

Sender Authentication Technology Update

Steven M Jones
Executive Director, DMARC.org
Senior Software Engineer, LinkedIn
<http://linkedin.com/in/stevenmjones>

JPAAWG 2nd General Meeting
ペルサール飯田橋ファースト B1F
2019.11.14
Session A7, Hall A

Topics

- Why Do We Focus On Sender Authentication?
- Standards and Protocols
- DMARC Use Update
- Common Problems With DMARC Records

Why Do We Focus On Sender Authentication?

- Easier To Identify Legitimate Email
- Best Practices = Better Delivery
- Undelivered Mail = Wasted ¥

- Criminals Exploit Email Effectively
 - Phishing is #1 Cause - Data Breach
 - Business Email Compromise

BEC Scam Costs Media Giant Nikkei \$29 Million



Author:
Lindsey O'Donnell
November 4, 2019
10:34 am

In September, a Nikkei America employee transferred \$29 million to BEC scammers who were purporting to be a Nikkei executive.

Media conglomerate Nikkei Inc. has fallen victim to a business email compromise (BEC) scam that fleeced the company out of \$29 million.

3 minute read

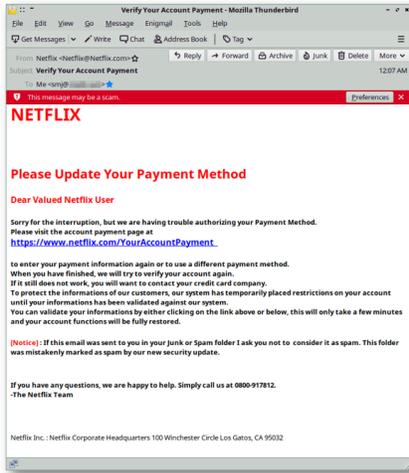
BEC totals:

<https://www.tripwire.com/state-of-security/security-data-protection/bec-scams-cost-victims-26b-over-a-three-year-period-finds-fbi/>

Phishing -> data breach:

<https://info.phishlabs.com/blog/phishing-number-1-data-breaches-lessons-verizon>

Easier Detection, Better Protection



```
From: Netflix <Netflix@Netflix.com>
Authentication-Results: XXX.XXXXX.com/xACJ8inv058374;
  dmarc=fail (p=reject dis=none) header.from=Netflix.com
Authentication-Results: XXX.XXXXX.com; spf=fail smtp.mailfrom=Netflix@Netflix.com
Authentication-Results: XXX.XXXXX.com; dkim=pass (2048-bit key; unprotected)
  header.d=uttarauniversity.edu.bd header.i=@uttarauniversity.edu.bd
  header.b=kU8F/hqO
```

- Consistent authentication makes your legitimate email stand out, easy to model
- Machine Learning leverages this to detect cousin domains / “display name” attacks

Standards and Protocols



Overview of Common Protocols



- Sender Policy Framework (SPF)
RFC 7208
- Domain Keys Identified Message (DKIM)
RFC 6376
- Domain-based Message Authentication, Reporting & Conformance (DMARC)
RFC 7489
- Authenticated Received Chain (ARC)

SPF – <http://www.open-spf.org>

DKIM – <http://dkim.org>

DMARC – <https://dmarg.org>

ARC – <http://arc-spec.org>

Refining Protections Over Time

SPF: Combat “backscatter” from spamming	2002 – 2004
- Left header From: unprotected	
- Easily misconfigured, rarely enforced	
DKIM: Protect header From:, message forgery	2004 – 2007
- No accepted policy mechanism	
- Third-party signatures problematic	
DMARC: Has policy mechanism, enforced at ISP	2009 - 2015
- Cousin domains and “display name” attacks	
- Problems with mailing lists, forwarding	

Example of an Indirect Mail Flow



- Intermediary sends the message from a new IP address, causing SPF to fail to verify for Sender's domain
- Intermediary changes the message contents (subject:), causing Sender's DKIM signature to fail to verify

