A High Performance, Collision-Free Packet Radio Network

Phil Karn, KA9Q

ABSTRACT

For the past several years, those discussing "level 3 networking" have made much of the performance gains to be had through hop-by-hop acknowledgements. In this paper I will show that, while sometimes helpful, hop-by-hop ACKing is not the panacea it is generally perceived to be. Only fundamental changes in the way we allocate and use frequencies will really fix the problem.

1. Introduction

At present, our networks can best be described as "anarchistic." Single frequency digipeaters abound, and everyone knows just how likely you are to get a packet across five digipeater hops on a heavily loaded frequency [2]. Given this situation, software that provides hop-by-hop acknowledgements (e.g., NET/ROM [4]) is clearly a major win. Actively retransmitting ACKs, as in the ACK-ACK protocol [3] would yield an additional improvement.

Yet NET/ROM and ACK-ACK both fail to attack the fundamental problem: *carrier sense multiple access (CSMA) simply doesn't work very well on an open-access simplex radio channel.* Two things contribute to this. The first is the well-known *hidden terminal problem:* not sensing carrier on the channel does NOT guarantee that you won't interfere with someone if you transmit.

The second problem is less well known. Because it is the converse of the hidden terminal problem I will call it the *exposed terminal problem!* A station in a good location (e.g., a mountaintop) may hear local traffic from within a distant area. Not knowing that it would not interfere with that traffic by transmitting, it defers unnecessarily and wastes an opportunity to reuse the frequency locally.

In short, the carrier detect line from the modem is often useless. There is no guarantee that you won't interfere with someone if you transmit when you don't hear a carrier, and conversely there is no guarantee that you *would* interfere with another transmission even if you transmit when you *do* hear a carrier.

It is well known (and proven in practice!) that CSMA breaks down in the presence of hidden terminals, degrading rapidly to the performance of pure Aloha (where stations transmit at will, without first monitoring the channel). With the standard Aloha assumptions (many terminals each generating a tiny fraction of the total channel load) the maximum attainable channel throughput is only 18%. This occurs at an offered load of 50%, i.e., each packet has to be transmitted about 2.7 times on the average for it to be received once. Although hop-by-hop acknowledgements keep these figures from getting exponentially worse across a multi-hop path, they do *not* fix the fundamental problem: *CHANNEL COLLISIONS*!

This is a very important point. *Using link level ACKs to improve performance is, at best, a **band-aid solution.*** Because they represent overhead, sometimes they are actually counterproductive. The real challenge, therefore, is to make collisions impossible in normal operation. I will now discuss two of the traditional methods for collision avoidance when hidden terminals are present.

---

[1] George Flammer, WB6RAL, calls this the white lightning effect. [1]

2. Token Passing

One way to avoid collisions is to require each station to wait for explicit, one-at-a-time permission to transmit. When a station has sent its traffic, it passes this authority on to the next station. Since the message that grants permission to transmit is known as a *token,* this scheme is known as *token passing.*

Token passing works well in small networks with reliable nodes and links, but it doesn't scale well. Complex recovery algorithms must be worked out to recover from lost tokens caused either by failing nodes or transmission errors. In a packet radio network with many hidden terminals, the route that the token will take must be mapped out in advance; it cannot be passed between stations that cannot communicate. This complicates the addition of new stations to the network. In addition, much time is wasted passing the token when there are many stations in the network but only a few are actually sending traffic. Nevertheless, token passing is a completely unexplored technique in amateur radio, one that deserves serious consideration for special circumstances.

3. Busy Tone Multiple Access (BTMA)

Another effective technique for eliminating collisions when hidden terminals are present is for each station to transmit a signal on a separate frequency whenever it is actively receiving a packet. If another node hears this *busy tone,* it avoids transmitting knowing that it would interfere with the reception in progress. It is not necessary for a node to couple its busy tone directly to the receiver carrier detect indication; it may drop the busy tone once it determines by examining the packet destination address that the packet is for another station. This allows *frequency reuse* (successful simultaneous use of the same frequency by two pairs of stations far enough apart not to interfere with each other).

In theory, BTMA can be an effective solution to the hidden terminal problem. However, extra radio hardware is required since the busy tone transmitter must operate without interfering with data reception. In practice this means using separate frequency bands, and it may be difficult to get the range of the busy tone transmitter to match that of the data transmitter -- a fundamental assumption in BTMA. It is also difficult to get BTMA to solve the exposed terminal problem. Hearing a busy tone doesn't *always* mean that you'd interfere with a receiver if its desired signal is much stronger than yours, depending on the capture ratio of the modulation method in use. Setting the busy tone's amplitude in inverse relationship to the level of the signal being received, plus lots of tricky threshold adjustments in the busy tone receivers, might make this work.

4. Contention-Free Channels

The discussion so far has centered on reducing or eliminating collisions when a single frequency must be shared by more than one transmitter. Contention channels are likely to be with us for some time where random end-users are involved. However, the emerging network of dedicated, "backbone" sites need not follow the same anarchistic model. The rest of this paper discusses a more disciplined approach that appears extremely attractive for such stations.

