A Third-Generation Multicast Protocol for HF Wireless Networks

Huiyan Zhang and Eric E. Johnson
Klipsch School of Electrical and Computer Engineering
New Mexico State University

ABSTRACT

Multicasting is an efficient mechanism for disseminating messages through wireless networks, and is commonly used in military C4ISR systems. In this paper we introduce a multicasting protocol for third-generation (3G) HF radio networks that was designed to support such applications as P_MUL (ACP-142) and NATO STANAG 4406. This Multicast Data Link (MDL) employs the robust burst waveforms already in use for 3G ARQ, and operates in either acknowledged or unacknowledged mode. The paper concludes with preliminary performance estimates for the new protocol.

1. INTRODUCTION

High frequency (HF) is a term used to describe the 1.6 to 30 megahertz (MHz) portion of the radio spectrum. This frequency range can provide both short-range and long-haul beyond-line-of-sight communications. HF is widely used for long distance communication since HF is the only way of achieving global communication coverage without using expensive terrestrial or satellite infrastructure.

However, global coverage relies upon radio wave propagation via the ionosphere. There are significant day-to-day changes in ionospheric conditions as well as interference caused by distant transmissions. These make it difficult to find and keep a good frequency for point-to-point links. Modern technology offers a solution to solve the problem of ionospheric channel variability. The first generation of adaptive HF systems was developed in the late 1970s. More functionality was added during the 1980s to enable fully automatic link establishment (ALE) and to ensure link maintenance during message transfer. The major components in an HF radio subnetwork are automatic link establishment (ALE), a data link protocol (DLP), and automatic link maintenance (ALM). ALE is an automated technique that permits HF radio stations to find a suitable frequency and set up a link without operator assistance.

In common with most wireless technologies, an HF network is characterized by low bandwidth broadcast (versus point-to-point) channels. Unlike other wireless networks, however, most or all member stations in an HF network are in direct contact with each other; in other words, network diameter is usually 1 hop. Thus, a multicast protocol will be especially effective in HF networks because there may be no requirement to create and maintain multicast trees (as in RMTP [1] and SRM [2]).

Military wireless networks often must accommodate one-way communication to nodes in "radio silence" (EMCON) and operate efficiently in low bandwidth channels (especially in the HF band). The P_MUL protocol [3, 4] was designed to fulfill multicasting requirements in such networks.

2. P_MUL PROTOCOL

P_MUL (ACP-142) is a reliable multicast protocol for messaging in subnetworks with bandwidth constraints and delayed acknowledgements (e.g., the tactical military message networks supported by NATO STANAG 4406 [5]). It provides users with reliable multicast services and allows for acknowledgements from the receivers to be delayed for a rather long time.

As an application-layer protocol, P_MUL runs on top of a connectionless transport protocol such as the User Datagram Protocol (UDP). P_MUL has been applied to provide communication between X.400 Message Transfer Agents (MTA) and between RFC822 mail servers, as well as other kinds of reliable multicast transmission.

P_MUL operation is depicted in Figure 1. All P_MUL PDUs are transmitted using UDP. The sending processes track the state of the message transfer between the transmitting node and all receiving nodes. The transmitting node delivers all PDUs using the multicast communication service provided in the lower layer. All PDUs sent from receiving nodes back to the transmitting node are transmitted in unicast (point-to-point) mode.

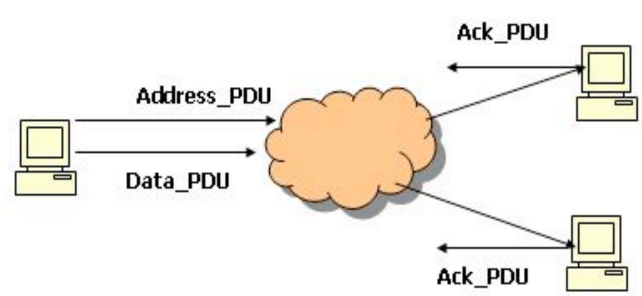

Figure 1. P_MUL protocol

2.1 Protocol Data Units for P_MUL

P_MUL uses four different PDUs for the data transfer:

- transmitter identification of message address (Address_PDU),
- transmission of message fragments (Data_PDU),
- receiver acknowledgement message (ACK_PDU),
- transmitter termination of the transmission of a specific message (Discard_PDU).