ARC Protocol

- ARC assists with authentication of “indirect mailflows”
- Under development since 2014
- Part of the IETF DMARC Working Group since 2016
- 8+ Interoperability testing sessions
- Draft of Usage Guide / FAQ available



ARC is being developed by the IETF
DMARC working group

See

<https://datatracker.ietf.org/wg/dmarc/>

Usage Guide:

<https://datatracker.ietf.org/doc/draft-ietf-dmarc-arc-usage/>

More information at <http://arc-spec.org>

ARC Published As RFC 8617 on 2019.07.09



ARC Implementations

- FastMail, Google, Microsoft – hosted email services
- Cloudmark ,Halon, MailerQ and MessageSystems (SparkPost) – Mail Transfer Agent (MTA)
- Mailman and Sympa - Mailing List Manager (MLM)
- Free Software – dkimpy, Mail::DKIM, OpenARC
- More at arc-spec.org → Resources



ARC being developed by the IETF DMARC working group

See

<https://datatracker.ietf.org/wg/dmarc/>

Microsoft and ARC

- 2019.10.24 – Announces ARC support on Microsoft 365 Roadmap
- Testing in May 2019
- Began using production key in July
- Messages from many Office 365 tenants sent with ARC headers since July



Microsoft announcement:

<https://www.microsoft.com/en-us/microsoft-365/roadmap?filters=Launched&searchterms=57057>

Microsoft logo image obtained from Wikimedia

Looking For Users of ARC

- ARC supports mailing lists – look there
- `arc-discuss@dmARC.org` mailing list
 - First message 2018.01.31 from an OpenARC user
 - 8.6% of posts have included an ARC Seal
- IETF's `ietf@ietf.org` mailing list
 - First message 2019.06.25 from Office 365 customer
 - 4.7% of posts have included an ARC Seal

ARC Discussion list:

<http://lists.dmarc.org/mailman/listinfo/arc-discuss>

IETF lists can be found at

<https://www.ietf.org/how/lists>

`ietf@ietf.org` list:

<https://www.ietf.org/mailman/listinfo/ietf>

RFC 8616: Email Authentication for Internationalized Mail

- Use of Unicode characters in domains and email addresses has been evolving
- RFC 8616 updates the core SPF, DKIM and DMARC specifications to clarify which form of Internationalized Domain Name (IDN) each uses
- Published on 2019.06.30

名がドメイン.co.jp
xn--v8jxj3d1dzdz08w.co.jp

<https://tools.ietf.org/html/rfc8616>

DMARC and Public Suffix Domains

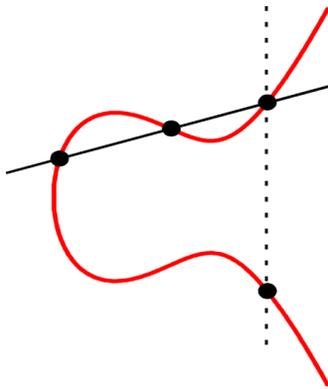
- Allow for DMARC to be applied at ccTLD, like `.uk` or `.jp`
- Also cover intermediate domains, ex. `gov.uk`
- Allow TLDs to have a DMARC policy for *non-existent* domains, ex. `nodomain.gov.uk`
- Proposed at M³AAWG 44 (Brooklyn) in 2018.10
- Several revisions in the IETF DMARC Working Group
- Nearing publication (as of November 2019)

PSD draft:

<https://datatracker.ietf.org/doc/draft-ietf-dmarc-psd/>

Cryptography Changes From 2018

Changes in DKIM Cryptography (RFC 8463)



- RSA algorithm was deprecated under RFC 8017
- Elliptic Curve signing algorithm standardized under RFC 8031
- DKIM may now use PureEdDSA variant Ed25519
- Smaller keys for equivalent strength

Shorter DKIM keys should be easier for small domains to publish in DNS without errors – see issues with 2,048 bit keys being too long for some DNS software/services.

