Addressing and Routing Issues in Amateur Packet Radio

Phil Karn, KA9Q

Radio Amateur Satellite Corporation

ABSTRACT

As amateur packet radio evolves from scattered, ad-hoc collections of local area digipeaters into a large, automatic, and interconnected network, several issues related to naming, addressing and routing will have to be faced and overcome.

Routing, in particular, has long been a fertile research area in computer networking. I make no claim to knowing the answers to many of these problems; however, I believe that they can at least be stated, and that certain decisions can be made early to ease experimentation with various solutions. In particular, the problem of address assignment is discussed with particular emphasis on making the routing problem easier.

1. Introduction: Terminology

I will begin by defining several important terms. A *link* is any transmission line, radio channel or the like capable of carrying a packet directly between two points. *Nodes* are the end points of links. A node may generate packets for other nodes, consume packets addressed to itself, or act as a relay point for packets originating from and addressed to other nodes.

Reference [1] gives a concise but effective definition of three more concepts: "In simple terms, a *name* tells what an object is; an *address* tells where it is; and a *route* tells how to get there."

To elaborate:

1. A *name* is an arbitrary string of characters, chosen for human convenience, to designate a particular person, node or service. Examples include people's names and the network names given to computers. Amateur radio callsigns might also qualify, although many may dispute their convenience!
2. An *address* is a number corresponding to a name, significant to a communications network. It is generally smaller than a name and has a well-defined format. Examples include telephone numbers (10 decimal digits in North America), Internet Protocol addresses (32 bits) and, of course, postal addresses.
3. A *route* is the path over which a specified address may be reached from a given point within a network.

Some communication systems blur the distinction between these concepts. For example, an old-time telephone system with a human operator might accept either a person's name or a telephone number in placing a call. A courier might accept a route ("go to the third (red) house on the right after making a left turn at the light") in place of an address. However, machines demand short, precise identifiers that are often not convenient for humans to remember, so translation from names to addresses then to routes must be provided. Telephone books and directory assistance systems provide translation between names and telephone numbers, while maps provide the information necessary to find a route to a given street address.

While names are generally arbitrary (except that they must be unique, at least in the context in which they are used), addresses may or may not imply something about physical location to facilitate route selection. For example, telephone numbers in North America are assigned by a multi-level location-dependent hierarchy; the first level is the 3-digit area code and the second level is the 3-digit central office code. If you move from New York to Los Angeles you are not allowed to keep the same telephone number; you must get a new one that reflects your new location. You are allowed to keep your name, however; the telephone company grants you this one concession!

Because of the arbitrary nature of names, full-blown database systems are generally required to convert names into addresses. The telephone network provides name-to-address mapping through telephone books and directory assistance bureaus. Computer networks may use either or both of these techniques, e.g., a directory may be maintained locally in each user system, or a central site may be set up as a "name server."

2. Naming In Computer Networks

The main problem here is in verifying the uniqueness of a selected name. For small networks, a single central clearinghouse for name assignments is practical. For large networks, however, hierarchical allocation is the most practical approach.

The ARPA Internet has grown rapidly over the past 5 years, and its central name registry is showing considerable signs of strain. To answer this problem, the ARPA Internet will be evolving from a single, globally administered name space to a hierarchical "domain-based" system [12,13]. In the domain system, a name may consist of several words separated by periods: FOOBAR.ARPA, meaning host name "FOOBAR" under domain "ARPA." The same system might be more fully named as FOOBAR.MIT.ARPA, distinguishing it from any other hosts that might also be named "FOOBAR" in other locations within the ARPA "top-level domain." Partially specified domains are interpreted based on the location in the hierarchy where the name is encountered, and "nearby" (in a hierarchical sense) hosts might be named without any domain names at all. This is analogous to the way people refer to others by their first names (e.g., within a family or work group), but would give full names when referring to someone on the "outside."

The ARPA Internet has only begun its conversion to the domain system, and several details need to be worked out. Within amateur radio, the easiest naming convention would simply be to use our callsigns, since they are already unique. A domain name could be allocated (e.g., AMPRNET) so that when the ARPA Internet finishes its conversion our callsign-names would simply become, e.g., "KA9Q.AMPRNET" outside our network.

3. The Routing Problem

Once the network is given an address, it must select a route to reach it. Again, this may be done centrally, in a network manager that keeps track of the entire network (e.g., TYMNET [4]), or it may be done in a distributed fashion by local routing algorithms that operate from partial, locally constructed views of the network state and topology (ARPANET, many others [2,5,8]).

