UTILIZING INTERNET ROUTES AS EXPRESSWAYS THROUGH SLOW NETS

IEN #147
PRTN #290
Radia Perlman
June 1980

INTRODUCTION

Some networks in the catenet are very slow relative to others. An example is a very large packet radio network, which might be 100 PR hops from end to end.

Distributed throughout the slow net might be several gateways to a faster net. It would be desirable to utilize the internet to quickly route packets through the slow net. So in the particular example, rather than delivering a packet through 100 PR hops, it would be desirable to deliver the packet to the nearest gateway in the PR net, and have that gateway deliver the packet, through the internet, to a gateway close to the destination of the packet.

This problem was presented by Dr. Vinton Cerf at the May 1980 Internet meeting. This paper is an approach to solving the problem on a specific case basis, and an argument for solving it that way rather than as a generalized internet routing problem.

WHY IT IS NOT THE CATENET'S JOB

The catenet is a network of nets. Its job is to deliver a packet as efficiently as possible to the correct net. It should not have to worry about the internal structure of a network to the extent of knowing which hosts are closest to which gateways.

It is reasonable for the catenet to know enough about nets to take into account their properties when calculating internet routes. For instance, according to whatever cost function gateways are using to calculate catenet distances, one net can be many times the cost of another. And if gateways wished to calculate distances to their neighbor gateways, and if there were three gateways G1, G2, and G3 on a net N, the link G1-G2 could have a different cost from the link G1-G3. However, the catenet should route to destination networks based solely on the most efficient internet route to that network, and let the destination network handle delivery to the host on the network as it sees fit.

HOW A SPECIFIC NETWORK CAN USE A NEARBY FAST NET

The previous section does not imply that the internet cannot be utilized as an express route through a slow net. What it says is that it is the job of a local net, rather than the catenet, to utilize the capability. And it is not difficult to design this into a net.

```
          net S        I           net F
    ________________________________________
    I                  I
    I     A            I
    I     .            I
    I     ...N1--G1S-G1-G1F..
    I                  I    .
    I                  I    .
    I                  I    .
    I                  I
    I     .N2--G2S-G2-G2F..
    I       .          I
    I       .          I
    I       .B         I
    I                  I
    ________________________________________
```

We are assuming a slow net with a few gateways to a faster network. Call the slow net S, and the fast net F. Assume two gateways, both on both S and F. The addresses of the first gateway are G1S and G1F, with G1S being G1's local address on net S. The addresses of the second gateway are G2S and G2F. Assume G1S is connected to node N1 on net S, and G2S is connected to node N2 on S. (I am assuming a network like the ARPANET or the PRnet, in which hosts are not nodes on the net, but are attached to nodes on the net.)

Node A on net S wishes to communicate with node B on net S. A and B are distant in net S, and A is close to G1, and B is close to G2.

A can communicate with B via A's route to G1, G1's route through net F to G2, and G2's route to B in S by prepending to the packet an internet header with a one level source route, and sending the packet to gateway G1. The one level source route is the address on the fast net of the destination gateway, i.e., <net F, host G2F>. The destination in the internet header is the internet address of the destination of the packet, i.e., <net S, B>. When G1S receives the packet, it will route it to internet address F,G2F. Since that internet address is on net F, G1 will route the packet through net F to G2. When G2 receives the packet, it will advance the source route and forward the packet to internet address <S,B>. So the packet travels A=>G1 through S, G1=>G2 through F, and finally G2=>B through S, as desired.

The source route is essential because if G1 is given a packet with an internet address on net S, it will just throw the packet back into net S directly, because no matter how cheap its link to G2 is through net F, it cannot be cheaper than 0, which is G1's distance to net S, because G1 is directly connected to S.

In most network designs, you would want most of net S to be unburdened with knowledge of the internet, neighboring fast net, internet headers, etc. Thus a reasonable design is that nodes on S have some local net link-link protocol, and some routing information about links that gets passed around. The nodes directly attached to the gateways, N1 and N2, utilize the foreign net as a "link" between themselves. To do this, they must encapsulate local net packets for transmission across this "link", within internet packets, since the "link" is actually the internet.

N1 adds an internet header to the packet with a one-level source route being the internet address of the second gateway on the fast net, and internet destination address <S,N2>, and passes the packet off to the gateway. N2 strips off the internet header and continues the packet along the local net route. The rest of the nodes on S can be ignorant of just what mechanism is employed when N1 and N2 communicate. All they know is that N1 and N2 report that they can get packets to each other with some miraculously low cost. They believe N1 and N2, and find when they compute routes that many local routes will take advantage of that high speed link. N1 and N2 do not have to know very much about net F. For N1 to send a packet to N2, all it needs to do is send its attached gateway the packet with a prepended internet header with single-level source route <F,G2> and destination internet address <S,N2>. N2 must be able to strip off the added internet header and continue the packet in the local net as if the N1-N2 link were an ordinary link in that net. (In IEN #140, "Mutual Encapsulation of Internetwork Protocols", the authors give another example of using one net as if it were a link in another net -- the encapsulation of PUPs in IP datagrams.)