Quantum Computing and Encryption

- 2019.10.23 – Google claims “Quantum Supremacy”
- What are the implications for traditional cryptography
- M3AAWG 46 (Montreal) had sessions on this topic
- Impacts most online activity, communications
- Directly impacts DKIM and ARC; indirectly DMARC
- How quickly can the IETF address this issue?

Google announcement:

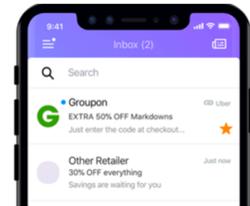
<https://www.blog.google/technology/ai/computing-takes-quantum-leap-forward/>

The Verge article on same:

<https://www.theverge.com/2019/10/23/20928294/google-quantum-supremacy-sycamore-computer-qubit-milestone>

BIMI

- Brand Indicators for Message Identification (BIMI)
- Email clients would show sender's logo with messages
- Entrust Datacard issued first Verified Mark Certificate (VMC) in September 2019
- Yahoo US running a trial; Google in 2020
<https://www.brandindicators.org>



Note that JIPDEC, Yahoo Japan, and GMX.de / 1&1 have been doing this for some time

<https://www.entrustdatacard.com/about/newsroom/press-releases/2019/verified-mark-certificates>

<https://www.infosecurity-magazine.com/news/verified-mark-certificate-issued/>

DMARC Use Update

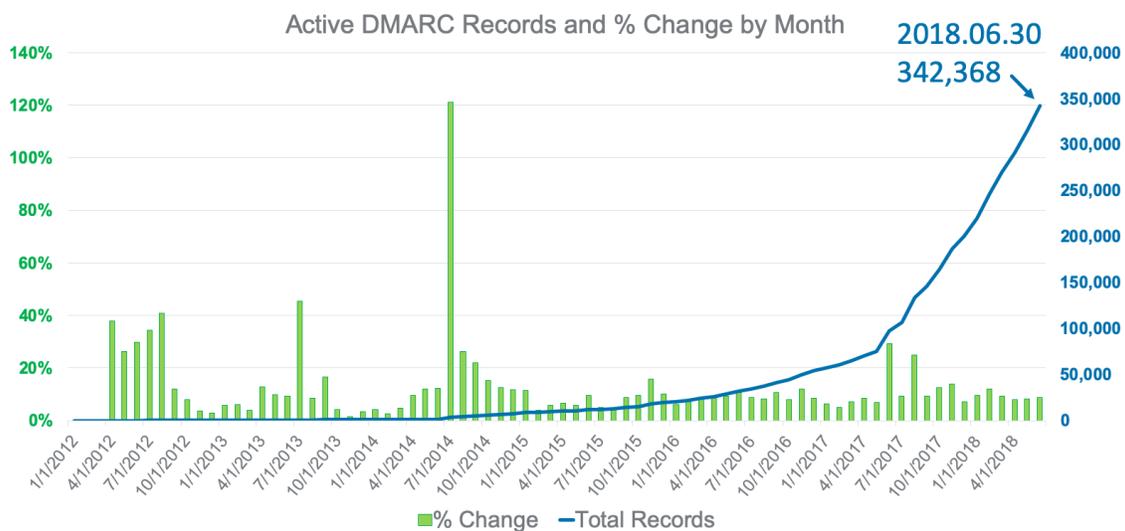


Farsight Security DNS Data

- Sensors located at network providers around the world
- Response data – the answer – is timestamped and stored
- Sensors only see records when somebody looks them up

- DMARC.org only includes valid records still published in DNS, and are tracked by when they were first published
 - The set of active records changes over time

