Experimentation on Live DNS with DIINER

Wes Hardaker and John Heidemann
USC/Information Sciences Institute

As a central protocol in the Internet, any change to the DNS protocol must be carefully evaluated. Evaluation should consider real-world traffic and constraints. In recent years, experiments have been essential to show that TCP and TLS have manageable performance costs [5] and how QNAME minimization changes traffic [1].

DNS experiments today typically happen at small scales in researcher's labs. Tools exist to support experiments, with artificial traffic generators [2, 3] and trace playback [4**?**], but it still still the burden of the researcher to show that those tools and the laboratory setup capture real-world traffic and operational constraints. **[other tools? —johnh 2020-07-02]**

Our goal is to assist researchers in carrying out realistic experiments with the DIINER experimental infrastructure. It will support experiments on real-world data, or on real-world data that is mutated to test some variation on traffic. Alternatively, one should be able to replay a stored trace repeatedly. The test platform should provide hardware and software identical or similar to an operational deployment. Finally, in the long run we plan to allow comparison of experimental answers against current production software to check for correctness and to compare performance.

Evaluating Against Real-world Traffic in DIINER Experimentation Today: Our approach to support experimentation on real traffic involves *Parallel Resolution Evaluation*, as shown in Figure 1. A typical deployment has a load balancer that splits data out to several backend servers. To support experiments, we make a copy of a fraction of the live traffic and send it to a copy of a production node running experimental software, or to experimental hardware. The output of these queries is not sent to users, but may be examined as part of the experiment. With dedicated experimental hardware and a live query stream we can compare the experiment to real-world operations to evaluate latency, memory use, or other factors of performance. In addition, we plan to compare results to evaluate correctness, as we describe below.

This setup provides an experimental stream of real-world data, allowing safe experiments over realistic traffic. In addition to traffic diversity, some experiments require rare events or controlled, repeatable traffic. We can handle both of these cases with replay of saved traffic. We propose to use tools like LDPlayer [4] that can replay root-server and other traffic at scale with good timing accuracy [5]. Replay of a single event allows experiments on DDoS events, and repeated replay allows controlled experiments to gather statistics or compare alternatives.

In addition to replay of saved traffic, we can also run artificial DNS traffic generators such as dnsperf [2] and dnsblast [3].

Mutating queries: Often experiments are designed to explore not what *is*, but what *will be* or what *might be*. For example, we have previously examined root-server performance if all traffic used DNS-over-TLS [5]. The key component to test alternate-future scenarios is a *query mutator*—a tool that takes a real-world query stream and modifies it in some way.

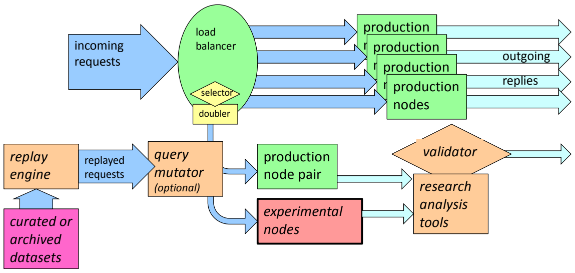

Figure 1: Parallel Resolution Evaluation

Our current query mutator runs off-line, producing a stored query stream that can be replayed. We also plan to explore near-real-time mutators that can modify live traffic.

Comparing Answers in DIINER Tomorrow: An important experimental question is to determine if the experimental approach provides correct answers? We are planning to develop a *query validator* that can consume the output of experiments and compare it to answers from a production system.

A challenge with query validation of DNS traffic is that the protocol allows latitude in what responses are correct. Record order can vary, DNS glue can sometimes be dropped, and implementation differences result in small variation inside the "envelope" of the DNS specification. An ideal query validator should have options to check exact replies and also to allow a broader set of acceptable replies. Work on a query validator is still early, but we plan to provide one in the next two years.

Current status: As of July 2020, our testbed has been used for internal evaluation and we are looking for interested partners for alpha testing by external users. Please talk to us if you are interested.

This work is partially supported by the project "DNS, Identity, and Internet Naming for Experimentation and Research (DIINER)" (NSF CNS-1925737).

REFERENCES

[1] Wouter B. de Vries, Quirin Scheitle, Moritz Müller, Willem Toorop, Ralph Dolmans, and Roland van Rijswijk-Deij. A first look at QNAME minimization in the Domain Name System. In *Proceedings of the Passive and Active Measurement Conference*, pages 147–160, Puerto Varas, Chile, March 2019. Springer.

[2] DNS-OARC. Dnsperf. https://github.com/DNS-OARC/dnsperf, 2019.

[3] github/jedisct1. DNSBlast.

[4] Liang Zhu and John Heidemann. LDplayer: DNS experimentation at scale. In *Proceedings of the ACM Internet Measurement Conference*, Boston, Massachusetts, USA, October 2018. ACM.

[5] Liang Zhu, Zi Hu, John Heidemann, Duane Wessels, Allison Mankin, and Nikita Somaiya. Connection-oriented DNS to improve privacy and security. In *Proceedings of the 36th IEEE Symposium on Security and Privacy*, pages 171–186, San Jose, Californa, USA, May 2015. IEEE.

DINR'20, 2020, Marina del Rey, CA, USA
2020.