starting at 0 from the root node downwards in the topology. Each node sends packets when (slot number) *modulo* (number of nodes) matches its own node id. The contention slots are unused. All data packets are destined either to the root node or the leaf node, and static routing entries facilitate routing of data. The number of control, contention, and data slots, and the slot interval are all configurable in the user space through a /proc entry. We have used a 100-slot frame: 3 control, 5 contention and 92 data slots in this setup and the slot interval is varied as described in individual experiments. There is a PC at either end of the 4-hop wireless topology, and UDP/TCP throughput are calculated by using the *iperf* tool between the two PCs.

5.2 Expected Throughput

All nodes are set to transmit at 54Mbps and can transmit only in their own transmission slots. With the configuration described in Sec. 5.1, with five nodes transmitting, we use 87 of the 92 available data slots¹ in a round-robin fashion, so that each node gets 87/5 = 17.4 slots per frame. The number of packets sent in each slot depends on the slot size and the size of the packet. Table 2 shows the transmit time for the various parts of a packet at 54Mbps. Equation 3 calculates the theoretical throughput for the 4-hop case with the slot size of 2ms and a $100\mu s$ guard band giving 87 slots per second to each node. Similar calculations can be performed to derive the theoretical maximum throughput for any number of hops and for any slot size.

Table 2: Time taken to transmit various portions of the packet at 54Mbps

Description	Bytes	Time (μs)
UDP Payload	1470	217.77
UDP Header	8	1.185
IP Header	20	2.962
Ethernet Header	14	2.074
CRC Trailer	4	0.592
Fractel Data Header	32	4.740
PLCP Header	-	20.444
Total	-	249.767

 $Slot tx time = 1900\mu s (2000 - 100\mu s guard time)$ $Packets/slot = \lfloor 1900/249.767 \rfloor = \lfloor 7.607 \rfloor = 7$ $Packets/sec = (\#of slots/sec) * (packets/slot) \qquad (3)$ = 87 * 7 = 609

 $Throughput = 609 * 1470 * 8/10^{6} = 7.16 Mbps$

5.3 Number of Hops and Throughput

We conducted experiments for TCP and UDP throughput on linear topology with varying number of hops from 1 to 4 and slot size 2ms. The results are shown in Figure 5. The variability across different measurements was small, and hence we do not report it in the plot. We can see that UDP throughput decreases with increasing number of hops, since the number of slots per node per frame decreases with increasing number of nodes. TCP throughput decreases much faster than UDP throughput because an increase in number of hops means an increase in end-to-end error probability and also an increase in the round trip delay. Moreover, since we have disabled per link retransmissions, TCP throughput suffers drastically. When we use multiple TCP connections, depicted in the graph as TCP-2 and TCP-4, for two and four TCP connections respectively, the total available bandwidth is shared between them. Thus the cumulative throughput approaches that shown by UDP.

5.4 Slot size and Throughput

¹The last x data slots in a frame are not used so that the control slot timer for the next frame is triggered precisely. x is equal to the number of nodes in the network.



Figure 5: UDP throughput decreases with increasing number of hops. TCP throughput decreases much faster due to no link-link retransmissions.

The slot size variations should ideally have no impact on the UDP throughput. However, since we do not fragment packets at the MAC layer, an increase in slot size causes lesser overhead. Also, we implement a small guard band of $100\mu s$ for every slot. Thus UDP performs better with increasing slot size as seen in Figure 6. However, TCP throughput is adversely affected by a large slot size since it experiences a higher delay in receiving ACKs; and this effect is magnified due to the absence of link-level retransmissions. Since TCP never fully uses the available bandwidth even in smaller slot sizes, the reduced overheads in larger slot sizes do not benefit TCP throughput much. Since a single TCP connection does not fully utilize the available bandwidth, we experimented with multiple connections. We found an increase in the cumulative throughput, as shown by 4-TCPs-4-Hops readings in Figure 6.



Figure 6: UDP throughput increases with increasing the slot size and decreases with increasing number of hops.

5.5 Delay Characteristics

In our experimental setup described in the Sec. 5.1, a packet sent by the root node, will be transmitted over consecutive data slots to reach the leaf node. Since the data slots are numbered from 0 to 4 and then the numbering is restarted from 0, a packet sent from the leaf node to the root node, will be transmitted by the intermediate nodes only when their transmission turn occurs. The round trip of a ping packet is depicted in the Figure 7 and formalized in Equation 4, where x is the number of hops. We note that the equation shown here is a function of the way we have numbered the slots (i.e. the scheduling).

Best case
$$RTT = x + ((x - 1) * (x - 2))$$
 (4)



Figure 7: The best case round trip of a ping packet.

The best case round trip time for the 4-hop network is 85ms as calculated from Equation 4; we also observed the same value in our experiments.