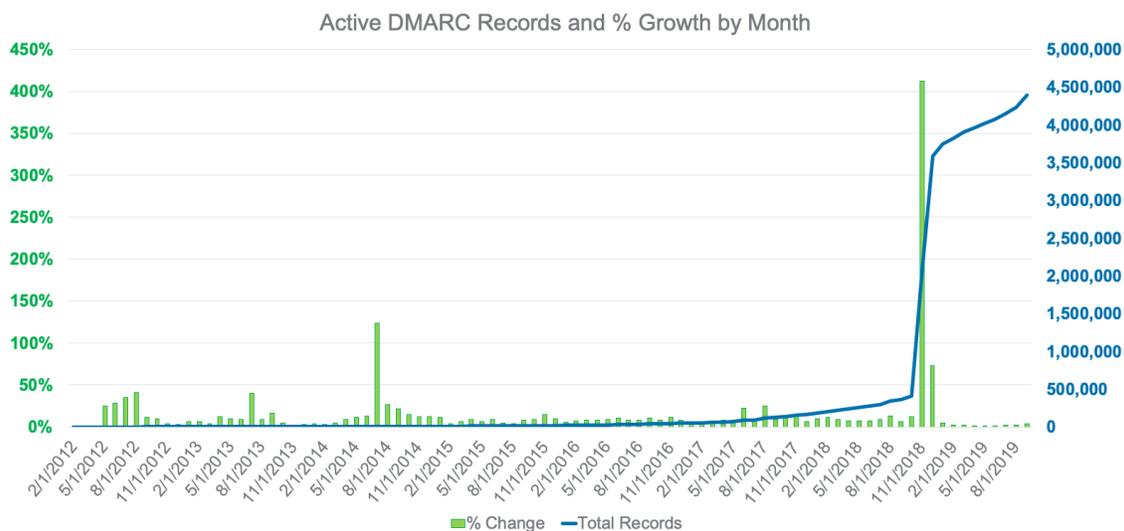
Active DMARC Records – 2Q 2018



Data provided by Farsight Security
Graph © 2018 Trusted Domain Project

~194,000 in August 2017, ~75% increase

Active DMARC Records – 3Q 2019



Data provided by Farsight Security
Graph © 2019 Trusted Domain Project

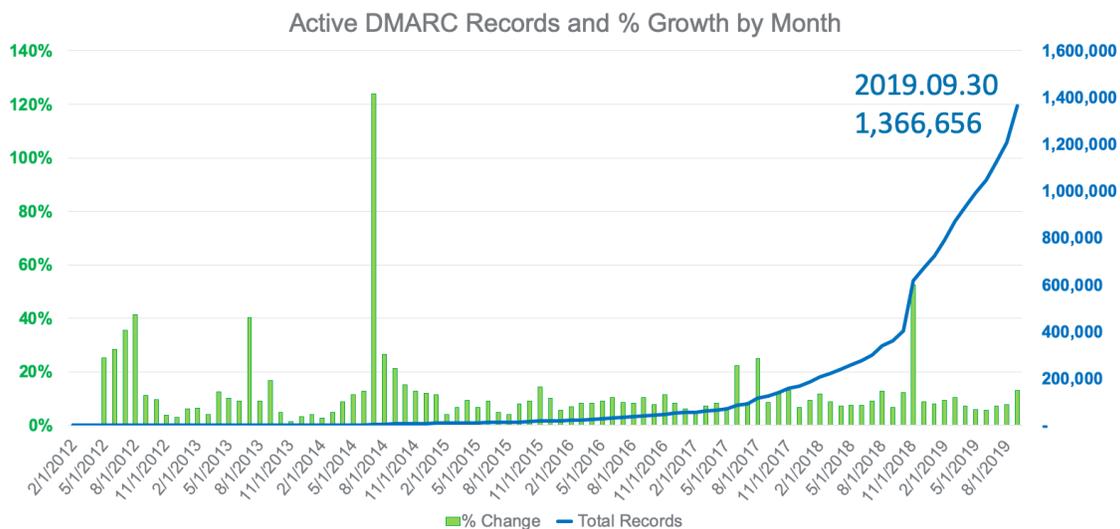
This graph shows 4.4MM active DMARC records as of 9/30/2019

3 Million New DMARC Records?

- Millions of DMARC records with strange names
 - `_dmarc.mx.mx.mx.mx.mx.ichiban.example.com`
- Most appear to trace back to “X”
- Nobody was aware of “X” behaving badly
- Exclude these records for now...



