

Tanny

A PROPOSAL FOR ADDRESSING AND ROUTING IN THE INTERNET

David D. Clark
Laboratory for Computer Science
Massachusetts Institute of Technology

Danny Cohen
USC/Information Sciences Institute

1. Introduction

Within the internet, there is a need for both addressing and routing information as part of every packet. Using the distinction articulated by John Shoch (1), we take the address of a packet to be where the packet is destined, and the route, the specification of how the packet will get there. Currently, an internet header contains only an address, and the route is derived implicitly from that address. That addressing/routing strategy is quite suitable given the current internet topology, but two problems may arise as the internet continues to grow. First, unless the internet experiment is truncated artificially, it can be expected to continue, as has the ARPANET, for some period of time, in which case the number of networks involved may grow to exceed the size of the field allocated to number them. Second, as the topology grows more complex, it may not always be possible to deduce the desired or effective route from the address. This proposal attempts to address those two problems.

2. Addressing of Networks

The current internet header has space to name 256 networks. The assumption, at least for the time being, is that any network entering the internet will be assigned one of these numbers. While it is not likely that a great number of large nets, such as the ARPANET, will join the internet, the trend toward local area networking suggests that a very large number of small networks can be expected in the internet in the not too distant future. We should thus begin to prepare for the day when there are more than 256 networks participating in the internet.

To cope with this problem, we propose that the top level entity named in an internet address not be a single network, but, optionally, an aggregate of networks which we will refer to as a region. Thus, an address now begins with a region number, perhaps followed by a network number, in turn followed by network dependent fields. Large networks, such as the ARPANET, will presumably continue to be a region unto themselves. In fact, all of the currently existing networks in the internet can be viewed as regions, which means that no reimplementations are required if the concept of regions is accepted. However, in the future, as more and more nets enter the internet, we can, at our discretion, lump various networks together into regions. Put another way, a network can only join the internet by first joining a region.

In fact, the concept of a region was always available to the internet, although in an informal manner. The structure of a network address was unspecified except that it began with a fixed size field naming the network. It was always permissible to use the component of the internet address next after the network field to identify a subnet of the named network. Making more explicit this hierarchy of regions and networks is important because, as

we discussed above, the address is used not only as an address but as a basis for deriving a default route. We must thus consider how, using the addressing structure of regions and nets within regions, the address can be used to generate an effective and unambiguous route.

3. Default Routing

Currently, every gateway in the internet knows how to generate a route to every network. If the number of networks grows substantially above 100 or 200, the gateways can no longer be expected to understand this much information. One of the major purposes of the region concept is to control the amount of information which every gateway must know. This scheme would imply that an arbitrary gateway now only need to know how to reach every region, but need not be concerned with routing to the individual network within a distant region. Only for gateways connected to or completely inside a particular region would be necessary to understand how to route packets to all of the networks within the region.

Let us consider an example of how these regions might be used. There are already two local networks at M.I.T., with more on the way. These might quite properly be considered as one region. That would imply that a gateway located in England need not be concerned with the internal structure of the local networks at M.I.T. If M.I.T. were one region, such a gateway would merely need to know how to reach M.I.T. Only when the packet has reached a gateway connecting to one of the nets at M.I.T., would it be necessary to begin to worry about how to reach the correct local network. It is possible that deriving the route in this manner will not produce the optimum route; the packet may arrive at a gateway to M.I.T. which leads to the wrong local

network, but presumably this is an acceptable inefficiency. (We will return later to consider what might be done if it is not an acceptable inefficiency.) More generally, this example suggests that the success of deriving routes from addresses depends critically on the way networks are grouped together to form regions. A region containing two networks, one in California and the other in England, is not structured in such a way that effective routes can be derived from the network address alone. A naive gateway, routing its packet only to the region without regard to the particular network for which the packet is destined, may discover that the packet has been routed to the wrong side of the world. In fact, it is probable, in most cases, that regions should be connected. If not, it may be impossible to get a packet from one part of the region to the other, since the packet will have to leave the region in order to do so and may encounter a gateway which, not understanding anything about the network structure of the region, blindly sends the packet back to the part of the region from which it came. If, however, the appropriate gateways can be specially trained, there is nothing to prevent a disjoint region, and in particular circumstances it may be quite appropriate.