3.1 Centralized Routing

Centralized routers [5] have the advantage that a complete, coherent "picture" of the entire network can be maintained at a single point. Decisions can be based on the greatest possible amount of information, and attempt to optimize resource usage over a large area. However, centralization has several serious disadvantages. Communication overhead is involved in the collection of status reports from and the dissemination of routing decisions to the individual packet switches. The reliability of the communication paths to the central router (and that of the central site itself) is critical to the entire network.

The communication overhead with a centralized router more practical in a virtual circuit network, since routing is done only at circuit setup time. Such an approach is harder in datagram networks such as those based on ARPA IP because routing must be done on a per-packet basis (although one might "cache" routing information at the switches to minimize overhead).

3.2 Distributed Routing

An alternative is *distributed routing*, where each switch makes its own routing decisions based on a "local view" of the network. Switches usually exchange information with each other, either automatically or on a demand basis, thus maintaining a composite "snapshot" of the network. Such algorithms work well in datagram environments. They have other advantages, such as improved flexibility and reliability because of the lack of dependence on a central site.

For these reasons, distributed routing algorithms are popular, especially in datagram based networks such as the ARPA Internet, and I recommend its use in amateur packet radio. In the remainder of this paper, I will assume the use of distributed routing.

4. Routing Implications for Addressing

A packet switched network consists of a collection of nodes acting as packet sources, sinks and relay points. Depending on the topology of the network, a node may have to decide where to send a packet not destined to itself. In some cases, this is trivial. If the nodes share a transmission media allowing each node to communicate directly with all other nodes (e.g., Ethernet, closely spaced terrestrial packet radio nodes or nodes sharing a satellite channel) routing becomes trivial; packets can be send directly to the destination. Similarly, if the node is on a "stub" (i.e., it cm only communicate with one other mde) there is obviously only a single possible choice.

In the general case, however, packet network nodes are only partially interconnected with links and a packet must often be sent first to the neighbor which can best relay it onward to the destination.

One solution that works well in small networks (such as the ARPANET) is for each node to maintain a list of all other nodes in the network, giving the appropriate neighbor to which packets for each destination should be sent. (It is unspecified at the moment how these entries are determined.) As the network grows, however, each node's routing table will grow as well and the total amount of memory required at all nodes for routing tables will grow as the square of the number of nodes. Clearly, much memory could be saved if the list entries for nodes sharing the same "next hop" could somehow be condensed. This is possible if the addresses, instead of being arbitrary numbers, are related to the location of the node within the topology of the network.

At a node far removed from a given set of destinations, it is likely that the same neighbor would be used to reach any node within this set of destinations. If their addresses are "similar," in some sense, then it might be possible to "condense" these addresses into a single routing table entry.

5. Address Assignment Within IP

Bearing in mind the desirability of somehow encoding the topological location of a node into its address, I will now turn to the specific problem of address assignment within the ARPA Internet Protocol, IP.

An IP address field is 32 bits wide. IP addresses are further subdivided, primarily for administrative reasons, into three classes: A, B and C. The major difference between these three classes is the number of bits within this 32 bit field that may be assigned by the network administrator and how many are assigned by ARPA. This procedure is necessary if a network is ever to communicate with the existing ARPA Internet, since two sites might pick the same IP address unless there was some form of central coordination.

Thanks to the foresight of Hank Magnuski, KA6M, ARPA has assigned a Class A network number to amateur packet radio. This is a very valuable commodity, in that it fixes only the first byte of our addresses (to be decimal 44), leaving us the largest possible number of bits for our own use while keeping open the possibility of direct interconnection with the ARPA Internet. With the remaining 24 bits, we can address 16,777,216 nodes, easily enough to give every amateur in the world his or her own IP address if we allocate them efficiently.

Since AMPRNET is to be primarily a terrestrial radio network, it seems reasonable to encode a node's geographical location into its IP address. However, amateurs are distributed very unevenly throughout the world. Schemes that are based solely on geographic coordinates (e.g., grid squares), although aesthetically pleasing, are inefficient because they concentrate most of their address space over the poles, places with remarkably few amateurs.