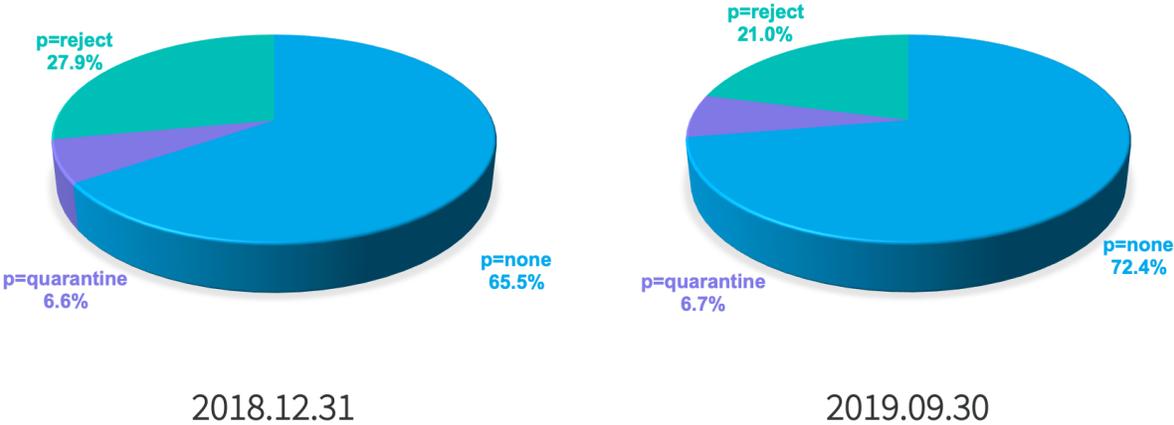
Active DMARC Records – 3Q 2019



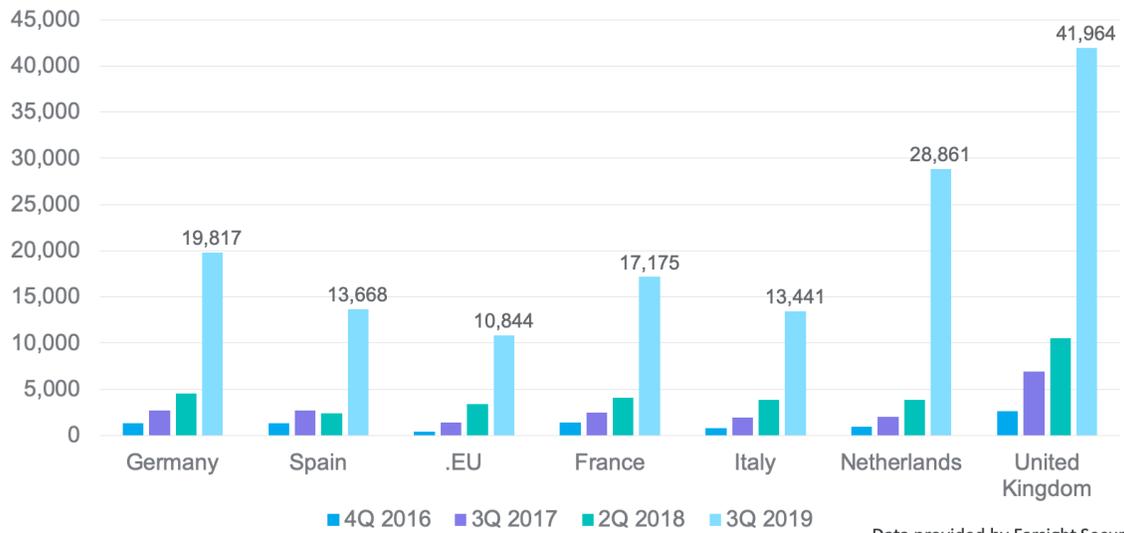
Data provided by Farsight Security
Graph © 2019 Trusted Domain Project