There are some nets in which only hosts are addressable, not nodes ("imps"). It is important, however, as discussed above, to have N1 be able to put N2 as the internet destination. If N1 merely put B (the destination host) as the internet destination, N2 would not know to strip off the internet header. Packets should not arrive looking different depending on what path they take, so it is at least ineligent to have some packets arriving at B with internet headers. Worse yet, B might not know about internet headers -- it might be a host that only communicates with other entities on the local net. And another problem is that if the route through the net traversed several express hops, several internet headers would be prepended, one after another, onto the packet. So the answer to this problem is that N1 must be able to have as the internet address N2. The code in N2 to strip off internet addresses can be thought of as a virtual host attached to N2, so there should be no problem assigning it an internet address.

THE MULTISTATION PACKET RADIO NET

For those of you uninterested in the packet radio net, skip this section.

Briefly, the packet radio network consists of nodes called PRs, each of which has a unique ID. Hosts, terminals, etc., are known as devices, and can be attached to any PR, and unplugged from a PR and moved to another PR without impact since device/PR correspondence is dynamically bound by the network. Devices also each have a unique ID.

Routing is calculated by special devices known as stations. A route consists of a sequence of IDs, currently all of which would be PR IDs. Presumably each PR in the route is in direct radio connectivity of the next PR along the route.

In the current design, station-station links are assumed to be ordinary packet radio multihop routes. Suppose that some of the stations were directly connected to each other with some sort of high speed, high bandwidth lines. The only modifications necessary to the design of the net to utilize these lines are that:

1) a station ID can appear in a route
2) a PR checks to see if the next ID in the route is its directly attached station. If so, rather than broadcasting the packet, it sends the packet to its attached station.
3) a station attached to other station(s) via high speed lines must be able to handle a packet by looking to see the next ID in the route and handing the packet to the correct next station, or to its attached PR (if the next ID was the station's attached PR)
4) stations would take these high speed lines into account in calculating routing in their subnets, and in calculating their distance to their neighbor stations in the station-net.

Suppose instead that there are gateways (which are devices) to a faster net interspersed throughout the PRnet. Assume the gateways are not in stations. The modifications necessary to the PRnet design to utilize them are:

1) a device ID must be able to appear in a route
2) a PR must check to see if the next ID in the route is its attached device. If so, the PR must prepend an internet header with single level source route being <net F (a constant assembled into the PR's program), host X> (where X is the next ID in the route (and if the internet local name doesn't fit into a PRnet unique ID, use two PRnet IDs, the

next 2 hops, to calculate what to put into X in the internet header)), and internet destination <PRnet (a constant assembled into the PR's program), ID of following hop of route>.

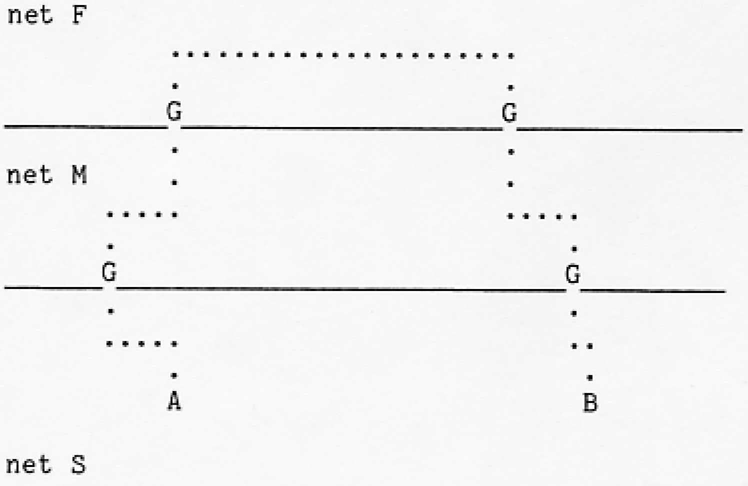
3) The PR receiving the packet from its attached device should strip off the internet header and continue the packet along the PRnet route.
4) Stations must know the gateway IDs so that they can recognize them as such when they are reported

Suppose instead that the stations themselves are gateways, in that they are themselves connected to both nets. Then there is a choice of how to treat the design. It can either be treated as the standard gateway design just covered, or it can be treated as the slightly simpler case of direct lines between the stations, where the "direct lines" are actually routes through the fast net. In other words the stations can use the internet as their "direct link", or the fast net alone.