Clearly a more efficient scheme is needed; one possibility is the binary tree. One way to illustrate this form of address assignment is with the game "Twenty Questions." Experienced players of this game know that the best strategy consists of asking questions for which "yes" and "no" answers are equally probable. In information theory, this corresponds to "maximizing the entropy of the source." For example, suppose that half of all the amateur packet stations in the world are in the United States. Then it would be reasonable to assign the first bit of the 24-bit address subfield to mean "US/non-US." Within the United States, one might determine that half of the packeteers are east of the Mississippi River and half are west, and so forth. Eventually, you reach a single "RF community" and you would assign the remaining address bits sequentially to the individual amateurs in that area.[1]

A major practical advantage of such a scheme is that the job of assigning addresses can be delegated to a hierarchy of organizations. An international organization (e.g., the IARU) would define only enough leading bits to uniquely designate each region or country in the world. National organizations within countries would then assign additional bits denoting regions within the country based on national concerns (i.e., the ARRL in the United States might handle this job based on American geopolitical boundaries). Other countries would have maximum freedom to devise their own national level addressing plans which might take into account unique national

[1] This is Huffman encoding, similar in principle to the popular CP/M programs "SQUEEZE" and "UNSQUEEZE." Huffman coding compresses files by recoding them with variable length "characters" assigned according to the relative character distributions in the file.

requirements or conditions. At the lowest level, an individual packet station would only have to contact his local packet radio coordinating body for a specific address assignment, and these "front line" organizations would have maximum flexibility in devising an allocation scheme suitable for the local environment. Individual assignments would then be forwarded back up the organizational hierarchy (or maintained in a "well known" directory server) so that the network as a whole may have convenient access.

Since node addresses in a given area have common prefixes, it is likely that a distant node would only have to keep a single routing table entry for a large collection of nodes. For example, a packet switch in New York would only have to maintain the information that all packets to nodes west of the Mississippi are sent to node X, thereby "condensing" half of the packet nodes in the USA into a single routing table entry in the New York switch.

Depending on the network topology and address assignments, routing table entries may consist of variable length prefixes. These prefixes might vary from 0 bits long (corresponding to a "wild card" or "default" routing entry to be used when on the end of a stub, for example) to a full 32 bits when used to describe an special entry for a specific address. The latter case would be useful to handle special cases, such as point-to-point connections via satellite, or a node whose entry cannot be condensed with any other existing entry.

There is no guarantee that a routing table would not be larger than average if a node were located near a boundary in the address scheme, e.g., the US/Canadian border. However, appears that such a scheme would reduce the AVERAGE size of routing tables in the network. More work on this problem is needed, particularly a comparative estimate of the routing table sizes and growth rates for a variety of address boundaries and population distributions.

It should be noted here that the issue of hierarchical address assignment is drawing much interest in the ARPA community. Currently, IP addresses are assigned according to a two-level hierarchy: a Class A, B, or C "network number" part and a host part. Assumptions are made by the rest of the Internet that all hosts within a network (even a Class A network with 16 million hosts) are capable of "direct" connectivity without (externally visible) routing.[2] Several people have proposed that extra, optional levels be added to the two-level hierarchy, and four RFCs (ARPA memos) have been released with various proposals over the last several months. As of this writing, the issue is not yet settled.

6. Implementing a Distributed Routing Algorithm

A variety of distributed routing algorithms have been used, with the ARPA Internet serving as one important example. I will now describe an algorithm often used in Internet Protocol networks; many variations exist on this common theme.

For each destination, a routing table entry contains the following information:

1. The hardware interface on which such packets should be sent.
2. The node to which such packets should be addressed at the link level (same as the destination if the hardware can reach it directly, the link address of an intermediate gateway otherwise).
3. A *metric* that indicates the route's "cost". (Cost is typically just a "hop count," although it might also indicate the relative speed or loading factor of a link).

Each node starts with its routing table containing only its directly accessible neighbors ("neighbors" could mean a node on the far end of a point-to-point link, or, collectively, all the nodes on a shared network such as an Ethernet or local packet radio channel). The metric for each neighbor reflects the cost of the link to that neighbor.

Each node periodically broadcasts its routing tables to all neighbors. When a node receives such a table broadcast from a neighbor, it examines each entry to see if it refers to a destination that was previously unknown in its own routing table, or reflects a metric that is lower than the value associated with an destination already in the node's routing table. If either condition is true, then the node inserts the new entry into its own table after incrementing the metric to indicate the "cost" of the link to its neighbor.

In this way, connectivity information "diffuses" throughout the network, and packets are routed along paths that favor the minimum cost or hop count (depending on the meaning of the metric). When a node receives a routing table entry from a neighbor that contains a metric equal to or higher than an entry already in its own table for the same destination, the node might decide to accept the new entry anyway, keeping it in reserve when the preferred route fails.

