Experiências com um Honeypot DNS: Caracterização e Evolução do Tráfego Malicioso

Tiago Heinrich[1], **Felipe de Souza Longo**[1], **Rafael R. Obelheiro**[1]

[1] Programa de Pós-Graduação em Computação Aplicada (PPGCA)
Universidade do Estado de Santa Catarina (UDESC) – Joinville, SC – Brasil

tiagoheinrich1995@gmail.com, felipeslongo@gmail.com, rafael.obelheiro@udesc.br

Abstract. *The Domain Name System (DNS) plays a central role in the operation of the Internet, being responsible for translating user-friendly names into machine-friendly IP addresses. However, the DNS has certain structural security vulnerabilities which allow it to be attacked or used as a tool for attacking third parties. Currently, a major concern are Distributed Reflection Denial of Service (DRDoS) attacks, which leverage misconfigured DNS servers to flood victims with traffic. This paper introduces DNSpot, a DNS-specific honeypot that allows attackers to interact with an open recursive DNS server in a controlled manner. We also analyze the DNS traffic observed by this honeypot over two periods, in 2015 (49 days) and 2016–2017 (250 days), with a focus on DRDoS attacks, and highlight some noteworthy aspects of attacker behavior.*

Resumo. *O Domain Name System (DNS) desempenha um papel central no funcionamento da Internet, sendo responsável por traduzir nomes mnemônicos em endereços IP. No entanto, o DNS possui certas vulnerabilidades estruturais de segurança, que permitem que ele seja atacado ou usado como instrumento de ataque a terceiros. Atualmente, uma grande preocupação são os ataques distribuídos de negação de serviço por reflexão (Distributed Reflection Denial of Service, DRDoS), que se aproveitam de servidores DNS mal configurados para saturar vítimas com tráfego. Este artigo introduz o DNSpot, um honeypot específico para DNS que permite que atacantes interajam com um servidor DNS recursivo aberto de forma controlada. O artigo também apresenta uma análise do tráfego DNS observado pelo honeypot durante dois períodos, em 2015 (49 dias) e 2016–2017 (250 dias), com foco em ataques DRDoS, destacando também aspectos notáveis do comportamento de atacantes.*

1. Introdução

O DNS (Domain Name System) [Mockapetris 1987] é um sistema distribuído de resolução de nomes que desempenha um papel vital na Internet. Ele possibilita que usuários refiram-se a nós da rede empregando nomes mnemônicos (como `www.google.com`) em vez de endereços IP (como 216.58.219.110), sendo responsável, entre diversas outras funcionalidades, pela tradução de nomes em endereços.

Em vista de sua ampla utilização, o DNS também é alvo e vetor de ataques. As principais ameaças envolvendo o DNS são resumidas por [Conrad 2012], que as divide em duas classes, ameaças ao próprio DNS e ameaças oportunizadas pelo DNS. A classe de ameaças ao DNS inclui:

- Negação de serviço: bloqueio do acesso de usuários ao DNS, com isso prejudicando ou mesmo impedindo (na prática) o seu acesso à Internet;
- Corrupção de dados: modificação não autorizada dos dados publicados no DNS, o que pode, por exemplo, levar usuários a acessar sites ilegítimos (como sites falsos de bancos ou de comércio eletrônico);
- Exposição de informação: revelação de informações sobre o comportamento dos usuários, como histórico de sites web acessados.

O DNS também pode ser usado como vetor de ataques. A classe de ameaças oportunizadas pelo DNS abrange:

- Ataques de amplificação: servidores DNS mal configurados podem ser usados para realizar ataques distribuídos de negação de serviço por reflexão (*Distributed Reflection Denial of Service*, DRDoS) contra terceiros [CERT.br 2016];
- *Fast flux* DNS: servidores usados para propósitos nefastos, como propagação de software malicioso, podem receber nomes que são mudados com alta frequência, com o propósito de dificultar a localização dos servidores e a identificação dos seus responsáveis;
- Exfiltração de dados: como o tráfego DNS geralmente passa incólume por *firewalls*, ele é usado com frequência para transmissão de dados sensíveis (capturados no curso de uma invasão) de forma sub-reptícia.

Dentre as ameaças acima, os ataques DRDoS despertam particular interesse. Em seu relatório anual sobre segurança na Internet [Arbor 2017], a Arbor Networks reportou que, de acordo com tráfego observado por diversos sensores espalhados pela Internet, em 2016 o DNS foi usado em 47% dos ataques DRDoS, gerando um volume médio de 3.083 Mbps por ataque, ficando à frente de outros protocolos em ambos os quesitos. Ao longo de 2016, os ataques DDoS baseados no DNS aumentaram de 10.500 para 18.500 por semana, em média. Um outro relatório recente da Akamai [Akamai 2017] reporta que, no primeiro trimestre de 2017, 57% dos ataques DDoS observados eram baseados em reflexão, e que o DNS foi um dos protocolos mais usados.

A Figura 1 ilustra o funcionamento de um ataque DRDoS usando o DNS. Um atacante envia a um servidor DNS recursivo[1] um número elevado de consultas com o endereço IP de origem forjado como sendo o da vítima do ataque (passo 1); o servidor recursivo interage com servidores autoritativos para obter a resposta para a consulta (passos 2 e 3); a resposta é enviada para a vítima (passo 4). O ataque é facilitado pelo fato de consultas DNS pequenas poderem gerar respostas grandes; como será discutido na Seção 4, esse fator de amplificação pode chegar a 100 (um tráfego de consulta de 1 Mbps gera um tráfego de resposta de 100 Mbps). Se forem usados diversos clientes e servidores simultaneamente (por exemplo, uma *botnet* de clientes usando servidores DNS distintos), é relativamente fácil gerar uma intensidade de tráfego suficiente para sobrecarregar a vítima, indisponibilizando sua conectividade.