The ACK_PDU is generated and transmitted by each receiver and is evaluated by the sender. The other three PDU types are generated and transmitted by the sender and are evaluated by the receivers.

2.2 Example Message Transfer Using P_MUL

The simple example in Figure 2 illustrates the P_MUL message transfer. S0 sends one message to the nodes R0, R1, R2 and R3. R2 and R3 are under EMCON. The message is split into three fragments.

- S0 creates an Address_PDU and three Data_PDUs and sends them to the multicast address.
- R0, R1, R2, and R3 receive the Address_PDU.
- R0, R2, and R3 receive the Data_PDUs without errors.
- R1 receives the first and third Data_PDU with errors, but receives the second PDU error-free.
- R0 sends a complete ACK_PDU
- R1 sends an ACK_PDU indicating the missing Data PDUs.
- R2 and R3 cannot send any ACK_PDU because they are in EMCON mode.

Note that R0 and R1 may have to contend for the channel to send their ACK_PDUs.

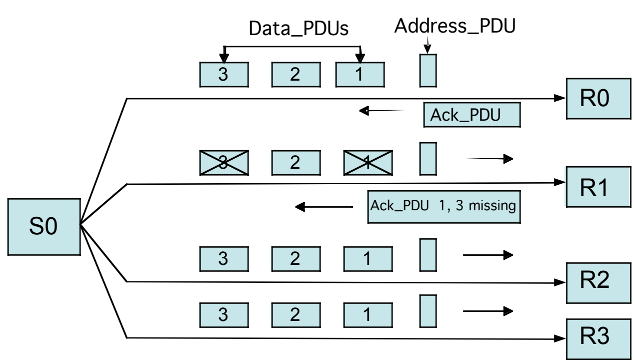

Figure 2. S0 sends message to R0-R4; R0 and R1 acknowledge

S0 will immediately retransmit the first and third PDUs, after deleting R0 from the Address_PDU. S0 will await ACK_PDUs from R2 and R3 until the *Transmitter_Expiry_time* for the message is exceeded, at which time it will send a Discard_PDU to any stations that have not acknowledged the entire message.

3. OVERVIEW OF 3G HF TECHNOLOGY

Currently two generations of HF automation are in use, commonly referred to as second generation (2G) and third generation (3G). Both generations of automated HF protocols are standardized in MIL-STD-188-141B: 2G in Appendix A and 3G in Appendix C.

Second-generation HF automation provided a sufficiently robust, reliable, and interoperable ALE technology to spark a resurgence of interest in HF radio for long-haul and mobile voice networks beginning in the 1980s. A robust data link protocol [6] was added to 2G ALE to support data applications over HF as well. However, the second generation technology did not scale well for large networks with heavy traffic within the limited HF channel capacity because of its overhead traffic.

3G HF automation technology introduces several improvements over 2G ALE, including synchronous scanning (for faster linking), a burst phase-shift-keying waveform (more robust), and a slotted carrier sense multiple access with collision avoidance (CSMA/CA) channel access procedure (for heavier traffic).

The third-generation ALE (3G ALE) protocol, the traffic management (TM) protocol, the High-Rate Data Link (HDL) and Low-Latency Data Link (LDL) protocols, and the circuit link management (CLC) protocol form an integrated protocol suite (Figure 3).