4. Source Routing

In the previous section, it was shown that an internet address of the form <region, network,...> could be used to derive a default route for a packet, much as a route is now derived by the gateways from the current internetwork address. Can we presume that this route will always be sufficient, even if it is not optimal? Unfortunately, in a few cases we cannot. First, it is easy to imagine circumstances in which the default route is hopelessly inefficient. A network may be connected by gateways to several

regions, even though for addressing purposes it is identified as belonging to one particular region. To send a packet to that region in order to reach the network may grossly lengthen the path of the packet. As we said before, this problem can be minimized by proper use of the region mechanism, but we cannot expect the region mechanism to be perfect. Second, it may be necessary to route a packet in such a way that it explicitly does not go through certain networks. For example, speech packets may be hopelessly delayed if they inadvertently travel by a network involving a satellite. It may be necessary to ensure that speech packets travel by some network with lower bandwidth but better response characteristics.

How can the internet come up with better routing information in those cases where it is required? In many cases, additional intelligence can be built into the gateways. What is required is that gateways not immediately adjacent to a region be prepared to understand the network field as well as the region field of packets destined to that region. This is analogous to something which happened in the telephone system, where a central office originating a phone call will usually examine only the area code in order to generate a route, but may, if it detects a particular area code, then further examine the destinations central office to discover if use of a particular optimized route is applicable. Building this additional routing knowledge into the gateways is very desirable in general, since it means that it will apply to all packets. However, we cannot expect all routing information to be embedded in the gateways. First, in order to solve the problem offered above of properly routing the speech packet, it would be necessary for the gateways to base their routing decisions on type of service information. This sounds like a rather complex decision for the gateways, especially since type of

service is not well understood. Second, network topology will change with time, and it is not reasonable to expect that all gateways will be constantly updated. Thus, we can expect the situation where only the originator of the packet has sufficient information to specify the proper route.

The solution which has been proposed in the past to cope with this is to replace the address in the packet with a route, called a source route since it is provided by the source of the packet. The disadvantage of having a route in a packet instead of an address is that the concept of an address is very useful one. For example, for accounting purposes it may be necessary to note the source and destination of a packet as it passes through a transit net. Clearly, it is desirable that the source and destination be uniquely identified for this purpose, something not easily done if the source and destination are specified only by a route. Thus, we propose that the address continue to be the primary piece of information in the packet, but that it be possible to include, in addition, an optional source route. This new internet option field will, if present, be used by the gateways instead of the default route which would normally be derived from the address.

We do not propose, in this report, a specific syntax for the option field. However, we make the following general observations. The source route should be structured in such a way that it need not contain more information than that required to augment the deficiencies in the default route. Thus, for example, it should be possible to source route a packet into a particular region, then specify that the default route should be used to get from there to some other region, and then specify additional explicit source information. In a later section we propose a particular semantics for source routes.

5. Migration

What is the relationship between the scheme proposed here and the current internet header with a fixed size address field? Happily, adoption of the addressing strategy involving regions together with the optional internet source route implies no immediate upheaval to packet formats or gateway code. Currently, every network is a region, and every gateway thus contains code for doing inter-region routing. Eventually, gateways will want to be modified as follows. When a region finally is defined which contains more than one network, then gateways inside that region will need to understand an additional component of the internet address. Presumably, since different regions may have a different number of networks, we can expect the size of the network field to differ between regions. Thus, unless gateway code is rewritten for different regions, it will be necessary to write code which can deal, eventually, with a variable size component of the address. The address itself, however, can reasonably be a fixed size, since it is merely an address and not a route. In fact, it seems that the field as specified for the current internet header is sufficient in size, although perhaps marginally so. Given that certain implementations of this header already exist, I would suggest that the correct field size in 3.1 be accepted unless strong complaints are heard from someone in the near future.

The next step in adopting this scheme, after the gateways have learned that for certain regions they must also look at some additional address bits, is to arrange that gateways selectively use this additional information, even when it is for a region for which they are not immediately adjacent. This facility, discussed above, can be used to provide more efficient routing than the default which would otherwise result from simplistic use of the address

information. The technical problem here is not implementation of additional gateway code for address manipulation, but rather the development of proper policies for dissemination of routing information so that the appropriate gateways are correctly informed of the routing decisions to make.