Uma forma de observar o comportamente de atacantes é usando *honeypots*, que são recursos computacionais com o objetivo de serem sondados, atacados ou comprometidos [Steding-Jessen et al. 2008]. Geralmente um *honeypot* é um *host* Internet que possui

[1]Servidores DNS autoritativos são responsáveis pelos dados de domínios específicos e precisam atender a consultas de quaisquer clientes a respeito desses domínios. Servidores DNS recursivos fazem resolução de nomes (recorrendo a servidores autoritativos) em benefício de clientes (*hosts*), e devem ter seu acesso limitado a um conjunto de clientes autorizados [CERT.br 2016].

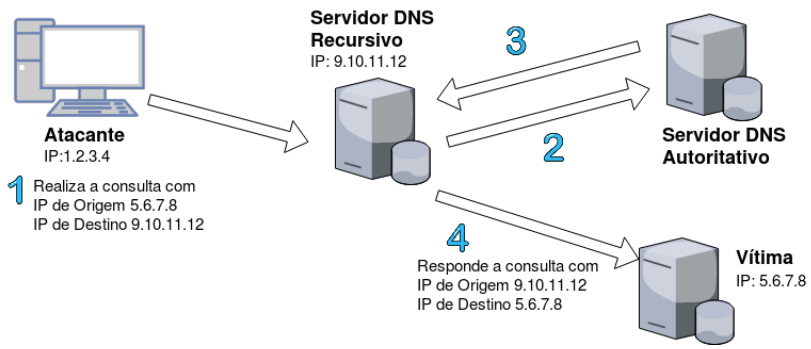


Figura 1. Ataque DRDoS usando o DNS

um endereço IP público (isto é, um endereço IP roteável globalmente), e que hospeda serviços de rede que não são anunciados publicamente. Qualquer interação realizada com um *honeypot* já pode ser considerada suspeita, já que é necessária a realização de varreduras para a descoberta do endereço IP do *honeypot* e de seus serviços.

Este artigo aplica o conceito de *honeypot* à análise de tráfego DNS malicioso. Sua primeira contribuição é a introdução do DNSpot, um *honeypot* DNS que tem por objetivo permitir a interação controlada de atacantes com um servidor DNS recursivo aberto. O diferencial do DNSpot é permitir a observação de tráfego suspeito evitando que esse tráfego cause danos ao próprio servidor ou a terceiros. A segunda contribuição do artigo é uma análise do tráfego coletado pelo DNSpot durante dois períodos, em 2015 (49 dias) e em 2016–2017 (250 dias), mostrando várias características do tráfego de ataque, com foco em ataques DRDoS, e a sua evolução de um período para o outro.

O restante deste artigo está organizado como segue. A Seção 2 discute trabalhos relacionados. A Seção 3 descreve o DNSpot. A Seção 4 apresenta a análise do tráfego coletado pelo *honeypot*, e a Seção 5 conclui o artigo.

2. Trabalhos Relacionados

Estudos envolvendo análise de tráfego DNS vêm sendo conduzidos há vários anos. A primeira linha de investigação foi a caracterização de tráfego DNS sob o prisma de servidores DNS raiz, tanto globais [Danzig et al. 1992, Brownlee et al. 2001, Castro et al. 2010] quanto do Brasil [Barbosa and Pereira 2009]. Outros estudos caracterizam o tráfego DNS sob a ótica de clientes e/ou servidores DNS recursivos [Jung et al. 2002, Gao et al. 2013]. De forma geral, esses trabalhos têm foco em descrever a natureza e a intensidade do tráfego DNS, com pouca ênfase em aspectos de segurança.

Uma outra linha de investigação envolve a detecção de tráfego DNS anômalo e/ou malicioso. [Zdrnja et al. 2007] tem o objetivo de detectar tráfego DNS anômalo, incluindo domínios similares a domínios conhecidos (*typo squatter domains*), que podem ser usados para enganar usuários, nomes *fast-flux*, e nomes associados a campanhas de envio de SPAM. [Perdisci et al. 2009] buscam detectar nomes DNS *fast-flux*. [Zhao et al. 2015] têm o foco na detecção de nomes usados para canais de comunicação e controle (C&C) de *malware*, que são usados para o controle remoto do *malware*, especialmente de *botnets*. O DNSpot introduzido neste trabalho poderia ser adaptado sem

grande dificuldade para coletar dados que seriam usados pelos algoritmos propostos nesses trabalhos.

Uma terceira linha de investigação usa varreduras (*scans*) para identificar servidores DNS recursivos abertos na Internet e realiza uma caracterização desses servidores. Em [Takano et al. 2013] são identificados 30 milhões de servidores, dos quais 25 milhões eram recursivos abertos; os servidores são classificados de acordo com sua versão de *software* e domínio Internet. [Kührer et al. 2015] analisam a evolução do número de recursivos abertos encontrados por varreduras (uma queda de 26,8 para 17,8 milhões entre 2014 e 2015) e caracterizam versão de *software*, tipo de *hardware*, e sistema operacional desses servidores. O estudo revela uma incidência significativa de servidores que manipulam respostas DNS, e que essa manipulação pode tanto legítima (*e.g.*, portais captivos) quanto ilegítima (*e.g.*, censura, redirecionamento de anúncios, *phishing* e *malware*).

Alguns trabalhos usam medições para avaliar o fator de amplificação de ataques usando o DNS. [Anagnostopoulos et al. 2013] discute o uso de registros DNSSEC nos ataques DRDoS, e realiza medições com servidores DNS recursivos, onde é observado um fator de amplificação máximo de 44. [van Rijswijk-Deij et al. 2014] realiza um estudo sobre o fator de amplificação obtido com o uso de registros DNSSEC. Foram realizadas consultas em servidores autoritativos responsáveis por 70% dos domínios assinados à época, sendo constatado que os dois fatores de amplificação mais comuns eram 40 e 55, e que o fator máximo era 178,6. [MacFarland et al. 2017] investiga o potencial de ataques DRDoS usando apenas servidores autoritativos como refletores, mostrando que, entre 2013 e 2015, o fator de amplificação médio aumentou de 32,77 para 41,49, e destacando ainda o papel dos registros DNSSEC na amplificação. Esse estudo também mede a adoção de limitação da taxa de resposta (RRL, *response rate limiting*), um mecanismo recomendado para reduzir o impacto de ataques DRDoS [CERT.br 2016]. Enquanto esses trabalhos realizam consultas a servidores DNS recursivos ou autoritativos para mensurar o fator de amplificação (medição ativa), o presente estudo usa uma abordagem de medição passiva, caracterizando o tráfego de ataque associado às consultas DNS recebidas por um *honeypot*, que é efetivamente (e não apenas potencialmente) ofensivo.