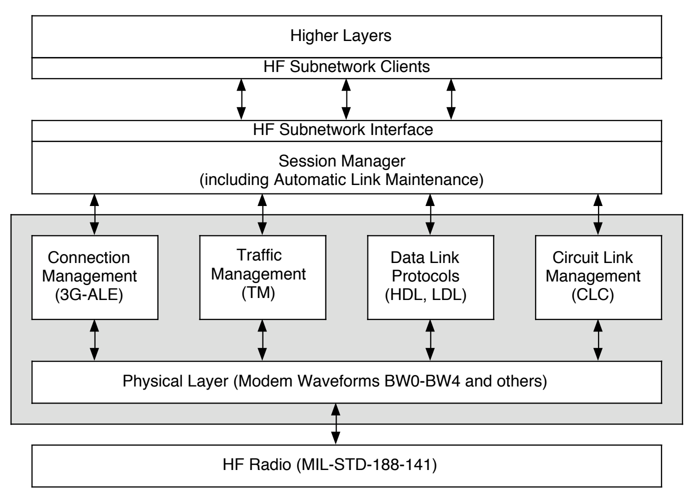

Figure 3. 3G HF protocol suite

3.1 3G Automatic Link Establishment

The Connection Manager in the 3G suite is responsible for making and breaking links via HF channels. Its Automatic Channel Selection (ACS) function fuses all available information about propagation characteristics of available channels (measured and/or predicted). The ALE function relies upon ACS when choosing channels for new traffic links requested by the Session Manager process. 3G ALE provides functionality similar to second-generation ALE, but it has improved ability to link in stressed channels, and operates more quickly and efficiently in large, data-oriented networks.

3G ALE stations synchronously[1] scan an assigned pool of frequencies listening for calls. Receivers dwell on each frequency for 5.4 seconds. Each synchronous dwell time is divided into six slots of 900 ms each. The first slot is reserved for retuning RF components for the new dwell frequency and for traffic monitoring. Two-way handshakes that establish links may begin in any of the next four slots. The last slot is used for completing handshakes that began in the fourth slot and for other special functions.

3.1.1 3G-ALE synchronous mode individual calling

The one-to-one linking protocol quickly identifies a frequency for traffic use and minimizes channel occupancy (Figure 5). A two-phase handshake on the current calling channel of the destination station is used to establish a link. First a Call PDU is sent by the calling station (Caller in Figure 5). This PDU contains the address of the destination station along with the type of traffic to be carried on this link. 3G ALE PDUs use Burst Waveform 0 (BW0), a robust PSK burst lasting only 613 ms.

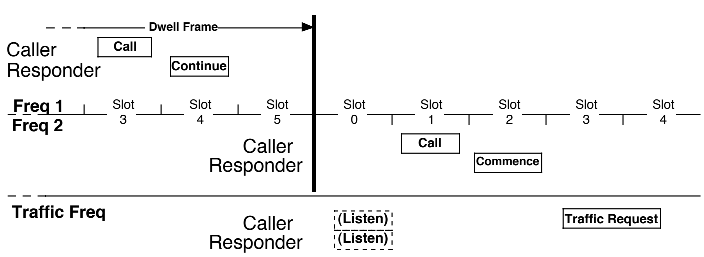

Figure 5. Synchronous mode 3G ALE protocol

If the destination station receives the Call, it checks its ACS database for a suitable traffic channel for the requested link. If no suitable channel is known, the Responder returns an Handshake PDU with a Continue Calling command (as seen on Freq 1 in Figure 5).

However, if the Responder ACS database finds a suitable channel for the announced traffic, the Responder instead replies with a Commence Traffic Handshake PDU (see Freq 2 in Figure 5). This PDU specifies the traffic channel to be used. Both stations then re-tune to the specified frequency and begin the traffic setup protocol.

3.1.2 3G-ALE synchronous mode multicast calling

A Multicast call is used to contact selected stations concurrently and direct them to a traffic channel selected by the calling station. In this case, the Call PDU sent by the Caller is addressed to a multicast address, and contains a Traffic Type of Multicast.

No station responds to a Multicast-type Call PDU. Instead, the caller sends a Handshake PDU immediately following the Call PDU that directs the called stations to a traffic channel. The called stations tune to that channel and listen for traffic. If the announced traffic does not begin within a traffic wait timeout, the stations return to scan.

3.2 TM Protocol

The 3G Traffic Manager (TM) protocol is used to coordinate traffic exchanges on connections established using the 3G ALE protocol. The participating stations use a two-way exchange of TM PDUs to determine the data link protocol, waveform, traffic priority, and so on. The TM protocol uses three PDUs (TM_Request, TM_Confirm, and TM_Term), sent using Burst Waveform 1 (BW1).

