

Realizando o Potencial da Plataforma RIPE Atlas

Lucas Barsand, Artur Vaz, João Paulo Bastos, Osvaldo Fonseca, Ítalo Cunha

Universidade Federal de Minas Gerais

{lucas.barsand, arturvaz, joaopaulosr, osvaldo.morais, cunha}@dcc.ufmg.br

Abstract. *Executing network measurements in the Internet is essential to characterization efforts, traffic engineering, and failure troubleshooting. The RIPE Atlas platform is one of the largest publicly accessible active measurement platforms. RIPE Atlas, however, imposes two constraints on user-defined measurements. First, measurement budgets protect networks hosting vantage points and destinations from overload. Second, the RIPE Atlas API limits which measurements can be requested and when they will be executed. We introduce AtlasDNS, a framework that interposes between users and the RIPE Atlas API to provide flexible and centralized control of RIPE atlas measurements, allowing more precise control of what measurement to perform and reducing measurement cost to allow more measurements under a fixed budget.*

Resumo. *A execução de medições de rede é parte integral de esforços de caracterização, engenharia de tráfego e resolução de falhas na Internet. A plataforma RIPE Atlas é uma das maiores plataformas abertas de medições ativas. O RIPE Atlas, porém, impõe duas restrições sobre medições definidas por usuários. Primeiro, orçamentos de medição protegem redes hospedando pontos de medição e destinos contra sobrecargas. Segundo, a API do RIPE Atlas limita quais medições podem ser requisitadas e quando serão executadas. Nós apresentamos o AtlasDNS, um arcabouço que se interpõe entre usuários e a API do RIPE Atlas para prover controle flexível e centralizado de medições, permitindo controle mais preciso de quais medições executar e reduzindo o custo de medições para permitir mais medições dentro do mesmo orçamento.*

1. Introdução

Medir e monitorar a topologia da Internet é uma tarefa integral para modelagem e caracterização de suas propriedades [Spring et al. 2004, Cunha et al. 2014, Anwar et al. 2015, Lone et al. 2017]. Além disso, o monitoramento da topologia da Internet é útil para operadores que precisam entender e realizar engenharia de tráfego ou solucionar falhas [Katz-Bassett et al. 2008, Katz-Bassett et al. 2012, Giotsas et al. 2017]. A ferramenta mais comumente utilizada para realizar medições ativas de rotas entre *dois* dispositivos na Internet é o traceroute. Estas tarefas, entretanto, requerem medições com grande cobertura da rede para permitir identificação de rotas alternativas e isolamento de potenciais enlaces com falhas. Em geral, estas medições requerem execução sistemática do traceroute entre conjuntos representativos de origens e destinos.

Plataformas colaborativas de medição são a abordagem mais comum para obter acesso a conjuntos de pontos de medição diversos, permitindo coleta de medições topologicamente distribuídas pela Internet [Cunha et al. 2016]. O RIPE Atlas¹ é uma das

¹<https://atlas.ripe.net>

maiores destas plataformas e oferece aos seus usuários mais de 10.000 pontos de medição (*sondas*) instalados em mais de 3.000 sistemas autônomos (ASes), e permite que seus usuários realizem medições ativas como ping, traceroute e resolução de nomes (DNS). Um diferencial do RIPE Atlas é o paradigma de controle de acesso, que incentiva a colaboração para o crescimento e manutenção da plataforma por meio de um *sistema de créditos*. Usuários que hospedam sondas em sua rede recebem créditos, que podem ser gastos para realizar medições a partir de qualquer sonda disponível na plataforma. Este sistema motiva usuários a expandir a plataforma na mesma proporção que sua demanda.

Apesar da significativa contribuição do RIPE Atlas para realização de medições,² a plataforma impõe restrições como quotas diárias de utilização e limite de medições concorrentes à condução de experimentos para evitar sobrecarregar redes hospedando sondas e destinos com tráfego de medição excessivo. Estas restrições dificultam a condução de experimentos que precisam de controle mais preciso das medições.