O trabalho mais proximamente relacionado é [Fachkha et al. 2015], que apresenta uma análise de 1,44 TB de tráfego DNS enviado a uma *darknet* /13 (mais de 512.000 endereços IP) durante seis meses. Como os endereços IP de uma *darknet* não são utilizados, todo o tráfego destinado a ela é suspeito, podendo ser tráfego de ataque ou devido a erros de configuração. Ao longo do período de monitoração foram observados 134 ataques DDoS, com variados graus de intensidade. Por outro lado, não existem servidores DNS nessa rede, e portanto a análise restringe-se a tráfego enviado às cegas, sem que o atacante saiba se ele será respondido ou não. Como o DNSpot aparenta ser um servidor DNS, os atacantes podem interagir com ele a ponto de acreditar que ele produzirá as respostas necessárias para ataques DDoS. Isso torna o tráfego recebido pelo DNSpot mais representativo que tráfego enviado a endereços não usados em uma *darknet*, cuja eficácia é desconhecida. Por outro lado, o DNSpot oferece um único ponto de vista, o que dificulta a avaliação da real extensão dos ataques DDoS que são observados pelo *honeypot*.

Fora do contexto acadêmico, [Tanasi 2014] implementou um *honeypot* DNS que compartilha similaridades com o DNSpot, mas que registra apenas as consultas recebidas, não as respostas. [Van Impe 2015] implementou um servidor DNS que envia uma

resposta fixa para qualquer consulta, e analisou o tráfego recebido entre o final de dezembro de 2014 e meados de fevereiro de 2015, igualmente desconsiderando o tráfego de resposta. Alguns resultados são consistentes com os observados no DNSpot. Contudo, é inviável afirmar se as divergências devem-se a mudanças no tráfego malicioso entre os períodos monitorados por Van Impe e os descritos neste trabalho ou ao fato do servidor de Van Impe fabricar respostas fixas (e inconsistentes) para qualquer consulta, o que pode afastar atacantes que testam a funcionalidade de servidores recursivos abertos antes de usá-los em ataques.

3. DNSpot

O DNSpot é um *honeypot* projetado especificamente para monitorar e analisar o tráfego DNS. Seu objetivo é permitir a observação das interações de usuários potencialmente maliciosos com servidores DNS recursivos. A arquitetura do DNSpot pode ser observada na Figura 2. Ele possui um *proxy* que escuta na porta 53/UDP, que é a porta padrão do serviço DNS. Ao receber uma consulta, o *proxy* armazena a consulta em um banco de dados e logo em seguida a repassa a um servidor DNS recursivo real. Esse servidor real, que aceita apenas consultas originadas na própria máquina, obtém uma resposta do seu próprio *cache* ou interagindo com servidores autoritativos na Internet. Por último, o *proxy* recebe a resposta do servidor recursivo, armazena no banco de dados e encaminha para o cliente que enviou a consulta.

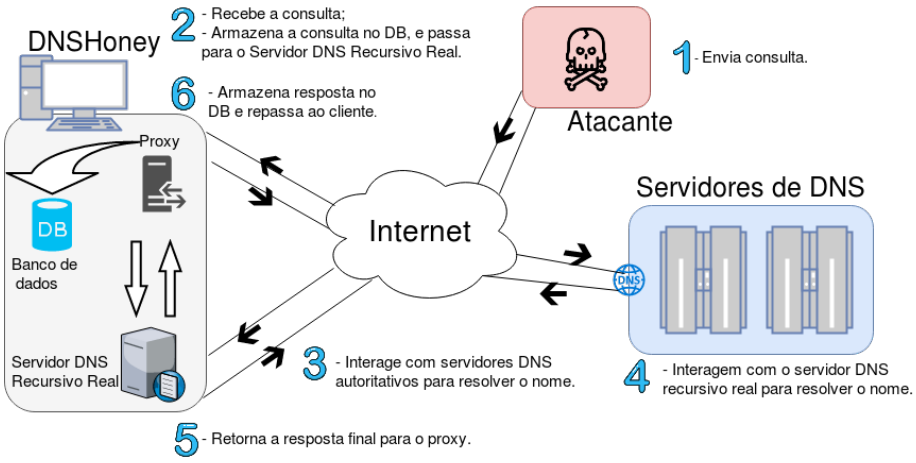


Figura 2. Arquitetura do DNSpot

Ocasionalmente, o DNSpot retorna uma mensagem falsa de erro (*ServFail*, que sinaliza uma falha inespecífica do servidor) para o cliente; a frequência dessas mensagens é configurável. O objetivo é simular, perante um atacante humano (não uma ferramenta automatizada), um servidor DNS inconfiável, tentando não levantar suspeitas caso o DNSpot seja desligado ou não consiga processar algumas consultas.

Para reduzir o impacto caso o DNSpot seja usado como refletor em um ataque DRDoS, é estabelecido um limite diário (também configurável) para o número de consultas atendidas para cada endereço IP de origem. Caso esse limite seja atingido, o *proxy* deixa de enviar respostas para consultas do mesmo endereço, até o dia seguinte. Essa contagem de consultas por endereço IP é zerada diariamente.

4. Análise de Tráfego

Esta seção apresenta uma análise do tráfego coletado com o DNSpot, abrangendo implantação (Seção 4.1), estatísticas gerais do tráfego (Seção 4.2), caracterização de ataques DoS (Seção 4.3), resultados notáveis (Seção 4.4), e discussão geral (Seção 4.5).