3.3 Data-Link Protocols

A 3G HF network uses two[2] data link protocols to transmit messages. A 2-way TM handshake on a channel synchronizes the time bases of the data link terminals, and determines the direction and mode of data transfer. Following this handshake, the link runs in either high-throughput or low-latency mode. The high-throughput data link (HDL) protocol is for large messages and/or good channel conditions, while the low-latency data link (LDL) protocol is for short messages and poor channel conditions.

HDL is a selective repeat hybrid ARQ protocol that uses code combining to correct data received in error. The HDL_Data PDU uses Burst Waveform 2 (BW2), which carries data at rate R=1 in the initial transmission. Transmissions of a packet cycle through four sets of error-correction bits for code combining.

LDL uses the very robust Burst Waveform 3 (BW3) in a code combining stop-and-wait protocol.

---

[1] Lower-performance asynchronous 2G and 3G modes are also available for special situations.

[2] A third protocol (HDL+) is under development. The TM protocol can also set up traffic using other data protocols (e.g., STANAG 5066), as well as handing off to analog and digital voice systems.

4. MULTICASTING IN 3G HF NETWORKS

P_MUL is emerging as the standard reliable multicast protocol for military messaging in bandwidth-constrained subnetworks. The initial work in using P_MUL over HF channels [7] has used 2G protocols because 3G support is currently incomplete. In this paper, we propose modifications to the 3G protocol suite to fully support P_MUL in 3G HF networks.

4.1 Requirements to support P_MUL

As an application-layer protocol, P_MUL attempts to provide reliable message delivery using UDP or other best effort transport protocols, and requires no support from the network apart from multicast routing (since the destination of the traffic is a multicast address).

Efficient support for P_MUL-style multicasting in an HF network requires the following:

1. ALE with multicast calling, so that the implicitly addressed stations can all be directed to a channel.
2. A data link protocol for sending the data with forward error correction. Multi-party ARQ operation at the link layer is optional.
3. An efficient mechanism for returning the application-layer acknowledgements to the sender if reliable multicast is desired.

The existing 3G protocol suite meets requirements 1 and 3. To satisfy requirement 2, we propose to add a new Multicast Data Link (MDL) protocol and extend the Traffic management (TM) protocol to support the new MDL. The expanded 3G protocol suite is depicted in Figure 4.

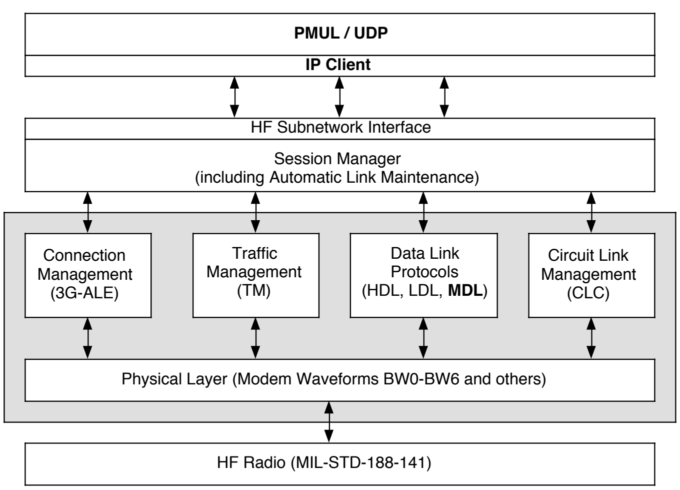

Figure 4. 3G Suite with Multicast Support

Packets from a multicast application are processed by the P_MUL, UDP and IP processes, and are routed to the 3G session manager for HF transmission.

The sections below describe the changes to the 3G suite for supporting application-layer multicasting.

4.1.1 Client interface changes

Address resolution in the IP client will need to accommodate mapping IP multicast addresses to 3G HF multicast addresses. The mechanism for this, and for indicating the multicast traffic type via the subnetwork interface is beyond the scope of this paper.

4.1.2 Connection manager

The Connection Manager will handle multicast traffic by directing ALE to establish a multicast link. This will place new demands on the ACS function, as it will need to find a channel with usable propagation to multiple destinations. Due to the complex temporal and spatial variations in ionospheric propagation, it may not be possible to find a single suitable channel for all multicast subscribers, so multiple multicasts may be required. In such cases, a roll call after linking would be useful.