Neste trabalho apresentamos o ATLASDNS, um arcabouço interposto entre usuários e a API do RIPE Atlas para permitir controle mais preciso de medições e otimizar a utilização de recursos (seção 2). O ATLASDNS intermedia a interação com a API do RIPE Atlas para configuração de medições e possibilita o controle de origens e destinos de forma mais flexível. Além disso, o ATLASDNS permite a configuração de medições contínuas, que custam a metade da quantidade de créditos que medições individuais, permitindo a execução do dobro do número de medições.

Uma das contribuições do ATLASDNS é utilizar redirecionamento via DNS, similar ao utilizado em redes de distribuição de conteúdo, para medições do RIPE Atlas. Um nome de domínio gerenciado pelo ATLASDNS é primeiro configurado como o destino de medições do RIPE Atlas. Quando a medição é executada em uma sonda s , esta dispara uma requisição DNS. O ATLASDNS itera sobre a lista de endereços vinculada a este nome para retornar um endereço IP específico. O ATLASDNS permite que usuários especifiquem quais endereços IP devem ser retornados para cada nome alvo. O ATLASDNS permite ainda que usuários controlem quais sondas devem realizar medições para cada nome gerenciado pelo arcabouço. Estas funcionalidades, combinadas, permitem que usuários do ATLASDNS controlem de forma programática as origens e destinos das medições realizadas pelo RIPE Atlas.

Avaliamos a eficácia do ATLASDNS realizando um experimento de descoberta de interfaces na Internet. Inspirados em aplicações práticas que utilizam informações obtidas do protocolo BGP (AS-paths) para definir origens e destino de medições [Spring et al. 2004, Katz-Bassett et al. 2012], configuramos um experimento que visa minimizar a sobreposição entre medições de traceroute para maximizar o número de interfaces descobertas. Comparamos resultados experimentais obtidos utilizando as funcionalidades nativas do RIPE Atlas e utilizando o ATLASDNS, demonstrando que a reconfiguração de destinos e redução do custo das medições permite a descoberta de quase duas vezes mais interfaces com o mesmo orçamento de medição.

2. O RIPE Atlas

Atualmente, o RIPE Atlas é uma das plataformas com maior número de pontos de medição (*sondas*) disponível publicamente. O controle de acesso a seus serviços é reali-

²A página do RIPE Atlas na Wikipedia lista mais de 25 publicações que utilizaram a plataforma.

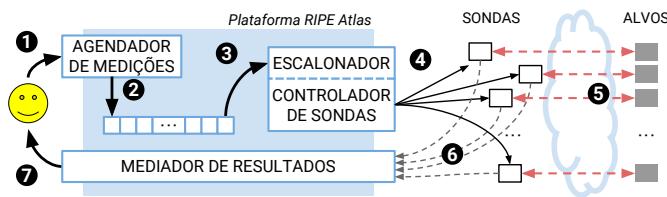


Figura 1. Fluxo de etapas para requisição, escalonamento, execução, disponibilização e acesso a medições no RIPE Atlas.

zado por meio de um sistema de créditos baseado em recompensa, que concede créditos diariamente para usuários que hospedam sondas em suas redes. Estes créditos, por sua vez, são utilizados pelos usuários na realização de medições a partir de outras sondas hospedadas por outros usuários. Medições podem ser avulsas (*one-off*) ou periódicas (com intervalo entre medições especificado pelo usuário). Medições avulsas custam o dobro de crédito que medições periódicas.

A figura 1 ilustra a requisição e execução de uma medição no RIPE Atlas [RIPE NCC Staff 2015]. Para configurar uma medição de traceroute, o usuário deve submeter (passo 1) parâmetros de uma medição (como o destino da medição e sondas a serem utilizadas como origem) à interface de agendamento, que a insere em uma fila junto às demais requisições (2). A requisição é processada (3) e enviada para as sondas (4), que executam o traceroute (5). Em seguida, os resultados da medição são enviados (6) para a interface de disponibilização de resultados, que finalmente são coletados pelo usuário (7). Como cada um dos passos 2 a 6 pode levar uma quantidade de tempo fora do controle do usuário, esta arquitetura dificulta o controle de quando as medições são realizadas.