4.1. Implementação do DNSpot e Coleta de Dados

O DNSpot foi implementado em Python, usando Unbound[2] para o servidor DNS real e SQLite[3] para o banco de dados, em plataforma OpenBSD.[4] O *honeypot* foi instalado na rede da Universidade, coletando dados de forma quase ininterrupta durante os períodos mostrados na Tabela 1. Em ambas as coletas houve alguns períodos de indisponibilidade, causados por falta de energia, interrupção do acesso à Internet ou travamentos do sistema; a indisponibilidade total é estimada em 0,5% a 1% dos 300,1 dias de coleta.

O DNSpot foi configurado para responder a 20% das requisições com uma mensagem de falha do servidor, de modo a aparentar confiabilidade limitada. O número de consultas diárias por endereço IP foi fixado em 30; as consultas excedentes são processadas e armazenadas no banco de dados, mas as respostas não são enviadas. Esses parâmetros foram estabelecidos empiricamente, visando a limitar o tráfego gerado em ataques DRDoS e ainda permitir que atacantes humanos interajam com o *honeypot*; o segundo objetivo não seria alcançado caso o DNSpot não enviasse nenhuma resposta.

Existem diversos projetos que realizam varreduras para procurar servidores DNS recursivos abertos na Internet.[5] Essas varreduras podem ser identificadas pelos nomes ou sufixos consultados, e o DNSpot foi configurado para ignorar tais consultas.

Dataset	Início	Fim	Total (dias)
DS1	09/09/2015 07:57	28/10/2015 22:29	49,6
DS2	17/09/2016 08:00	25/05/2017 20:47	250,5

Tabela 1. *Datasets* coletados com o DNSpot

4.2. Estatísticas de Tráfego

A Tabela 2 resume o tráfego processado pelo DNSpot nos dois períodos de coleta. Foram mais de 4 milhões de requisições no *dataset* DS1 e mais de 32 milhões em DS2, totalizando 36,4 milhões de requisições processadas. Destaca-se o número pequeno de consultas respondidas (5,7% do total) e o grande número de consultas ignoradas (98,5% das não respondidas, ou 92,8% do total), o que reflete o funcionamento do mecanismo de limitação diária de consultas por endereço IP. A fração de consultas respondidas foi menor em DS2 do que em DS1, o que se explica tanto pelo aumento de erros (requisições malformadas) quanto pela maior intensidade e duração dos ataques DoS (discutidas na Seção 4.3). A taxa média de processamento de requisições passou de 81.353,02 transações por dia (0,96 tps) em DS1 para 129.955,61 em DS2 (1,50 tps), o que representa um incremento de 59,7% na intensidade de tráfego processado pelo DNSpot.

[2] `https://www.unbound.net/`.
[3] `http://www.sqlite.org/`.
[4] `https://www.openbsd.org/`.
[5] Exemplos: `http://dnsresearch.cymru.com` e `http://openresolverproject.org`.

Transações	DS1		DS2		Total	
	Quantidade	%	Quantidade	%	Quantidade	%
Respondidas	488.289	12,1	1.600.386	4,9	2.088.675	5,7
– Válidas	391.050	*80,1*	1.280.425	*80,0*	1.671.475	*80,0*
– *ServFail* falso	97.249	*19,9*	319.961	*20,0*	417.210	*20,0*
Não respondidas	3.547.306	87,9	30.758.542	95,1	34.305.848	94,3
– Ignoradas	3.544.876	*99,9*	30.243.241	*98,3*	33.788.117	*98,5*
– Erros	2.370	*0,1*	515.301	*1,7*	517.671	*1,5*
Total	4.035.605	100,0	32.358.928	100,0	36.394.533	100,0

Tabela 2. Tráfego DNS processado pelo DNSpot. Porcentagens em itálico representam a subdivisão dentro de uma categoria.

A Tabela 3 apresenta o volume de tráfego processado pelo DNSpot. Ao todo foram processados 6.801,1 MB de tráfego pelo DNSpot, sendo 1.361,5 MB (20,0%) de consultas e 5.440,1 MB (80,0%) de respostas. Caso todas as consultas tivessem sido processadas, o tráfego esperado seria de quase 50 GB; restringir o número diário de consultas por IP levou a uma redução de 88,7% no tráfego total, mais uma vez comprovando a eficácia do mecanismo para limitar o tráfego gerado pelo DNSpot. A proporção de consultas aumentou de 10,6% em DS1 para 22,8% em DS2, levando a uma menor redução do tráfego de resposta no segundo *dataset*.

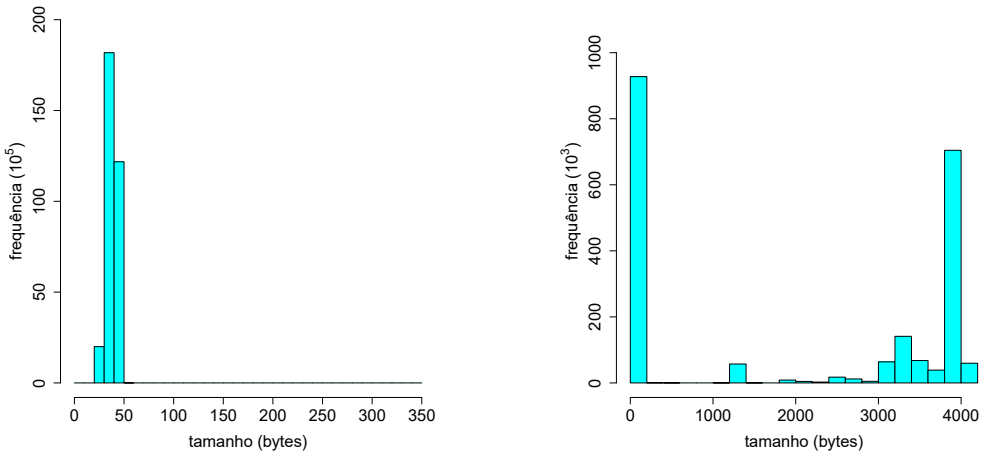
Tipo	DS1		DS2		Total	
	MB	%	MB	%	MB	%
Tráfego processado	1.560,1	100,0	5.241,5	100,0	6.801,6	100,0
– Consultas	165,1	*10,6*	1.196,4	*22,8*	1.361,5	*20,0*
– Respostas	1.395,0	*89,4*	4.045,1	*77,2*	5.440,1	*80,0*
Tráfego esperado	14.775,6	–	34.775,9	–	49.551,5	–
Respostas	14.610,4	–	33.579,5	–	48.189,9	–
Redução de tráfego de resposta	13.215,4	90,5	29.534,4	84,9	42.749,8	88,7

Tabela 3. Volume de tráfego processado e esperado.