4.1.3 Traffic manager

After a multicast link is established, the TM protocol will send a TM_Request PDU to the recipients that announces use of the MDL protocol for traffic. Unlike ARQ traffic, for which the TM_Request prompts a TM_Confirm response, the TM_ Request that announces MDL as the data link protocol specifies either no response or a roll call to identify which stations are present to receive the multicast.

4.1.4 MDL protocol

A range of approaches is possible for adding multicast support to the 3G HF suite:

1. Each application-layer message could be sent as 3G packets, with an opportunity for link-layer repeat request(s) after each packet or fixed-length series of packets.
2. The 3G packets composing a message could be sent in a continuous stream, with roll call selective-repeat acknowledgements at the end of the message.
3. The 3G link layer could operate in broadcast mode after link establishment, using application layer acks for reliability. In this case, we could send the data using either 3G bursts or a continuous-data modem.

The first two approaches provide a reliable link layer which can deliver error-free multicast data in-order. However, error-tolerant applications may prefer the latter approach to avoid retransmission delays in the link layer.

Multicast approaches 2 and 3 separate packet retransmissions in time by the duration of the message (tens of seconds to many minutes). This time diversity should generally provide uncorrelated channel conditions for the retransmissions, which helps in correcting errors.

We here define a multicast data-link protocol (MDL) that supports multicast approaches 2 and 3. Data is sent in broadcast mode using one of the code-combing waveforms, BW2 or BW3, with optional roll-call acknowledgements after the message. For simplicity, the MDL protocol uses existing HDL and LDL PDUs: LDL_DATA, LDL_EOM, HDL_DATA, and HDL_EOM.

For efficiency, MDL data is sent continuously until the message is complete, with no pauses for acknowledgements. In this investigation, we consider four data modes for sending MDL data: BW3 (32 or 512 bytes per frame) and BW2 (message sent twice or four times).

- In the BW3 modes, the message is sent once, with all packets sent in order. Error correction via code combining requires retransmission at the application layer.
- In the BW2 modes, the entire message is sent either twice or four times, with different FEC bits sent in each repetition. The repeated packets are separated in time by the length of the message, so code combining error correction should be effective within a single transmission.

4.1.5 MDL acknowledgements

P_MUL ACK PDUs will be passed down the protocol stack as short point-to-point messages, and will be delivered using the 3G LDL protocol. Note that the end of the P_MUL message will prompt all of the active recipients to contend for the channel at the same time to send their acks. The resulting contention delays could be substantial in a large network.

To avoid this, a link-layer roll call may be used to provide reliability, and P_MUL acks may be proxied by the link layer to avoid channel contention.

5. PERFORMANCE ANALYSIS

In this section, we develop analytical models for estimates of MDL performance using approach 3, followed by analysis of MDL with a link-layer roll call (approach 2).

5.1 Analytical model for MDL without roll-call

We first compute throughput for the 3G Multicast Data link (MDL) including the time required for link setup and receipt of the application-layer acknowledgements (including contention for channel access).

- The multicast Call is assumed to reach all members of the multicast group on the first attempt.
- The links between the sender and receivers are modeled as single hop channels (i.e., no relaying is needed).
- All channels are mutually independent and have the same average SNR. Frame error probabilities as a function of SNR were computed using a simple model extracted from measurements of 3G bursts.

We use the following statistics to calculate the throughput for multicast transmission:

Multicast message delivery time (T_{cycle}): the time measured from the start of link establishment through the moment all active recipients have sent acknowledgements (including the effects of contention).

Throughput: the average number of user data bits received error-free during a cycle divided by T_{cycle}.

Multicast transmission time: the time required for the multicast originator to set up the multicast link, announce the MDL protocol, and send the message. Assuming that all recipients can be reached with a single 3G call, this is

$$T_{MDL-Tx} = 2 * T_{slot} + T_{tune} + T_{BW1} + T_{BW2-enc} + T_{MDL-Data}$$

Acknowledgement time: the time for the recipients to successfully send their acknowledgements, including channel contention delays.