One sure way to eliminate collisions is to eliminate all but one transmitter on each frequency. All other transmitters on the same frequency must be placed far enough apart so that their coverage areas do not overlap. Each station uses a separate, dedicated receiver to hear each of its neighbors; it does not listen on its own transmit frequency. A network node might look like this:

```
Beam or Omni    Beam or Omni    Beam or Omni
  Antenna         Antenna         Antenna
     ↓               ↓               ↓
 Receiver 1      Receiver 2      Receiver N
     ↓               ↓               ↓
+---------------------------------------------+
|                Packet switch                |
+---------------------------------------------+
                     ↓
                Transmitter
                     ↓
                Omni antenna
```

Many things now become easier or perhaps even possible for the first time. As it is no longer necessary to "get off the frequency" quickly when a station has sent its traffic, fast transmit-receive switching is no longer required. Transmitters and power amplifiers with relays or slow-lockup synthesizers need not be modified; they could operate either continuously, or with tail timers like those in conventional voice repeaters. Similarly, coherent receiver demodulators (which work well with very low signal levels but require relatively long acquisition times) need not penalize network performance. The link receivers may be cheap pocket scanners since they need not transmit. If adjacent nodes transmit on different bands, the expense of repeater-style duplexers can be avoided, although filter cavities ("trashcans") may still be needed (especially at hilltop sites) to reject strong out-of-band signals.

Since the design of this network makes collisions impossible, with proper modem design and adequate RF link margins the raw packet loss rate should be very low. The occasional end-to-end retransmission of a dropped packet will be more than offset by the savings in overhead gained by avoiding link level acknowledgements. High channel speeds are much easier to handle since the packet switches are much simpler, and real time applications such as packet voice become practical. Since the nodes are inherently full duplex, sliding-window transport protocols (with data packets and acknowledgements flowing simultaneously in both directions) finally make sense, as data/ack collisions are avoided.

5. Broadcasting

In addition, some very powerful broadcast techniques become possible. Much of the traffic now handled by bulletin boards consists of undirected messages read by a wide audience. At present, our virtual circuit protocols require that a separate copy be sent to and acknowledged by every interested reader. This wastes one of the most useful and unique properties of radio: the ability of more than one receiver to hear a single transmitter. Efficient but reliable broadcasting on a very unreliable channel (e.g., an existing digipeater network) is almost impossible. However, the situation changes completely if the raw packet loss rate can be lowered to a reasonable level.

Consider the operation of an ordinary voice bulletin net, one organized to disseminate information of general interest to many stations. (A good example is the Tuesday night AMSAT net on 75 meters). After the control station finishes reading,% he invites requests for repeats. If conditions are good, only a few stations will respond, and the requested message fragments are retransmitted. As with the original transmission, all receiving stations are free to make use of the retransmitted information; this often preempts a second station's request for a fill. If conditions are bad, the control station may first read the entire bulletin several times (a simple form of forward error correction) to cut down the number of fill requests.

6. Flood Routing

Given a reasonably reliable channel (i.e., one with only a single transmitter) this scheme should be easy to automate. Wide-area bulletin coverage could be achieved with a flood routing scheme similar to the USENET bulletin board network. In flooding, a node originating a message transmits it to all of its neighbors. Each message contains a unique network-wide identifier (e.g., the node address concatenated with a serial number). Each receiving node maintains a list of messages it has already seen and ignores duplicates. A non-duplicate message is entered into the list and retransmitted to its neighbors until it has spread to every reachable node in the network.

Flooding is extremely robust, as it tries every possible route to each node in parallel. USENET has proven this in practice, despite an amazingly anarchistic network management style. It is the preferred way to reach large numbers of people, since a given message crosses each link in the network exactly once. Because of its reliability, flooding is a useful fallback for high priority point-to-point traffic when ordinary routing schemes have failed. (One often finds person-to-person messages posted on USENET because direct mail routing hasn't worked. Clearly this is to be discouraged except as a last resort because of the unnecessary load this generates.)

7. Summary

The use of contention-based channel access algorithms is perhaps unavoidable where end users are involved. However, such free-for-alls are inappropriate on backbone links in light of the severe performance problems involved. The evolving backbone networks should take a more enlightened approach. Instead of just attempting to patch things up at a higher layer by adding hop-by-hop acknowledgements, they should be carefully planned to avoid collisions altogether. Not only can the extra overhead of hop-by-hop acknowledgements be avoided, but qualitatively new and vastly more efficient bulletin dissemination techniques fall out almost for free. Considering the vastly improved performance and functionality that would result, the extra costs of doing so are minimal.

8. References

1. Flammer, G., "Survival Training for Mountaintop Digipeaters," *73 Magazine,* August 1986 p. 68.
2. Clark, T., "The Trouble with Digipeaters," *Gateway.*
3. Karn, P. and Lloyd, B.: "Link Level Protocols Revisited," *ARRL Amateur Radio Fifth Computer Networking Conference,* pp. 5.25-5.37, Orlando, 9 March 1986.
4. Raikes, R., and Busch, M., "NET/ROM Version 1 Documentation," Software 2000 Inc, May, 1987.