Cada consulta especifica um registro de recurso (RR), que é a unidade básica de informação do DNS [Mockapetris 1987]. Cada RR é caracterizado pelo seu nome, tipo e classe, e possui um conteúdo que depende do tipo e da classe. A classe padrão é IN (Internet), e será implicitamente assumida no restante do texto. É possível ter vários RRs com o mesmo nome e tipo (*e.g.*, múltiplos registros A, que armazenam endereços IP), assim como vários RRs para o mesmo nome mas com tipos distintos. A ampla maioria das consultas observadas pelo DNSpot é por RRs com tipo ANY: 94,6% no total, sendo 99,2% em DS1 e 94% em DS2. O tipo ANY pode ser usado apenas em consultas, e recupera todos os RRs (de qualquer tipo) associados a um dado nome. Isso pode ser feito *(i)* para depurar domínios DNS, *(ii)* para obter múltiplas informações com uma única consulta, *(iii)* para descobrir potenciais alvos de ataques, ou *(iv)* para produzir respostas grandes a partir de consultas pequenas, amplificando o tráfego (este último caso é o mais pertinente a este estudo). Em função dos riscos associados, as especificações do DNS estão sendo alteradas para recomendar a limitação de respostas a consultas ANY [Abley et al. 2017]. Em contraste com o DNSpot, [Van Impe 2015] observou 51,3% de consultas ANY e 47,9% de consultas A; uma possível explicação é que, como seu *honeypot* enviava apenas respostas

falsas, ele pode ter sido menos usado para ataques DRDoS.

A distribuição dos tamanhos de consultas e respostas no *dataset* DS1 é fortemente distorcida pela dominância de um único registro de recurso RR, `hehehey.ru ANY`, presente em 97% das consultas (com tamanho de 39 bytes) e 90,4% das respostas (3.850 bytes). Em DS2 a variabilidade é maior: o tamanho das consultas possui uma distribuição assimétrica positiva (concentração de valores pequenos), enquanto o tamanho das respostas possui uma distribuição bimodal, conforme pode ser visto na Figura 3. 99,999% das consultas têm até 50 bytes, enquanto que, dentre as respostas, 43,9% têm até 100 bytes, 54,9% têm mais de 3.000 bytes, e 36,4% têm mais de 3.800 bytes. As respostas grandes podem ser associadas a tráfego DRDoS, enquanto que as pequenas devem-se a uma conjugação de *(i)* ataques DRDoS malsucedidos (consultas por nomes inexistentes ou que geram respostas pequenas), *(ii)* tráfego de sondagem (*probes*), e *(iii)* uso do DNSpot como servidor recursivo por usuários finais (que será discutido na Seção 4.4).



(a) Consultas (b) Respostas

Figura 3. Distribuição dos tamanhos de consultas e respostas no *dataset* DS2.

A Tabela 4 resume a distribuição geográfica dos endereços IP que originaram consultas para o DNSpot.[6] Essa análise não distingue entre clientes que realizam consultas e endereços de vítimas de ataques DRDoS, os quais usam IPs de origem forjados. Comparando os *datasets* DS1 e DS2, percebe-se que no segundo há uma maior diversidade, tanto em número de países quanto em concentração de requisições: o número de países aumentou de 73 para 161, e a proporção de endereços IP dos cinco países mais frequentes caiu de 84,0% para 36,5%. China e EUA aparecem nas duas primeiras posições em ambos *datasets*, mas a proporção desses dois países caiu de 57,2% para 24,1%. Os resultados de geolocalização para o *dataset* DS1 são semelhantes aos de [Van Impe 2015], o que faz sentido devido aos períodos de monitoração serem mais próximos.

4.3. Ataques DoS

É esperado que uma parte do tráfego direcionado a um servidor DNS recursivo aberto seja devido a ataques DRDoS em que o servidor esteja sendo usado como refletor. Nesta

[6]Os dados de geolocalização foram obtidos de `http://freegeoip.net/`.

	DS1			DS2		
Posição	País	IPs distintos	%	País	IPs distintos	%
1	China	1.287	30,0	China	28.705	15,6
2	Estados Unidos	1.164	27,2	Estados Unidos	15.613	8,5
3	Rússia	759	17,7	Brasil	9.838	5,3
4	Alemanha	297	6,9	Coreia do Sul	6.815	3,7
5	Canadá	94	2,2	Japão	6.398	3,5
	outros	686	16,0	outros	117.195	63,5
Total	*73 países*	4.287	100,0	*161 países*	184.564	100,0

Tabela 4. Top 5 países de origem das consultas ao DNSpot.

seção são caracterizados os ataques DRDoS em que o DNSpot foi usado como refletor. Para essa caracterização, foi adotada a seguinte definição de ataque DoS:

> *Um ataque DoS é formado por um conjunto com no mínimo 5 consultas com o mesmo IP de origem e com espaçamento máximo de 60 segundos entre consultas consecutivas, e pelas respostas a essas consultas.*

Essa definição foi estabelecida com base em uma análise preliminar do tráfego do DNSpot. Embora cinco consultas representem no máximo 20 KB de tráfego, esse número mínimo de consultas foi estabelecido considerando que o DNSpot esteja sendo usado como um dos vários refletores envolvidos em um ataque DRDoS, e não como um servidor recursivo regular. Cabe destacar ainda que, como ataques DRDoS usam endereços IP de origem forjados, do ponto de vista do DNSpot não é possível distinguir entre ataques DRDoS em que o *honeypot* é usado como refletor e ataques DoS diretos contra o *honeypot*; estes são bem menos prováveis, já que o sistema não hospeda nenhum serviço anunciado publicamente.