All P_MUL ACK_PDUs sent from receiving nodes back to the transmitting node are transmitted in point-to-point mode. This means N receivers compete for the channel during the first contention window to send their acks.

From [8], the probability of a successful acquisition in slot i of S slots, $\sigma_{S|N}(i)$, and the overall probability of success with N nodes in contention, $P_{Success}(N)$, are

$$\sigma_{S|N}(i) = \left(\frac{N}{S}\right)\left(\frac{S-i}{S}\right)^{N-1} \text{ and } \quad P_{Success}(N) = \sum_{i=1}^{S} \sigma_{S|N}(i)$$

If the multicast recipient is successful in sending its ACK_PDU, the sender will send an LDL_EOM PDU to the recipient. The total time occupied in case of success is (see MIL-STD-188-141B App. C for timing constants)

$$T_{ack-success} = 900ms + 900ms + T_1 + T_{BW1} + T_2 + T_{BW1} + T_3 + T_{ack-data} + T_4 + T_{BW4}$$

$$T_1 = T_{prop} + T_{BW0proc} + T_{BW1enc} \qquad T_2 = T_{prop} + T_{BW1proc} + T_{BW1enc}$$

$$T_3 = T_{prop} + T_{BW1proc} + T_{BW3enc} \qquad T_4 = T_{prop} + T_{BW3proc} + T_{BW4enc}$$

If the slot in a contention window instead contains a collision, the call fails in that dwell time and the recipients will try again in the next dwell time. The channel time occupied for a failed attempt is

$$T_{fail} = T_{dwell} = 5.4s$$

The total acknowledgement time for one recipient is then

$$T_{ack} = T_{ack-success} + (\frac{1}{P_{success}} - 1) * T_{fail}$$

For N recipients, the total acknowledgement time is N successful ack times, plus the dwell times lost to collisions. Note that the probability of a collision in each dwell drops as successful acks reduce the number of contending nodes.

$$T_{N-ack} = N * T_{ack-success} + \sum_{j=1}^{N} (\frac{1}{P_{success}(j)} - 1) * T_{fail}$$

Using the above formula and values, we can compute the expected throughput for one cycle:

$$X = \frac{(frames/transmission)(bits/frame)(1 - FER)}{T_{Cycle}}$$

where FER is the error rate of the frames (a function of SNR and other channel characteristics). Note that frames lost to channel errors will be retransmitted in a subsequent cycle; we count good bits per cycle as throughput. This is the average throughput *per recipient.*

5.2 Analytical results

Figures 6 and 7 show the overall throughput of the four traffic modes for multicast delivery with short (5000 byte) messages. The number of active recipients is 4 in Figure 6 and 8 in Figure 7. The tradeoff between throughput and SNR is clear in these figures. Because of contention, the duration of the acknowledgement phase increases faster than linear with the number of recipients, reducing per-node throughput for N = 8. For these short messages, the traffic waveform of choice appears to be BW3-32 for low SNR (up to 8 or 10 dB). For higher SNR, either of the BW2 waveforms provides better throughput.

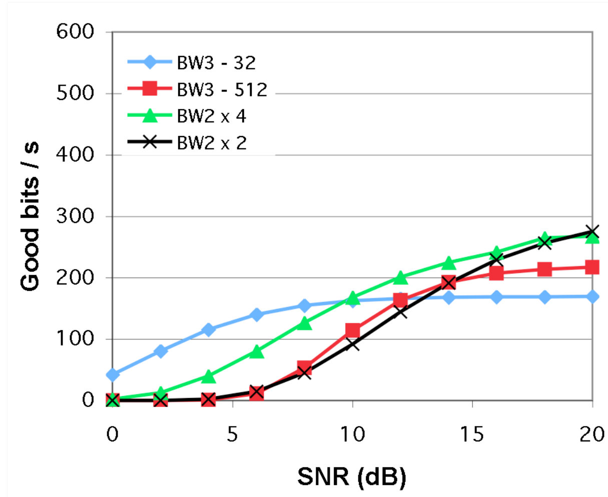

Figure 7. Throughput for multicast delivery (N=8, file size 5000 bytes)

Figures 8 and 9 show the throughput for larger payloads: 100,000 bytes. As expected, the setup and acknowledgement overhead is much less significant here, and throughput is correspondingly improved. For these larger messages, the higher asymptotic throughput of the BW2 x 2 traffic mode is clearly the best choice for voice-grade channels (10 dB SNR or better).