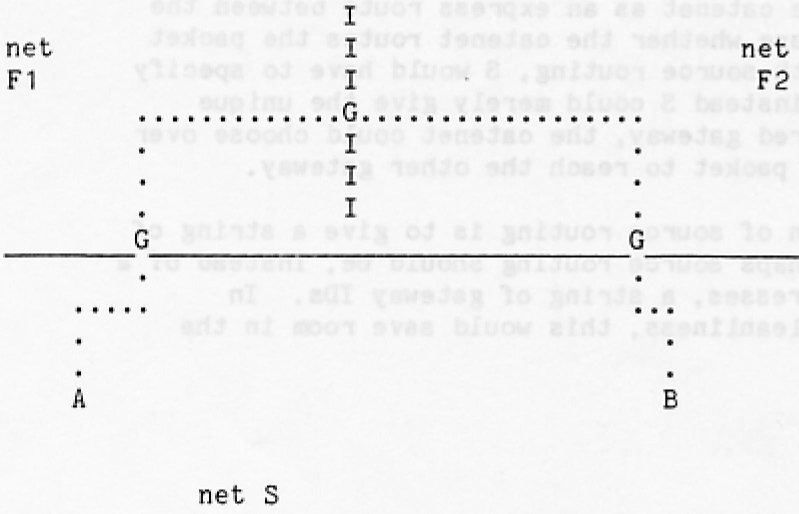
COMPLICATIONS?

In this section, I will present some variants of the basic problem, and show how they are not complications of the basic design.

The first variant is when there are three nets, S (Slow), M (Medium), and F (Fast). There are a few gateways between S and M. There are a few gateways between M and F. There are no gateways between S and F.



Routing in net M is exactly as described above. Net M can make use of the gateway links through net F. Routing in net S is also exactly as described above. Net S can make use of the gateway links through net M. S does not know anything about net F. It routes packets through net M, and is unaware that M cleverly makes use of net F. So a packet between two hosts A and B on S might travel to a nearby S-M gateway, through M to an S-M gateway near B, and then to B. The path through M between the two S-M gateways might travel through net F for some part of the way. The important thing is that none of the nets needs to know exactly how the other net routes packets. The strange thing is that each time this happens a new internet header will be prepended to the packet. So in this example, while the packet is travelling through net F, there will be two levels of internet header on the packet. One might prefer instead to have a two-level source route on the packet. The reason I don't think that would be the right design is that it requires net M to understand that the packet travelling through its local net is an internet packet. I am assuming that while a packet is travelling through net M it has a local M-net header on it, and as far as net M is concerned, the internet header, which follows the local header, is just data. So entities on M would not be smart enough to strip off the local header and manipulate the internet header by adding more levels of source route.



The next variant consists of three nets, S (Slow), F1 (fast net 1), and F2 (fast net 2). Assume a gateway between S and F1, a gateway between F1 and F2, and a gateway between F2 and S. Assume these are the only gateways, and the two gateways on S are distant from each other on S. Assume a host A on S close to the S-F1 gateway and a host B on S close to the S-F2 gateway. S should route packets between A and B through net F1, then net F2, and then to B.

S handles this case exactly as all the others. The route through S goes first to a node attached to the S-F1 gateway, and that node prepends an internet header with the source route being the F2 address of the S-F2 gateway, and the internet destination being the S node attached to the S-F2 gateway, and gives the packet to the S-F1 gateway. The S-F1 gateway routes the packet as desired. So net S treats the S-F1 gateway as if it really were an S-F2 gateway in the original example.

AN ALTERNATIVE TO SOURCE ROUTING FOR THIS CASE

This design does not really utilize the general capabilities of source routing. Assuming source routing is implemented, it provides a convenient method of implementing the catenet as an express route through a slow net. However, it is probably cleaner to look at the solution in a slightly different way.

With source routing, the packet header specifies a full internet address as an intermediate destination. In this case, the intermediate address is always a gateway. Assuming gateways were given unique IDs (as opposed to multiple internet addresses), it would be cleaner to specify a gateway ID rather than a <net, host ID> pair for a gateway. An example where it is cleaner is having three nets, S (slow), F1 (fast net #1), and F2 (fast net #2). Suppose two gateways, each connected to all three nets. When net

S wishes to utilize the catenet as an express route between the gateways, S does not care whether the catenet routes the packet over F1 or F2. But with source routing, S would have to specify one or the other. If instead S could merely give the unique gateway ID of the desired gateway, the catenet could choose over which path to send the packet to reach the other gateway.

If the only utilization of source routing is to give a string of gateway addresses, perhaps source routing should be, instead of a string of internet addresses, a string of gateway IDs. In addition to possible cleanliness, this would save room in the header.

CONCLUSIONS

The catenet should be protected from knowledge of how local networks route packets. All the catenet should know is the cost of each gateway-gateway link, not the mechanism of packet travel on the link.

Local networks should be mostly protected from knowledge of the internet. A network that was not originally designed as part of the catenet should be able to be connected to the catenet merely by adding a gateway to that net.

The problem of utilizing faster neighbor nets can be solved in each net in which the opportunity exists in a not very costly or complicated manner, by adding sophistication only to the nodes directly connected to gateways. The rest of the local network can be oblivious of what is going on.