In addition to the delay, the packet delay variation (jitter) is also an important metric for good quality of service for audio and video communication. A jitter below 100ms is generally believed to be good enough for such applications. In our experimental setup, we observed the jitter to be very low: about 2.5ms over the 4-hop path, with a slot size of 2ms (the jitter was more or less independent of the slot size).

5.6 Implications of Results

We have observed UDP throughput to be very close to the theoretically predicted value. Even though TCP throughput is quite low for multiple hops, multiple TCP connections together can provide good performance as evident from Figure 6. What these mean is that the implementation of the TDMA MAC does not have any system bottlenecks, even when the radio is operating at 54Mbps. The low delay and jitter are encouraging for real-time applications. During the testing of our prototype, we also played videos from one PC to another in the topology shown in Figure 4 and also made voice calls using Ekiga between the two machines. We empirically observed that both these applications showed good performance.

6. DISCUSSION AND CONCLUSION

We have shown that a multihop TDMA protocol can be built on top of commodity hardware by modifying the madwifi driver. We have also designed a complete multihop TDMA system to support multiple flows, allow new nodes to join and leave dynamically. We have experimented with a linear topology with statically allocated transmission opportunities. Nevertheless, the implementation is very close to what we expect from the full-fledged protocol since we have built place-holders for almost all aspects of the MAC design – we have control slots, contention slots, data slots and perform multihop synchronization and data transfer between nodes. In a real deployment with long distance links, channel switching and spatial reuse is possible. In our indoor setting, spatial reuse was not possible but avenues exist for channel switching. We have deferred this to future work.

Our experiments have revealed that there is no system bottleneck in the implementation: the throughput matches what we expect after accounting for the various overheads, and the delay/jitter values are very low. The TDMA MAC implementation thus holds promise for a rich set of applications on outdoor mesh networks. The high throughput and low delay/jitter means that we could also enable a longdistance WiFi mesh network to be used a back-haul, for providing cellular connectivity with a base-station in a rural region.

7. REFERENCES

- Bhaskaran Raman and Kameswari Chebrolu. Design and evaluation of a new mac protocol for long-distance 802.11 mesh networks. In *MobiCom '05: Proceedings* of the 11th annual international conference on Mobile computing and networking, pages 156–169, New York, NY, USA, 2005. ACM.
- [2] Rabin Patra, Sergiu Nedevschi, Sonesh Surana, Anmol Sheth, Lakshminarayanan Subramanian, and Eric Brewer. Wildnet: Design and implementation of high performance wifi based long distance networks. In NSDI, 2007. ACM, SIGCOMM, 2007.
- [3] Bhaskaran Raman, Kameswari Chebrolu, Dattatraya Gokhale, and Sayandeep Sen. On the Feasibility of the Link Abstraction in Wireless Mesh Networks. *IEEE Transactions on Networking*, 17(2):528–541, Apr 2009.
- [4] Kameswari Chebrolu and Bhaskaran Raman. FRACTEL: A Fresh Perspective on (Rural) Mesh Networks. In NSDR, Sep 2007. A Workshop in SIGCOMM 2007.
- [5] Michael Neufeld, Jeff Fifield, Christian Doerr, Anmol Sheth, and Dirk Grunwald. Softmac - flexible wireless research platform. In *Fourth Workshop on Hot Topics* in Networks (HotNets-IV), November 2005.

- [6] Ashish Sharma, Mohit Tiwari, and Haitao Zheng. MadMAC: Building a Reconfiguration Radio Testbed using Commodity 802.11 Hardware. In 1st IEEE Workshop on Networking Technologies for Software Defined Radio Networks, SDR'06, pages 78–83, Sep 2006.
- [7] Ashish Sharma and Elizabeth M. Belding. Freemac: framework for multi-channel mac development on 802.11 hardware. In *PRESTO '08: Proceedings of the ACM workshop on Programmable routers for extensible services of tomorrow*, pages 69–74, New York, NY, USA, 2008. ACM.
- [8] Ananth Rao and Ion Stoica. An overlay mac layer for 802.11 networks. In MobiSys '05: Proceedings of the 3rd international conference on Mobile systems, applications, and services, pages 135–148, New York, NY, USA, 2005. ACM.
- [9] Dimitrios Koutsonikolas, Theodoros Salonidis, Henrik Lundgren, Pascal LeGuyadec, Y. Charlie Hu, and Irfan Sheriff. TDM MAC Protocol Design and Implementation for Wireless Mesh Networks. In *CoNEXT*, Dec 2008.
- [10] Mikrotik Website. http://www.mikrotik.com.
- [11] OpenWRT Website. http://openwrt.org/.
- [12] The MadWifi project Website. http://madwifi-project.org/.
- [13] IEEE P802.11, The Working Group for Wireless LANs. http://grouper.ieee.org/groups/802/11/.
- [14] HAL source code released. http://madwifi-project.org/wiki/news/20080929/ hal-source-code-released.