A Tabela 5 mostra o número de ataques DoS segundo a definição proposta, e as quantidades de endereços IP de origem, RRs e requisições envolvidos em tais ataques. Observa-se uma redução tanto na frequência de ataques quanto nas proporções de endereços IP, RRs e requisições envolvidos em ataques DoS. Foram identificados dois fatores que contribuem para essa redução. De um lado, no *dataset* DS1 houve uma concentração de requisições – os 97% de consultas por `hehehey.ru ANY` –, o que indica que o DNSpot foi intensamente usado como refletor nesse período, provavelmente por uma única ferramenta automatizada. De outro lado, em DS2 verificou-se a partir de maio de 2017 um aumento significativo da proporção de requisições que não se encaixam na definição de ataque, o que ajuda a explicar as quedas mais expressivas, de endereços IP e RRs envolvidos em ataques. Essa mudança de característica é discutida na Seção 4.4; como ela é recente, é preciso acompanhar sua evolução para ver se é algo efêmero ou mais duradouro.

Em ataques DRDoS, o fator de amplificação mede quanto tráfego de resposta um refletor gera em função de um dado tráfego de entrada. Nos ataques usando o DNS, esse fator é dado pela razão entre o tamanho de respostas e o tamanho das consultas correspondentes. A Tabela 6 mostra os fatores de amplificação observados nos dois *datasets*. O fator máximo foi comparável nas duas coletas, mas o fator médio foi menor em DS2. Essa redução deve-se a um número expressivo de consultas em DS2 (mais de 2 milhões, ou 6,5% do total) para RRs com fator de amplificação inferior a 10, que pode ser explicada por nomes equivocados (Seção 4.4) e por medidas de contenção do tamanho de

	DS1		DS2	
N° de ataques	7.940		23.788	
Ataques/dia (média)	160,1		95,0	
Métricas	Total	Envolvidos em DoS	Total	Envolvidos em DoS
IPs	4.287	3.499 (81,6%)	184.564	23.745 (12,9%)
RRs	136	87 (64,0%)	4.982	840 (16,8%)
N° de requisições	4.035.605	4.032.778 (99,9%)	32.358.928	30.661.228 (94,7%)

Tabela 5. Métricas envolvidas em ataques DoS contra o DNSpot

respostas, como mudanças no processamento de consultas com tipo ANY em servidores DNS autoritativos [Abley et al. 2017]. No geral, os maiores fatores de amplificação foram observados para consultas ANY cujas respostas contêm diversos registros DNSSEC, corroborando o observado em [Anagnostopoulos et al. 2013, van Rijswijk-Deij et al. 2014, MacFarland et al. 2017].

Fator de amplificação	DS1	DS2
Médio	96,3	74,1
Máximo	110,7	103,6

Tabela 6. Fator de amplificação dos ataques DoS

A Figura 4 mostra as distribuições empíricas da duração dos ataques nos dois *datasets*. Em DS1, 95% dos ataques duraram até 538 s (menos de 9 min). Em DS2, 50% dos ataques duraram até 480 s (8 min), e 25% duraram 1.140 s (19 min) ou mais. Isso significa que, embora a frequência de ataques DoS tenha caído, a duração dos ataques foi maior. O número de requisições por ataque também aumentou bastante, conforme mostrado na Tabela 7; a mediana foi 17,1 vezes maior, e as demais medidas aumentaram cerca de 12 vezes. Em ambos os *datasets*, a ampla maioria dos endereços IP esteve envolvida (i.e., foi vítima) em apenas um ataque, sendo que os máximos observados foram de 73 ataques para um único IP em DS1 e 41 ataques em DS2.

Dataset	média	mediana	3° quartil	95° percentil	máximo
DS1	507,9	132,0	402,0	2.100,2	25.363
DS2	6.029,7	2.264,0	5.024,0	25.594,4	203.474

Tabela 7. Estatísticas de requisições por ataque DoS

4.4. Resultados Notáveis

Durante a operação do DNSpot, foram observados alguns resultados notáveis do ponto de vista qualitativo, que serão descritos a seguir.

Varreduras UDP e SIP. Foram recebidas requisições malformadas, que não respeitavam o formato das mensagens DNS. Algumas dessas requisições tinham conteúdo vazio, e foram classificadas como varreduras de portas UDP, como as realizadas pelo Nmap[7]. Também foram identificadas varreduras do protocolo SIP (*Session Initiation Protocol*),

[7] `https://nmap.org/`.

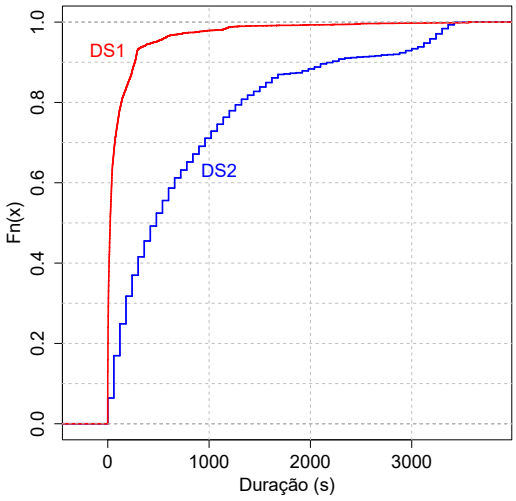


Figura 4. Distribuições empíricas da duração de ataques

que buscavam terminais e servidores de telefonia IP usando a porta 53/UDP, sendo que a porta padrão para sinalização SIP é 5060/UDP.

