Spoorploeter: Anycast Mapping with Fine Resolution

Joe Abley

***Abstract*—Anycast is a technique used to distribute services in multiple places across a network by the deployment of autonomous or semi-autonomous service nodes, commonly used with a variety of protocols to reduce transaction latencies, localise unwanted traffic and increase overall service availability. The extent to which these objectives are met depend on the combination of routing policies of the service operator, the operators of the networks that contain the service's clients and all the networks between them. While it is possible to make inferences about the routing policies of adjacent and distant networks and to influence the path service traffic will take, this is far from an exact science; an iterative, trial-and-error approach is usually required, often involving production traffic.**

Verfploeter is the name given to an effective approach for mapping anycast services that scales quickly to very large numbers of vantage points without requiring new test infrastructure to be deployed; its use has been demonstrated with 3.8M vantage points (VPs), considerably more than the VPs available (for example) in RIPE Atlas. Verfploeter makes novel use of ICMP echo, as used in the network troubleshooting tool "ping".

We propose spoorploeter, which extends Verfploeter by introducing variations in the IP time-to-live (TTL) header parameter, such that intermediate routers also function as VPs. This approach simultaneously extends the number of effective VPs available with the same initial configuration and adds more fine-grained detail, such that inflection points in anycast node selection within individual networks can be inferred. We show how this additional detail can be used as an effective tool in the diagnosis of anycast service performance, and also in the construction of effective peering and transit policies that can improve performance and reliability.

***Index Terms*—Internet topology, IP Networks, Routing protocols**

I. Introduction

Anycast [1] routing is an approach to service distribution that mimics the topology of a distributed, multi-homed network by the deployment of individual, autonomous service elements across a wider network. Since each service element behaves consistently with all the others, the appearance from a client's perspective is of consistent, single service. The advantage to such an approach from a service provider approach include performance, since the service appears to be deployed locally for a larger proportion of clients than would otherwise be the case, and reliability, since the service as a whole is at most partially reliant on the availability of any individual service element.

Anycast is properly a deployment technique, and not a property of the routing system or any particular IPv4 or IPv6 address. We use the phrase "anycast service" to denote a service, intended for use by clients, that has been deployed using anycast. The term "anycast address" is used to refer to an address associated with such an anycast service. An anycast address is said to be covered by an anycast prefix, corresponding to the unit of routing information that is provided to an adjacent network that would invite traffic addressed to the anycast address to be delivered to a particular "anycast site", a particular deployment of service infrastructure with a consistent routing policy with respect to its surrounding, adjacent networks.

This paper is concerned with anycast as used to deploy services on the global Internet, using BGP [2] to originate anycast prefixes, although the experiments described could be adapted to be used in other environments with other routing protocols.

Traceroute was a tool implemented by Van Jacobsen for exposing addresses bound to interfaces on the individual routers involved in forwarding a probe packet from source to destination. Jacobsen's approach was to originate probe packets as UDP datagrams [3] with the destination address field in the IPv4 [4] header set to the address of a target system. Subsequent probe packets to the same target are sent from a single vantage point address with the TTL field in the IP header set to monotonically increasing values. When a router between the source and destination receives an IP datagram with a TTL field of zero, forwarding of the probe packet ceases and an ICMP type 11 message [5] is returned to the vantage point sourced from an address on the router, indicating that the probe packet's TTL expired en route to the destination. Each such time exceeded message includes the IPv4 header plus 8 bytes of payload from the beginning of the probe packet, which can be used to associate particular responses with particular measurements.

While the original Van Jacobsen traceroute used UDP probe packets, other implementations (most notably Microsoft Windows) have successfully used ICMP echo requests as probes. Type 11 ICMP messages in IPv4 have analogues in IPv6 [6] in the form of ICMPv6 type 3 messages [7]; in IPv6 additional information may be returned in the ICMPv6 type 3 response since the payload is specified as "as much of the invoking packet as possible without the ICMPv6 packet exceeding the minimum IPv6 MTU"'. Since both ICMP and ICMPv6 support the core mechanisms required in this paper, that of echo request and response and time exceeded notifications, we consider them equivalent and do not distinguish between them in descriptions of our methodology (although since the IPv4 and IPv6 internets are distinct networks, we measure them separately).

Verfploeter [8] is a novel approach that can be used by the operator of an anycast service to measure the catchment of individual

anycast services independent of production service traffic. ICMP echo requests (and their IPv6 analogues) are originated towards a set of target systems from one or more anycast sites, with source addresses set to the anycast address of a service provided at each site; the responses that are triggered are routed to the anycast site whose catchment includes the originator of the response. Each system in the target set that responds provides a new vantage point for the catchment mapping without requiring dedicated probe infrastructure to be deployed, allowing arbitrarily dense data from which to construct a map.