O uso de funcionalidades nativas da plataforma restringe medições de rotas na Internet a medições para até 100 alvos distintos, o que é uma fração insignificante da quantidade de possíveis alvos na Internet. Além disso, o limite diário de uso de créditos também restringe a execução diária de traceroutes a menos de 17.000: uma fração insignificante do número de rotas na Internet. Estes problemas tornam o uso eficiente de créditos na plataforma essencial.

3. ATLASDNS

O ATLASDNS é um arcabouço para controlar origens e destinos de medições no RIPE Atlas, que gerencia um conjunto de medições previamente configuradas no RIPE Atlas e um conjunto de nomes de domínio \mathcal{N} . Cada nome de domínio $n \in \mathcal{N}$ (e.g., `targets0.atlas.winet.dcc.ufmg.br`) é configurado como destino de uma medição controlada pelo ATLASDNS. O controle sobre cada nome de domínio n e a medição correspondente é outorgado pelo administrador do sistema a um usuário.

Endpoint	Operações	Ações
<code>/targets</code>	GET	Consultar o estado dos nomes de domínio gerenciados.
<code>/ips/n/</code>	GET, POST, UPDATE	Gerenciar M_n (atualizações informam um dicionário).
<code>/origins/n/</code>	GET, POST	Gerenciar \mathcal{S}_n (atualizações informam uma lista).
<code>/admin/</code>	GET, POST	Gerenciar U_n (atualizações informam um usuário).

Tabela 1. Endpoints da interface REST do ATLASDNS

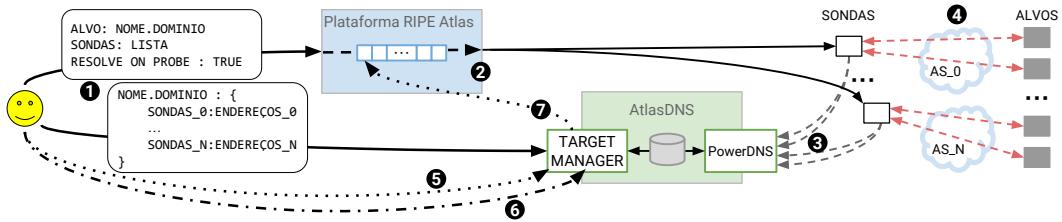


Figura 2. Fluxo de controle de medições desde a configuração do usuário até o seu recebimento nas sondas

Usuários podem modificar as origens (sondas) e destinos das medições para cada um dos nomes de domínio sob seu controle utilizando a interface REST provida pelo módulo ENDPOINT-MANAGER. Para cada nome de domínio n gerenciado pelo ATLASDNS, uma lista de sondas \mathcal{S}_n que realizam medições para aquele nome de domínio. Para gerenciar os destinos associados a um nome de domínio n , o ENDPOINT-MANAGER mantém um dicionário $M_n \rightarrow \mathcal{I}_n$ mapeando um nome $n \in \mathcal{N}$ para uma sequência de endereços IP \mathcal{I}_n para os quais sondas que resolvem n devem realizar medições. Além disso, o ENDPOINT-MANAGER mantém o último endereço IP em \mathcal{I}_s retornado a cada sonda s .

Quando uma requisição DNS para um nome n é recebida, o ENDPOINT-MANAGER consulta o próximo elemento i de \mathcal{I}_n , e o retorna para a sonda que realizou a requisição (se o último endereço retornado for o último da sequência, o ENDPOINT-MANAGER retorna ao início da sequência).