5.3 MDL with roll-call

The MDL described above operates in broadcast mode at the link layer for multicast data delivery, but each packet carries a sequence number that could be used in link-layer ARQ as in multicast approach 2. To estimate the performance of MDL with roll-call acks, we assume that the roll-call is initiated by a BW1 PDU, followed by BW3 acks.

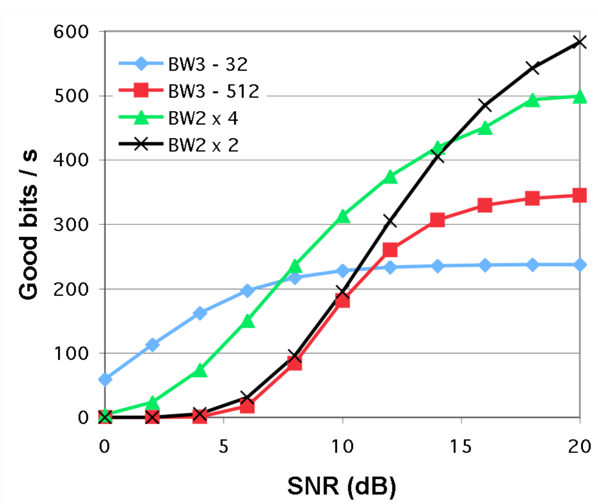

Figure 6. Throughput for multicast delivery (N=4, file size 5000 bytes)

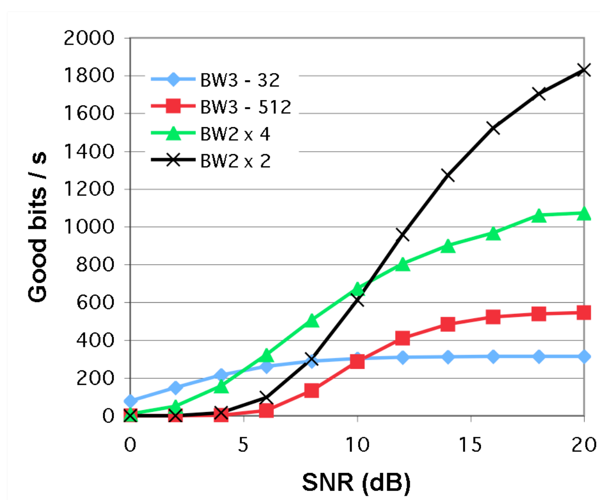

Figure 8. Throughput for multicast delivery (N=4, file size 100,000 bytes)

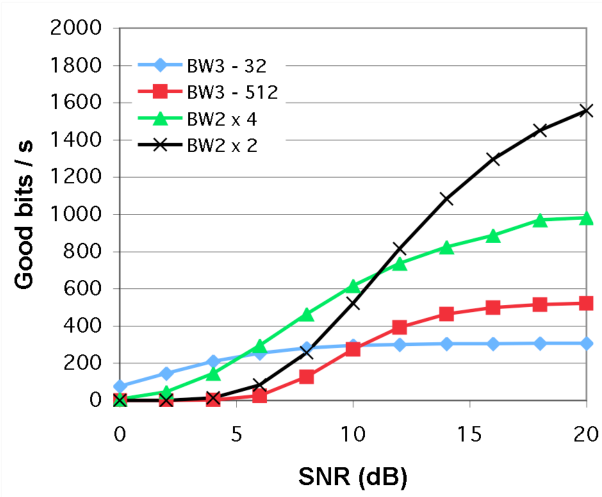

Figure 9. Throughput for multicast delivery (N=8, file size 100,000 bytes)

The link-layer acknowledgements returned via roll-call after the message must be able to carry an Ack bit for each packet in the message. Our 100,000 byte message contains 430 packets, which requires a 64-byte BW3 ack packet, while the 5000 byte message needs only 32 bytes.