To assure rapid recovery from a link failure or network reconfiguration, nodes often "poll" their neighbors periodically to assure themselves that they're still there. If a poll fails for some period of time, all routing table entries referring to that neighbor are removed, and an attempt is made to disseminate this information to the other neighbors that are still up.

As mentioned, many variations and enhancements are possible on this basic theme. For example, it has been observed that "good news" (the availability of a new node or link) "travels fast," while "bad news" (the failure of a node or link) "travels slowly." The polling rate is clearly a tradeoff; frequent polls minimize the time needed to detect and recover from a failure at the expense of extra network traffic. Other schemes attempt to avoid polling by acting only when a local client "complains" that a given node appears to be inaccessible.

With certain algorithms, it is possible to have transient "routing loops," where packets are forwarded endlessly. Fortunately, this need not be catastrophic in a network based in IP because the "time to live" (TTL) field bounds the number of hops a datagram is allowed to make. As long as the updated routing information is allowed to propagate, however, the network will eventually recover.

One problem that can occur in such a distributed scheme is that a node may advertise, either accidentally or maliciously, that it can reach *every* other node in the network with zero cost. Other nodes may then be gullible enough to accept this information and decide to route every packet to the offending node which discards them, effectively crashing the network [3]. It may be necessary to establish "sanity checks" or encryption-based procedures to establish the authenticity and reliability of routing information.

[2] The ARPANET appears as a single Class A "local" network in the ARPA Internet, even though it spans the continental US and parts of Europe and the Pacific. Even though the ARPANET is not fully interconnected at the physical level, it does its own internal routing and thus appears as a fully interconnected network to the IP gateways.

7. Conclusions

I have only superficially scratched an involved topic, one that has been the subject of many books and learned journal articles. Nevertheless, I believe I can make several early recommendations that should ease the construction of our network and experimentation with practical routing algorithms:

- The use of a common datagram protocol at the network level (i.e., the ARPA Internet Protocol, IP) greatly simplifies the routing problem. Since a datagram network does not need nor guarantee absolute reliability, a wider variety of routing strategies may be considered. Routing protocols exploiting a datagram network's ability to efficiently broadcast routing information may also be used. Datagram networks can take full advantage of routing algorithms that dynamically balance link traffic. With IP, the sender always has the option of taking partial or full manual control of routing with the "source route" option, if desired.
- Network addresses should encode, in a hierarchical way that conserves address bits, the location of a node to reduce the amount of routing information that must be stored at each node and propagated throughout the network.
- A distributed routing algorithm should be used to avoid dependence on a central site and to allow maximum flexibility.
- Early emphasis should be made on establishing a standard protocol for the exchange of routing information (the ARPA Exterior Gateway Protocol, EGP, may be suitable for this purpose).
- Existing routing algorithms, particularly those used internally in the ARPANET and in the ARPA Internet should be investigated and tested to determine their suitability for widespread amateur use.

8. References

[1] Cerf and Kirstein, "Issues in Packet Network Interconnection," Proc. IEEE, November 1978 (special issue on Packet Communication Networks), p. 1386.

[2] L. Kleinrock, "Principles and Lessons in Packet Communications," p. 1320, same issue.

[3] R. Kahn et al, "Advances in Packet Radio Technology," p. 1468, same issue.

[4] L. W. Tymes, "Routing and Flow Control in TYMNET," IEEE Transactions on Communications, April 1981, page 392 (special issue on Congestion Control in Computer Networks).

[5] M. Gerla, "Controlling Routes, Traffic Rates and Buffer Allocation in Packet Networks," IEEE Communications Magazine, November 1984, p. 11 (special issue on Progress in Computer Communications).

[6] J. Hahn and D. Stolle, "Packet Radio Network Routing Algorithms: A Survey," p. 41, same issue.

[7] D. Clark, "Names, Addresses, Ports and Routes," ARPA RFC 814.

[8] W. Hsieh and I. Gitman, "Routing Strategies in Computer Networks," IEEE Computer, June 1984, p 46.

[9] J. Wescott, "Issues in Distributed Routing for Mobile Packet Radio Networks," IEEE

[10] D. Mills, "Exterior Gateway Protocol Formal Specification," ARPA RFC 904.

[11] R. Hinden, A. Sheltzer, "The DARPA Internet Gateway," ARPA RFC 823.

[12] P. Mockapetris, "Domain Names - Concepts and Facilities," ARPA RFC 882.

[13] P. Mockapetris, "Domain Names - Implementation Specification," ARPA RFC 883.