O ATLASDNS permite que administradores modifiquem o conjunto de usuários U_n autorizados a gerenciar \mathcal{S}_n e M_n (para os nomes de domínio n que controlam) dinamicamente através de uma interface REST, mostrada na tabela 1. As modificações realizadas em U_n , \mathcal{S}_n , e M_n surtem efeito imediatamente, permitindo que usuários controlem de forma instantânea quais medições devem ser realizadas. Para evitar que usuários aumentem o custo das medições configuradas inicialmente, o ATLASDNS não permite modificação do número de sondas (tamanho) em \mathcal{S}_n . Em termos práticos, usuários podem apenas trocar uma sonda por outra. Porém, modificações do conjunto \mathcal{S}_n são efetivadas realizando *participation requests* na interface do RIPE Atlas, o que induz atraso de escalonamento de novas medições para sondas adicionadas a \mathcal{S}_n . Nenhuma dessas operações consome créditos do RIPE Atlas.

A configuração e utilização do funcionamento do ATLASDNS, bem como sua iteração com o RIPE Atlas, seguem as etapas ilustradas na figura 2. Inicialmente, a configuração das medições é realizada através das interfaces do RIPE Atlas (1). Para cada medição é necessário informar o nome de domínio n alvo, um conjunto de sondas inicial para a medição (i.e., o que define $|\mathcal{S}_n|$), e habilitar a resolução de nome pela própria sonda (caso contrário o RIPE Atlas resolve uma única vez e substitui o nome por um endereço IP). Os passos 2–4 são executados indefinidamente (até o final das medições ou desligamento do ATLASDNS): (2) o RIPE Atlas escalona as medições, (3) as sondas resolvem o nome e (4) realizam a medição. A disponibilização dos resultados segue o fluxo descrito na seção 2. A qualquer momento, um usuário pode modificar a configuração das origens \mathcal{S}_n (5) ou destinos M_n (6) associados a nomes de domínio que controla. Uma modificação de \mathcal{S}_n dispara uma sequência de *participation requests* para a API da RIPE (7).

Considerações Éticas: Reforçamos que o ATLASDNS não fere os termos de uso do RIPE Atlas. Pelo contrário: a ferramenta apresentada contribui para a prevenção de sobrecarga na plataforma, já que reduz a quantidade de requisições à API do RIPE Atlas por permitir reconfiguração dos destinos via DNS. Ressaltamos ainda que todas as restrições impostas a usuários do RIPE Atlas continuam válidas para medições gerenciadas pelo ATLASDNS.

3.1. Detalhes de Implementação

O ATLASDNS é modularizado em contêineres do Docker³, o que facilita a sua implantação, extensão e atualização de partes do sistema. Um dos contêineres executa o ENDPOINT-MANAGER, desenvolvido em cima de uma versão do Flask estendida para permitir autenticação HTTPS com certificados de cliente⁴ (a versão padrão só inclui o caso comum de certificados de servidor). O Flask é conectado ao banco de dados não-relacional MongoDB que armazena e persiste a configuração atual.

Outro contêiner executa o servidor PowerDNS com um *backend* desenvolvido em Python 3. O *backend* é chamado sempre que o servidor recebe uma requisição para um dos nomes de domínio gerenciados pelo ATLASDNS, e retorna um endereço IP de acordo o estado atual (i.e., $M_n[s]$ e o último endereço IP enviado a s). Este contêiner tem acesso ao banco de dados gerenciado pelo MongoDB.

Por fim, um contêiner executa o Redis, que serve como gerenciador de mensagens (*message broker*) do Celery, que é usado no ENDPOINT-MANAGER para agendar as execuções dos *participation requests* na API do RIPE Atlas sempre que um usuário atualiza a lista de sondas em um conjunto \mathcal{S}_n .

4. Avaliação

Para validar a eficácia do ATLASDNS, comparamos dois experimentos realizados no RIPE Atlas onde o objetivo é maximizar a quantidade de interfaces e sistemas autônomos (ASes) descobertos. Referenciamos os experimentos por E_{nativo} e E_{ADNS} , onde o primeiro utiliza a funcionalidade nativa do RIPE Atlas para requisitar medições avulsas e o segundo utiliza o ATLASDNS para fazer reconfiguração dinâmica.