Figure 10 compares the throughput of MDL with roll call (contention-free) to MDL with P_MUL acks, using the BW2 traffic modes for a 5000 byte message and 8 nodes. This roll-call mechanism for multicast acks is clearly attractive for short messages and larger networks.

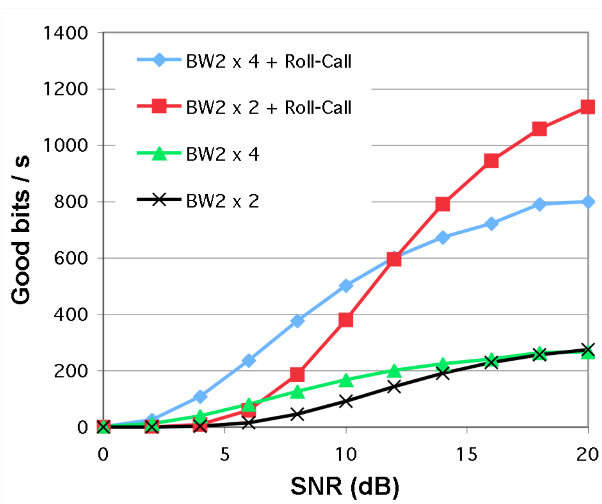

Figure 10. Roll-call *vs* P_MUL Acks (N=8, file size 5000 bytes)

6. CONCLUSIONS AND FUTURE WORK

One potential application for HF multicasting is maritime shore-to-ship transmissions. Today, maritime broadcasts often operate with a serial-tone PSK modem at 600 bps, which gives good performance down to the 5-10 dB SNR range. From the preliminary results in this research, it appears that 3G multicast, even in its less-efficient contention-based mode, could provide roughly the same throughput as the broadcast mode even while accommodating acknowledgements and retransmissions (for multicast groups of up to 8 nodes).

Four traffic modes were evaluated using a simple model of FER versus SNR. These results suggest that choice of waveform could be made considering only message length, with no knowledge of SNR on the multiple paths from the sender to individual multicast group members. For short messages, it appears that the BW3-32 mode would be a good choice, although its performance falls short of the BW2 modes at high SNR. For long messages the higher throughput of the BW2 modes is preferred in all but the worst conditions.

Future work will include simulation and measurements to get more accurate results. We also plan to investigate the scalability of the scheme to larger networks, and explore schemes to suppress the acknowledgements from the receivers and to reduce the retransmission from the sender, as well as the roll-call acknowledgment idea.

7. REFERENCES

1. S.Paul and K.K. Sabnani, "Reliable Multicast Transport Protocol (RMTP)," *IEEE Journal of Selected Areas in Communications*, **15** (3), pp. 407-421, 1997.

4. S. Floyd, V. Jacobson, and C.G. Liu, " A Reliable Multicast Framework for Light Weight Session and Application Level Framing," in *Proc. ACM SIGCOMM'95*, pp.342-356, August 1995.

5. ACP 142, P_MUL – A Protocol for the Reliable Multicast Messaging in Bandwidth Constrained and Delayed Acknowledgement Environments, Allied Communications Publication, August 2001.

6. J.Tovirac and W. Zhang "A Simulation Study of P_MUL Multicasting Protocol Enhanced by Ideal FEC Code." in Proc. of the LASTED International Conf. on Wireless and Optical Comms.(WOC2002), Banff, Canada, Jul.2002.

7. NATO Standardization Agreement (STANAG) 4406, "Military Message Handling System," Ed. 1, 1999.

8. STANAG 5066, "Profile for High Frequency (HF) Radio Data Communications," version 1.2, 1999.

9. V. Jodalen, A.Eggen, B.Solberg, and O.Grønnerud, "Military Messaging in IP Networks Using HF Links", *IEEE Communications Magazine*, November 2004

10. E.E. Johnson, M. Balakrishnan and Z.Tang, "Impact of Turnaround Time on Wireless MAC Protocols," *Proc. MILCOM 2003*, IEEE, Boston, October 2003