Domínios projetados para amplificação. Foram encontrados diversos domínios contendo RRs que não possuem nenhum significado ou utilidade para uma consulta normal DNS, cujo único propósito é gerar respostas grandes (próximas a 4 KB), úteis para ataques DRDoS, algo que já havia sido descrito na literatura (*e.g.*, [Anagnostopoulos et al. 2013]). No DNSpot foram observados nomes com mais de 250 registros A pertencentes a uma mesma sub-rede, e nomes com mais de 30 registros TXT cujo conteúdo é a letra "x" repetida 99 vezes combinada com um sufixo numérico. Alguns desses nomes pertencem a domínios distintos mas possuem estrutura idêntica, com os mesmos registros SOA, NS, MX e A associados ao nome, o que indica que foram criados pela mesma pessoa ou grupo.

Desaparecimento e redução no tamanho de domínios. Durante os períodos de observação, foram constatadas mudanças no conteúdo dos domínios consultados que levaram à redução no fator de amplificação. No *dataset* DS1, foi verificado o desaparecimento do quarto domínio com mais consultas, que teve 13.455 requisições. Inicialmente, esse domínio gerou 2.970 respostas com 3.875 bytes, e posteriormente houve 2.316 respostas com 96 bytes, que acusavam domínio inexistente; aparentemente o domínio continuou sendo usado por ferramentas automatizadas para ataques DRDoS mesmo depois de seu desaparecimento (que ocorreu pela expiração do registro do domínio). Em DS2 desapareceram 17 RRs. Diferente do que ocorreu com DS1, em DS2 foram registradas poucas consultas após o desaparecimento dos RRs, o que sugere que as ferramentas de ataque passaram a atualizar com maior frequência suas listas de RRs usados em ataques DRDoS.

Consultas por nomes equivocados. Em ambos os *datasets* foram observadas quantidades expressivas de consultas por domínios de primeiro nível (DPNs) obviamente inexistentes, como `.3858` (13,5 mil consultas em DS1) e `.pkt` (1,74 milhões de consultas em DS2). Foram também registradas consultas com tipos aparentemente trocados; por exemplo, em DS1 houve 6,5 mil requisições para um registro TXT inexistente, sendo que uma consulta pelo mesmo nome com tipo ANY gerava uma resposta de 4.071 bytes. Acredita-se que essas consultas por nomes equivocados sejam fruto de erro na configuração de ferramentas usadas em ataques DRDoS. Incidências significativas de consultas por DPNs inexistentes também foram observadas em servidores DNS raiz [Castro et al. 2010], tendo sido atribuídas a roteadores domésticos mal configurados.

Consultas de usuários finais. No *dataset* DS2 foi observada uma quantidade expressiva de consultas por nomes tipicamente associados a usuários finais, incluindo *sites* populares (*e.g.*, `google.com`, `facebook.com`, `amazon.com`) e nomes usados por ferramentas anti-*malware* (*e.g.*, `avqs.mcafee.com`). Essas consultas, cujo fator de amplificação é baixo (< 10), sugerem que usuários estão usando o DNSpot como resolvedor DNS padrão. Elas tiveram um aumento significativo em maio de 2017, e estão associadas a um incremento no número de endereços IP distintos que interagiram com o *honeypot*. Também foram observadas consultas por nomes usados para descoberta de serviços de rede (DNS-SD, *DNS-Based Service Discovery*) [Cheshire and Krochmal 2013], no caso com endereços de redes privadas (192.168.*.*). Essas consultas podem ser reflexo do uso do DNSpot como resolvedor regular, embora também possam tratar-se de varreduras usando DNS-SD em busca de alvos para ataques [Atlasis 2017].

4.5. Discussão dos Resultados

Uma análise dos dados coletados pelo DNSpot, considerando de forma conjunta as diferentes métricas observadas, permite chegar a algumas conclusões importantes:

- Ataques DRDoS constituem o principal abuso envolvendo servidores DNS recursivos abertos. A escolha adequada dos nomes consultados permite obter um fator de amplificação superior a 100, consideravelmente maior que os fatores típicos reportados para o DNS, entre 28 e 54 [CERT.br 2016]. Nesse sentido, o mecanismo de limitação diária de consultas, embora pouco sofisticado, foi eficaz em restringir o tráfego de ataque obtido com o abuso do DNSpot a 11,3% do volume pretendido.
- O volume de requisições maliciosas, especialmente considerando o fato de ser um servidor não anunciado publicamente, é significativo: considerando as duas coletas, o DNSpot recebeu em média 1,4 requisições por segundo, com um tráfego médio diário potencial de 165,1 MB. O DNSpot começou a receber as primeiras requisições segundos após o início do seu funcionamento, e menos de 28 h depois já estava sendo usado em ataques DRDoS.
- Uma análise da evolução dos ataques DoS entre as duas coletas revela uma diminuição na frequência dos ataques, contrabalançada por um aumento significativo na sua duração e intensidade. Também foi possível constatar uma redução no fator de amplificação médio dos ataques, embora não seja possível afirmar se tal redução foi provocada por uma mudança na configuração de servidores DNS auto-

ritativos (especialmente o tratamento de consultas do tipo ANY) ou por equívocos cometidos pelos atacantes.

5. Conclusão

Este artigo introduz o DNSpot, um *honeypot* específico para DNS que tem por objetivo propiciar a monitoração de requisições DNS sob o prisma de servidores recursivos abertos. O artigo também apresenta uma análise de dados coletados do DNSpot durante dois períodos, em 2015 (49 dias) e 2016–2017 (250 dias), no qual foram processadas 36,4 milhões de requisições. O estudo demonstrou a incidência de ataques DRDoS usando um servidor DNS recursivo aberto, e o alto fator de amplificação oferecido pelo DNS. Além disso, foram descobertos aspectos pouco conhecidos, como a existência de domínios projetados para amplificação. A comparação entre os dois períodos de coleta revelou algumas mudanças relevantes no tráfego, que incluem:

- aumento na intensidade do tráfego direcionado ao DNSpot;
- aumento na duração e no volume de tráfego em ataques DDoS que usam o DNSpot como refletor;
- evidências de uso do DNSpot como servidor DNS recursivo regular por parte de usuários finais ou em nome destes.