4.1. Escolha de Origens e Destinos dos Traceroutes

No experimento E_{nativo} , utilizamos uma heurística para escolher pares de origem e destino que maximizam a quantidade de ASes cobertos pelas medições. Primeiro obtemos os ASes hospedando cada sonda do RIPE Atlas utilizando a API do sistema. Depois processamos tabelas de roteamento BGP do RIPE RIS e do RouteViews para identificar os ASes nas rotas utilizadas por estes ASes e outros destinos na Internet. Por último, aplicamos um algoritmo guloso para escolher sondas e destinos de forma a maximizar o número de ASes cobertos pelos pares de origem e destino selecionados. A cada iteração, o algoritmo escolhe o par origem-destino que descobre o maior número de novos ASes. Inicialmente, o algoritmo escolhe o par origem-destino com a maior rota. Os ASes na rota escolhida são considerados cobertos e ignorados nas iterações seguintes. Executamos 16.000 iterações para escolher realizar 16.000 traceroutes avulsos no RIPE Atlas, praticamente o máximo

³ Embora nossa implementação use o Docker, poderia ser facilmente adaptada para outras tecnologias.

⁴ Disponível em <https://github.com/barsand/werkzeug/tree/cert-forwarding>

Experimento	Créditos	Traceroutes	Cobertura	
			Interfaces	ASes
E_{ADNS}	960.000	32.000	101.178	5.841
E_{nativo}		16.000	47.194	3.446

Tabela 2. Resultados dos experimentos utilizando funcionalidades nativas do RIPE Atlas e as do ATLASDNS

permitido pelo limite diário de uso de créditos (mantivemos uma margem de segurança para não atingir o máximo e comprometer o experimento).

Para o experimento E_{ADNS} , primeiro identificamos todos endereços IPv4 responsivos a *ping*, obtendo uma lista \mathcal{H} de 1.254.249 endereços. Em seguida, utilizamos a interface da RIPE Atlas para obter a relação de 8.743 sondas públicas ativas. Criamos então 7 nomes de domínio, e configuramos o mesmo número de medições periódicas, de forma que cada medição possua no máximo 1.000 sondas distintas (o máximo permitido pela plataforma). Ao final da configuração junto ao RIPE Atlas, foram realizadas medições a partir de 6.530 (um subconjunto das sondas ativas disponíveis para realizar traceroutes). Para cada nome de domínio, configuramos os alvos no ATLASDNS com os endereços de \mathcal{H} , de forma que cada lista de endereços fosse compartilhada por todas as sondas de um experimento vinculada ao seu respectivo domínio. Além disso, aleatorizamos a lista de endereços para evitar que vários domínios sejam sondados em um curto intervalo. Definimos o período entre medições consecutivas de cada sonda de forma a utilizarmos a mesma quantidade de créditos que no experimento E_{nativo} ; como medições periódicas custam a metade de medições avulsas, realizamos um total de 32000 traceroutes em E_{ADNS} .

4.2. Resultados

A tabela 2 mostra um sumário dos nossos resultados. A figura 3 mostra o número de interfaces e o número de ASes descobertos à medida que os experimentos são executados. No geral, o experimento E_{ADNS} observa mais que o dobro do número de interfaces e 61% mais ASes. Notamos que este resultado foi obtido utilizando uma heurística para escolher pares origem-destino no experimento E_{nativo} e sem escolha de pares para o experimento E_{ADNS} . Estes resultados demonstram a eficácia do ATLASDNS comparado à utilização da funcionalidade nativa do RIPE Atlas.

5. Aplicações Práticas

O PEERING Testbed [Schlinker et al. 2014] é uma plataforma que permite o estudo do protocolo de roteamento interdomínio (BGP) em um cenário real, oferecendo aos pesquisadores acesso à uma rede (AS47065) conectada a diversas outras redes na Internet, incluindo múltiplos provedores e alguns pontos de troca de tráfego (PTTs). Assim, o PEERING possibilita a condução de experimentos que demandam medição ativa do roteamento entre domínios na Internet.