We extend Verfploeter to Spoorploeter in much the same way that traceroute extended ping – that is, we originate Verfploeter's probe packets with variations in the IP time-to-live header parameter, in order to trigger time exceeded messages. This extension effectively extends the number of VPs available to include every responding router on the path between probe initiator and target.

II. Related Work

The Verfploeter paper described various related mapping techniques, including the use of open resolvers [9], the use of physical measurement platforms such as RIPE Atlas and PlanetLab, the use of traffic and log analysis to perform measurement based on production service traffic and specific measurement exercises that have targeted specific, prominent anycast networks such as those support root DNS nameserver infrastructure.

Services which incorporate dynamic processing capabilities in the client (those delivered through modern web browsers with client-side scripting capabilities, for example, or dedicated applications on mobile devices) have additional opportunities to perform representative measurements that can inform a service operator of topological inefficiencies. For example, an application running on a user's phone might be able to detect that the instance of an anycast service it sees is further away than a closer one it knows to exist; however, the capabilities of the client environment do not always facilitate further diagnosis to colour the observation with useful detail, or in the case of mobile devices, there may be a sensitivity to generating network traffic that might cost the user money. Using your own users as vantage points for measurement also only helps to the topological extent that you have users; it does not provide any ability to plan ahead for performance in the networks where new clients might appear.

Spoorploeter shares the same advantages that Verfploeter had over these other approaches, and brings the additional benefits of intermediate routers as additional vantage points.

III. Spoorploeter: Much More Splatter

Describe the separation between collection and probe transmission. Describe the probe formats, and considerations around the sending frequency and the maximum TTL. Describe the different approaches of doing a one-off measurement, perhaps repeated over time, and a continuous measurement. Describe the requirements for data collection and summarisation. Introduce toolsets that will have been written in order to carry out these various things.

IV. Experiments

Describe the use of Spoorploeter in mapping an Afilias anycast service.

V. Using Spoorploeter for Network Planning

Describe how we can use the results from this kind of measurement to answer questions about network planning, given knowledge about the AS topology extracted independently from the routing system.

References

[1] J. Abley and K. Lindqvist, "Operation of anycast services," Internet Requests for Comments, RFC Editor, BCP 126, December 2006. [Online]. Available: http://www.rfc-editor.org/rfc/rfc4786.txt

[2] Y. Rekhter, T. Li, and S. Hares, "A border gateway protocol 4 (bgp-4)," Internet Requests for Comments, RFC Editor, RFC 4271, January 2006. [Online]. Available: http://www.rfc-editor.org/rfc/rfc4271.txt

[3] J. Postel, "User datagram protocol," Internet Requests for Comments, RFC Editor, STD 6, August 1980. [Online]. Available: http://www.rfc-editor.org/rfc/rfc768.txt

[4] ——, "Internet protocol," Internet Requests for Comments, RFC Editor, STD 5, September 1981. [Online]. Available: http://www.rfc-editor.org/rfc/rfc791.txt

[5] ——, "Internet control message protocol," Internet Requests for Comments, RFC Editor, STD 5, September 1981. [Online]. Available: http://www.rfc-editor.org/rfc/rfc792.txt

[6] S. Deering and R. Hinden, "Internet protocol, version 6 (ipv6) specification," Internet Requests for Comments, RFC Editor, STD 86, July 2017. [Online]. Available: http://www.rfc-editor.org/rfc/rfc8200.txt

[7] A. Conta, S. Deering, and M. Gupta, "Internet control message protocol (icmpv6) for the internet protocol version 6 (ipv6) specification," Internet Requests for Comments, RFC Editor, RFC 4443, March 2006. [Online]. Available: http://www.rfc-editor.org/rfc/rfc4443.txt

[8] W. B. de Vries, R. de O. Schmidt, W. Hardaker, J. Heidemann, P.-T. de Boer, and A. Pras, "Verfploeter: Broad and load-aware anycast mapping," in *Proceedings of the ACM Internet Measurement Conference*, London, UK, 2017. [Online]. Available: http://www.isi.edu/%7ejohnh/PAPERS/Vries17b.html

[9] X. Fan, J. Heidemann, and R. Govindan, "Evaluating anycast in the domain name system," in *Proceedings of the IEEE Infocom*. Turin, Italy: IEEE, Apr. 2013, pp. 1681–1689. [Online]. Available: http://www.isi.edu/%7ejohnh/PAPERS/Fan13a.html