OP2: A richer network abstraction? (revision)

Choosing the network path of packets in the current Internet is not possible, but if we wanted to change this to allow it to processes we could implement that without violating layering. The proposed solution is to add a layer above the network layer that would provide this service, that I will call the *path layer*. This layer would support path-based addressing on top of IP, defined in more details below. Only a subset of routers would need to implement it, forming an overlay on the network. These specialized routers (I will call them *path routers*) would monitor and exchange information about the latency and throughput of their local traffic, in the same manner as RON nodes [1] would. Path routers at the edge of the network would then decide on a path via other such routers, using this information, to satisfy the requested properties that the packet specifies.

The path-layer's header would contain a `viaPath` field used by the node addressed in the IP packet destination to determine how to relay the packet. This idea is inspired by WRAP, the protocol used for the TRIAD project [2]. When processing the IP packet, a path router looks at the next hop on the path and routes it accordingly. The first path router encountered (maybe the gateway router, maybe further) would be a special router allowed to look at the transport-layer header and responsible to compute the path. Indeed, the transport-layer header should be modified to include flags telling the specifications the path should meet, but this would only be used by the first path router, all subsequent ones only look at the `viaPath` field. We could imagine deciding on a set of geographical regions (e.g., continents, or possibly more fine-grained), each corresponding to a bit, set to 1 if the path should not cross it. There would in addition be fields about performance, e.g., maximal throughput, minimal latency or "don't care". This implies to also change the interface between the transport and application-layer to offer the same choices to processes, such as `sendWithMinLatency()` and so on. These choices would be decided by e.g., combining the opinion of a small committee of experts, and the results of an open poll to users.

This solution is the cleanest since it does not require too many changes in the current layers. An alternative solution would be to allow the user to decide on a precise list of hops, but that would require to break the current abstractions, and expose the internal details of the network. Nevertheless, a process cannot exactly choose the network path of its packets with this solution, but rather select some desirable properties about it. Additionally, its quality depends on the demand, since if only a few processes were to use that functionality, all others would still need to pay the cost. This cost consists mostly of the increase in packet size, and the fact that path layer routers need to extract the path-layer header of all packets, which adds useless processing overhead if no path is specified.

References

[1] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris, "Resilient Overlay Networks", 18th ACM Symp. on Operating Systems Principles (SOSP), October 2001.
[2] D. R. Cheriton, and M. Gritter, "TRIAD: A New Next-Generation Internet Architecture", January 2000.