Diversos trabalhos na literatura abordam temas que envolvem a coleta de informações de rotas na Internet de forma dinâmica. Por exemplo, [Cunha et al. 2014] monitora mudanças de rota entre diversos pontos da Internet, e aloca medições dinamicamente de acordo com a estabilidade de cada rota; [Quan et al. 2013] avalia a confiabilidade da rede, direcionando medições para regiões não exploradas ou com confiabilidade

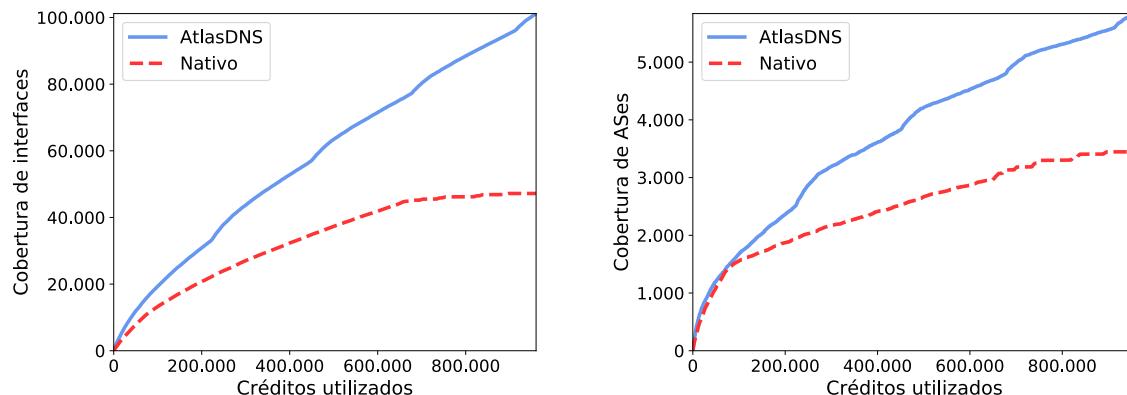


Figura 3. Comparação da cobertura de Interfaces (esquerda) e ASes (direita) entre os experimentos com AtlasDNS e RIPE Atlas convencional

reduzida. Na maioria dos experimentos utilizando o PEERING, ao realizar um anúncio BGP que irá propagar pela Internet, é necessário monitorar as rotas escolhidas pelos ASes para alcançar o PEERING. Duas formas utilizadas para encontrar essas rotas são: verificar os AS-paths nos coletores BGP ou realizar medições de traceroute para o prefixo anunciado, partindo de vários pontos de medição na Internet. A primeira solução fornece uma cobertura de ASes inferior a 400 ASes, sendo necessário aumentar essa cobertura através de pontos de medição espalhados na Internet.

O ATLASDNS oferece funcionalidades que facilitam a condução dos experimentos apresentados nos trabalhos citados e em vários outros. Sua utilização por pesquisadores e operadores de rede poderá potencializar o uso da plataforma RIPE Atlas e contribuir diretamente com a qualidade de trabalhos futuros na área de medição e monitoramento da Internet. Em especial, ao configurar anúncios de prefixos do PEERING, os usuários que executam experimentos a fim de provocar mudanças nas políticas de roteamento interdomínio podem configurar medições na RIPE com ajuda com ATLASDNS para monitorar as rotas entre sondas para os prefixos anunciados. Esta configuração pode ser gerenciada sob demanda, de acordo com os reflexos observados, para realizar a verificação da efetividade dos anúncios.

6. Planejamento de Demonstração no Salão de Ferramentas

O Código-fonte, documentação e tutorial de instalação do ATLASDNS estão disponíveis em <http://dcc.ufmg.br/~lucas.barsand/atlasdns>. Para demonstrá-lo ao público no salão de ferramentas, será realizada a implantação do ATLASDNS em um servidor da UFMG gerenciando medições reais no RIPE Atlas. Um segundo servidor poderá ser utilizado para apresentar o processo de implantação do sistema com o Docker. Utilizaremos um computador que suporte utilização de comandos Unix e conexão à Internet para configurar experimentos de medição e evidenciar a contribuição da ferramenta descrita neste trabalho no cenário da avaliação descrito na seção 4.2, apoiado por um poster com conceitos básicos ilustração da arquitetura do ATLASDNS.