Para a continuidade deste trabalho, algumas perspectivas podem ser apontadas, como manter a coleta de dados do DNSpot por períodos ainda maiores, conjugando análises de longo e curto prazo, e investigar como múltiplas instâncias do DNSpot, em diferentes pontos da Internet, podem ser usadas em um sistema distribuído de monitoramento e detecção de ataques ao DNS.

Referências

[Abley et al. 2017] Abley, J., Gudmundsson, O., and Majkowski, M. (2017). Providing minimal-sized responses to DNS queries that have QTYPE=ANY. IETF Draft draft-ietf-dnsop-refuse-any-04 (Proposed standard). Disponível em `https://tools.ietf.org/html/draft-ietf-dnsop-refuse-any-04`.

[Akamai 2017] Akamai (2017). Q1 2017 state of the Internet/security report. Technical report. Disponível em `http://www.akamai.com/`.

[Anagnostopoulos et al. 2013] Anagnostopoulos, M., Kambourakis, G., Kopanos, P., Louloudakis, G., and Gritzalis, S. (2013). DNS amplification attack revisited. *Computers & Security*, 39:475–485.

[Arbor 2017] Arbor (2017). Worldwide infrastructure security report, vol. XII. Technical report, Arbor Networks. Disponível em `http://www.arbornetworks.com/`.

[Atlasis 2017] Atlasis, A. (2017). An attack-in-depth analysis of multicast DNS and DNS service discovery. In *Hack in the Box*, Amsterdam. Disponível em `http://tinyurl.com/ycybxp59`.

[Barbosa and Pereira 2009] Barbosa, K. R. and Pereira, E. S. J. (2009). Análise passiva do tráfego DNS da Internet brasileira. In *Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais (SBSeg)*, pages 203–216.

[Brownlee et al. 2001] Brownlee, N., Claffy, K. C., and Nemeth, E. (2001). DNS measurements at a root server. In *IEEE Global Telecommunications Conference (GLOBECOM)*, San Antonio, TX.

[Castro et al. 2010] Castro, S., Zhang, M., John, W., Wessels, D., and Claffy, K. C. (2010). Understanding and preparing for DNS evolution. In *Traffic Monitoring and Analysis Workshop (TMA)*, pages 1–6, Zurich, Switzerland.

[CERT.br 2016] CERT.br (2016). Recomendações para melhorar o cenário de ataques distribuídos de negação de serviço (DDoS). Disponível em http://www.cert.br/docs/whitepapers/ddos/.

[Cheshire and Krochmal 2013] Cheshire, S. and Krochmal, M. (2013). DNS-based service discovery. RFC 6763.

[Conrad 2012] Conrad, D. (2012). Towards improving DNS security, stability, and resiliency. Technical report, Internet Society.

[Danzig et al. 1992] Danzig, P. B., Obraczka, K., and Kumar, A. (1992). An analysis of wide-area name server traffic: a study of the Internet Domain Name System. *ACM SIGCOMM Computer Communication Review*, 22(4):281–292.

[Fachkha et al. 2015] Fachkha, C., Bou-Harb, E., and Debbabi, M. (2015). Inferring distributed reflection denial of service attacks from darknet. *Computer Communications*, 62:59–71.

[Gao et al. 2013] Gao, H., Yegneswaran, V., Chen, Y., Porras, P., Ghosh, S., Jiang, J., and Duan, H. (2013). An empirical reexamination of global DNS behavior. *ACM SIGCOMM Computer Communication Review*, 43(4):267–278.

[Jung et al. 2002] Jung, J., Sit, E., Balakrishnan, H., and Morris, R. (2002). DNS performance and the effectiveness of caching. *IEEE/ACM Transactions on Networking*, 10(5):589–603.

[Kührer et al. 2015] Kührer, M., Hupperich, T., Bushart, J., Rossow, C., and Holz, T. (2015). Going wild: Large-scale classification of open DNS resolvers. In *Internet Measurement Conference (IMC)*, pages 355–368, Tokyo, Japan.

[MacFarland et al. 2017] MacFarland, D. C., Shue, C. A., and Kalafut, A. J. (2017). The best bang for the byte: Characterizing the potential of DNS amplification attacks. *Computer Networks*, 116:12–21.

[Mockapetris 1987] Mockapetris, P. (1987). Domain names – concepts and facilities. RFC 1034.

[Perdisci et al. 2009] Perdisci, R., Corona, I., Dagon, D., and Lee, W. (2009). Detecting malicious flux service networks through passive analysis of recursive DNS traces. In *Annual Computer Security Applications Conference (ACSAC)*, pages 311–320, Honolulu, HI.

[Steding-Jessen et al. 2008] Steding-Jessen, K., Vijaykumar, N. L., and Montes Filho, A. (2008). Using low-interaction honeypots to study the abuse of open proxies to send Spam. *InfoComp*, 7(1):44–52.

[Takano et al. 2013] Takano, Y., Ando, R., Takahashi, T., Uda, S., and Inoue, T. (2013). A measurement study of open resolvers and DNS server version. In *Internet Conference (IEICE)*.

[Tanasi 2014] Tanasi, A. (2014). Homemade custom interaction DNS honeypot. Disponível em http://tinyurl.com/y955ohns.

[Van Impe 2015] Van Impe, K. (2015). Analyzing queries on a honeypot name server for better DNS log quality. *Security Intelligence*. Disponível em http://tinyurl.com/ybsze9zd.

[van Rijswijk-Deij et al. 2014] van Rijswijk-Deij, R., Sperotto, A., and Pras, A. (2014). DNSSEC and its potential for DDoS attacks: A comprehensive measurement study. In *Internet Measurement Conference (IMC)*, pages 449–460, Vancouver, BC, Canada.

[Zdrnja et al. 2007] Zdrnja, B., Brownlee, N., and Wessels, D. (2007). Passive monitoring of DNS anomalies. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA)*, pages 129–139, Lucerne, Switzerland.

[Zhao et al. 2015] Zhao, G., Xu, K., Xu, L., and Wu, B. (2015). Detecting APT malware infections based on malicious DNS and traffic analysis. *IEEE Access*, 3:1132–1142.