The third and final step in adoption of this scheme is the implementation in the gateways and hosts of the new internet option which specifies explicit source routes. Presumably, the general mechanism for dealing with internet option fields will already exist, so this is not a major upheaval of the code which parses an internet header. The only issue related to parsing is that, as the packet flows from gateway to gateway the source route will need to be modified to indicate which portion has already been used. This can either be done by physically rewriting the route into the option field, or by providing a pointer into the field as part of the option. The pointer field has the advantage that it does not destroy the route, so that it can be used to backtrack to the source, which is an important feature.

There are two reasons why it is desirable to be able to use a source route in reverse. First, the recipient of the packet may have no idea how to get back to the source. Second, and more relevant, if the route has been formed incorrectly, a gateway may receive a packet and have no idea how to forward it, because the next component of the route is nonsense. If that intermediate gateway cannot figure out how to get an error message back to the originating host, packets sent with malformed routes will appear to fall into black holes. It is very difficult to debug systems with black holes. Thus, reversability of routes is very important.

The need for modification implies the option should not be checksummed as part of an end-to-end checksum. The packet will also contain an address which

can be used by the eventual recipient to confirm that the packet is indeed destined for him. The address field can be checksummed, and under certain circumstances even encrypted.

6. Semantics of Source Routes

We view the internet as being composed of three physical entities: networks, gateways, and hosts, and one logical entity, a region, which is an aggregate of networks. The default route algorithm examines the region and, optionally, the network within this region, and selects from a table the appropriate network and gateway on that network to which to send the packet. Thus, if the route followed by a packet were written out in advance, it would be an alternating concatenation of network names and gateway names, finally terminating with a network name followed by a host name. For symmetry, one should also precede the source route with the name of the originating host, to allow the route to be used in reverse. Thus, an initial structure for internet source routes might look as follows:

H,N,G,N,G,N,G,N,...,G,N,H

where H is a host identifier,

N is a network identifier,

G is a gateway identifier.

Each gateway, on receiving this route, finds his position in the route using the pointer into the route, updates the pointer to indicate the next gateway which is to receive the packet, and then routes the packet through the specified network to the next gateway.

The source route as shown above always specifies the complete route. For many cases this degree of specificity is not necessary. For example, once a

packet has been routed part of its way, the default route may then be effective. This generalization means that instead of specifying the particular network which is used as we pass from one gateway to the next, a default route can be used between two particular points in the route. Thus, we propose a more general form of a source route, as follows:

`<source route>:=<source><step>...<destination>`

The source route takes the packet from the source to a sequence of gateways to the destination. The progress from each of these specified points to the next is a step.

`<step>:=<explicit route step>|<default route step>`

`<explicit route step>:= RING`

If we are concerned with the exact route between one gateway and the next, we specify the step in this form, naming the particular network that is to be used. A sequence of explicit route steps can thus connect two gateways not immediately adjacent.

`<default route step>:=<starting net><ending net>G`

`<starting net>,<ending net>:= RN`

If the particular route to the next specified point is not critical, then the default route step is used. The originating gateway will generate a route to the network addressed by <ending net>. That net may be any distance away in the internet; intermediate gateways in this step will again generate the default route from <ending net> until the specified gateway G is reached, which will end the step. <starting net> is required so that the route can be

used in reverse. It must specify the network address of the gateway, originating this step.

`<destination>:= RNHTRNRNH`

The destination is in fact the final step and as such can be either explicit or default. Thus it has two forms, with the interpretation of the step with the equivalent form.

Note that while the string representing the source route is generated as a sequence of "forward" steps, there is another grammar that generates the same strings as a sequence of "reverse" steps. Also, in the <explicit route step>, the intervening network is identified using a full network address, including the region. In fact, any shorthand network identifier can be used, so long as it is unambiguously interpretable by the gateway at each end of the step.

7. Uniqueness of Names

Hosts will often be attached to more than one network. Thus, hosts may have more than one internet address. As long as the only routing algorithm is default routes based on addresses, there will be a strong desire to use these additional names to generate better routes. While this is fine in the short run, functions such as accounting will be easier to implement if hosts have a single unique address. To this end, when the route option is implemented we expect that it will be appropriate to address a host in only one way, and specify a route additionally.

REFERENCES:

- (1) Shoch, J.F., "A Note on Inter-Networking Naming, Addressing, and Routing," Xerox Palo Alto Research Center, Palo Alto, California, INTERNET Notebook, Note No. 19, Section 2.3.3.5, January 1978.