Adicionalmente, ofereceremos ao público a oportunidade de interagir com as interfaces do ENDPOINT-MANAGER em um interpretador Python pré-configurado, e gerenciar sob demanda configurações de alvos para observar seu funcionamento ao vivo.

7. Conclusão

Diante da importância da realização de traceroutes para pesquisa e operação de redes, apresentamos o ATLASDNS, um arcabouço para gerenciamento de medições no RIPE Atlas, uma das maiores plataformas abertas de monitoramento da atualidade. O ATLASDNS expande as funcionalidades do RIPE Atlas e é programaticamente acessível.

Avaliamos a eficácia do ATLASDNS realizando dois experimentos com o objetivo de identificar interfaces e sistemas autônomos na Internet, um experimento usando a funcionalidade nativa do RIPE Atlas e outro utilizando o ATLASDNS. Mostramos que o ATLASDNS identificou mais que o dobro de interfaces e 61% a mais de sistemas autônomos para o mesmo número de créditos utilizados no RIPE Atlas.

Agradecimentos: Este trabalho foi financiado por: RIPE NCC, RNP, FAPEMIG, CAPES e CNPq.

Referências

- [Anwar et al. 2015] Anwar, R., Niaz, H., Choffnes, D., Cunha, I., Gill, P., and Katz-Bassett, E. (2015). Investigating interdomain routing policies in the wild. In *ACM IMC*.
- [Cunha et al. 2016] Cunha, I., Marchetta, P., Calder, M., Chiu, Y.-C., Schlinker, B., Machado, B. V. A., Pescapè, A., Giotsas, V., Madhyastha, H. V. A., and Katz-Bassett, E. (2016). Sibyl: A practical internet route oracle. In *Usenix NSDI*.
- [Cunha et al. 2014] Cunha, I., Teixeira, R., Veitch, D., and Diot, C. (2014). Dtrack: A system to predict and track internet path changes. *IEEE/ACM Transactions on Networking*, 22(4):1025–1038.
- [Giotsas et al. 2017] Giotsas, V., Dietzel, C., Smaragdakis, G., Feldmann, A., Berger, A., and Aben, E. (2017). Detecting peering infrastructure outages in the wild. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*.
- [Katz-Bassett et al. 2008] Katz-Bassett, E., Madhyastha, H., John, J. P., Krishnamurthy, A., Wetherall, D., and Anderson, T. (2008). Studying Black Holes in the Internet with Hubble. In *Proc. USENIX NSDI*.
- [Katz-Bassett et al. 2012] Katz-Bassett, E., Scott, C., Choffnes, D. R., Cunha, I., Valancius, V., Feamster, N., Madhyastha, H. V., Anderson, T., and Krishnamurthy, A. (2012). Lifeguard: Practical repair of persistent route failures. In *ACM SIGCOMM*.
- [Lone et al. 2017] Lone, Q., Luckie, M., Korczyński, M., and van Eeten, M. (2017). Using Loops Observed in Traceroute to Infer the Ability to Spoof. In *Proc. PAM*.
- [Quan et al. 2013] Quan, L., Heidemann, J., and Pradkin, Y. (2013). Trinocular: Understanding internet reliability through adaptive probing. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM*.
- [RIPE NCC Staff 2015] RIPE NCC Staff (2015). Ripe atlas: A global internet measurement network. *The Internet Protocol Journal*, 18(3).
- [Schlinker et al. 2014] Schlinker, B., Zarifis, K., Cunha, I., Feamster, N., and Katz-Bassett, E. (2014). PEERING: An AS for Us. In *Proc. ACM HotNets*.
- [Spring et al. 2004] Spring, N., Mahajan, R., and Wetherall, D. (2004). Measuring isp topologies with rocketfuel. In *IEEE/ACM Transactions on Networking (TON)*.