

UFO: A Resilient Layered Routing Architecture

Yaping Zhu ^{*}, Jennifer Rexford ^{*}, Andy Bavier ^{*}, Nick Feamster [†]

^{*} Princeton University [†] Georgia Tech

Today’s “one size fits all” routing system scales well, but at the expense of availability. Routers do not respond automatically to congestion, forcing traffic to traverse overloaded paths even when better paths exist. Transient disruptions during routing-protocol convergence [9, 16] degrade the performance of interactive applications like Voice over IP (VoIP) [3, 8, 11] and online gaming. In addition, today’s routing protocols do not satisfy the diverse availability requirements of different applications. For example, some applications may prefer high-throughput paths while others prefer low loss and low delay, but today’s routing protocols route all traffic over the same set of paths. In this paper, we propose a routing architecture that scales well and provides high availability for diverse applications by combining ideas from conventional routing protocols and overlay networks.

Our architecture is motivated by the success of overlay networks at providing reliable communication services, albeit at limited scales. For example, a Resilient Overlay Network (RON) [5, 1] uses a small collection of hosts to form a topology where each overlay link traverses one or more hops in the underlying network. RON reacts more quickly than the underlying network to changes in network conditions; unfortunately, it does not scale to a large number of hosts, and the resulting paths can sometimes be inefficient. Traversing intermediate hosts adds delay and consumes bandwidth and processing resources [6]. Quickly masking failures requires monitoring the underlying paths with frequent probes, which does not scale to large overlay topologies [4, 12].

This paper argues that, to scale well *and* provide high availability, the routing architecture should consist of multiple layers. Overlay routing should be viewed as part of the routing system itself, not as “just another application”. The key questions that a multi-layer routing system must address concern: (1) What *functions* should be placed at each layer of the routing system? and (2) What *interfaces* should these layers expose to one another? To explore these questions, we present a two-level routing architecture that supports high availability by reacting quickly to failures by borrowing some design principles and mechanisms from overlay networks. We believe that the lower layer, the “underlay”, should support the following two functions:

- *Direct control over forwarding table entries.* We believe that routers should provide data-plane support for forwarding packets over “overlay” links, at the behest of the “overlay” control plane. This enables flexible

control over the end-to-end paths, without the efficiency drawbacks (e.g., higher latency and wasted bandwidth) of traditional overlay networks. Routers could also provide control-plane support for overlays by virtualization [10, 2, 13, 14, 15, 7].

- *Explicit notification about changing network conditions.* We believe that routers should notify the overlay about critical changes in the properties of an overlay link. Explicit notifications improve the efficiency of reactive routing at the “overlay” layer without compromising scalability. This raises important questions concerning who to notify about what kinds of events, and how to design a scalable notification system.

We call our system “underlay fused with overlays” (UFO) to emphasize the *cross-layer* nature of our design. In practice, we envision two main deployment scenarios. First, ISPs could view the two-layer design as an effective way to offer highly-available communication services to their customers. Second, ISPs could host third-party overlay services, offering them extra support for packet forwarding and explicit notification, as an extension of their server hosting business.

References

- [1] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris. Resilient overlay networks. *Proc. ACM SOSP*, 2001.
- [2] A. Bavier, N. Feamster, M. Huang, L. Peterson, and J. Rexford. In VINI veritas: Realistic and controlled network experimentation. *Proc. ACM SIGCOMM*, September 2006.
- [3] C. Boutremans, G. Iannaccone, and C. Diot. Impact of link failures on VoIP performance. *Proc. NOSSDAV Workshop*, 2002.
- [4] Y. Chen, D. Bindel, H. Song, and R. H. Katz. An algebraic approach to practical and scalable overlay network monitoring. *Proc. ACM SIGCOMM*, 2004.
- [5] A. Collins. The Detour Framework for Packet Rerouting. Master’s thesis, University of Washington, 1998.
- [6] J. Jannotti. *Network Layer Support for Overlay Networks*. PhD thesis, Massachusetts Institute of Technology, 2002.
- [7] S. Karlin and L. Peterson. VERA: An extensible router architecture. *Computer Networks*, 2002.
- [8] N. Kushman, S. Kandula, and D. Katabi. Can you hear me now?! it must be BGP. *ACM Computer Communication Review*, 2007.
- [9] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian. Delayed Internet routing convergence. *Proc. ACM SIGCOMM*, 2000.
- [10] K. Lakshminarayanan, I. Stoica, S. Shenker, and J. Rexford. Routing as a service. Technical report, University of California, Berkeley, 2006.

- [11] A. Markopoulou, F. Tobagi, and M. Karam. Assessment of VoIP quality over Internet backbones. *Proc. IEEE INFOCOM*, 2002.
- [12] A. Nakao, L. Peterson, and A. Bavier. A routing underlay for overlay networks. *Proc. ACM SIGCOMM*, August 2003.
- [13] L. Peterson, T. Anderson, D. Culler, and T. Roscoe. A blueprint for introducing disruptive technology into the Internet. *Proc. of HotNets*, 2002.
- [14] L. Peterson, A. Bavier, M. E. Fiuczynski, and S. Muir. Experiences building Planetlab. *Proc. OSDI*, 2006.
- [15] J. S. Turner. A proposed architecture for the GENI backbone platform. *ANCS*, 2006.
- [16] F. Wang, Z. M. Mao, J. Wang, and L. Gao. A measurement study on the impact of routing events on end-to-end Internet path performance. *Proc. ACM SIGCOMM*, 2006.