Current ending figure is 1.37MM active DMARC records

Policy Breakdown of Active DMARC Records

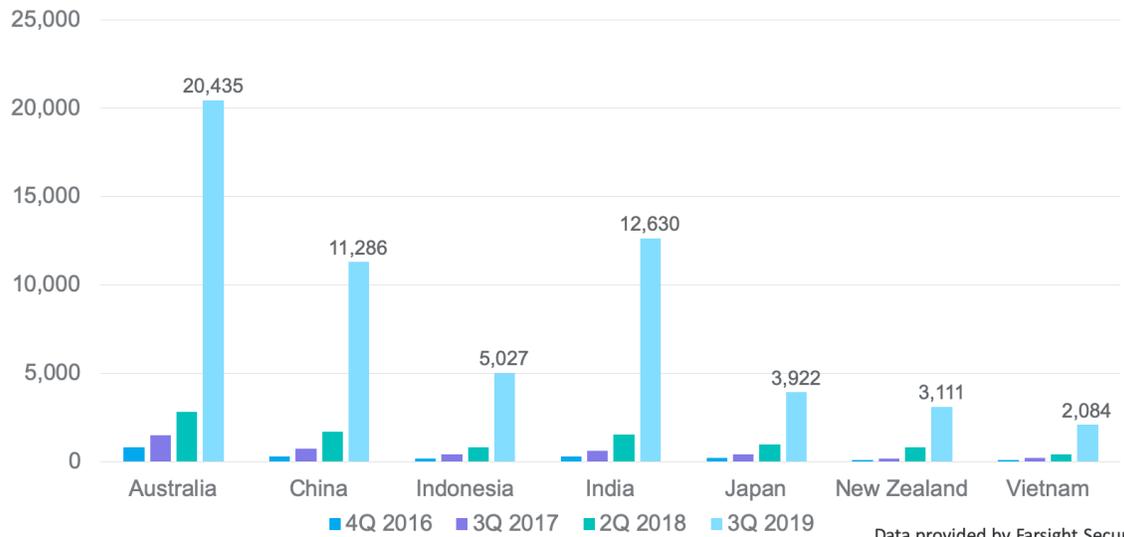


Active DMARC Records in Euro ccTLDs



Data provided by Farsight Security
Graph © 2019 Trusted Domain Project

Active DMARC Records in Asia ccTLDs

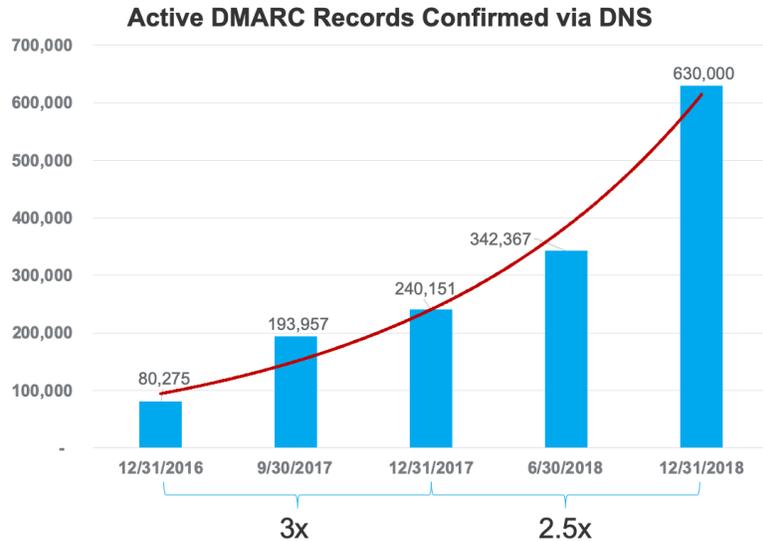


Data provided by Farsight Security
Graph © 2019 Trusted Domain Project

DMARC Records Increase 2.5x Year-over-Year

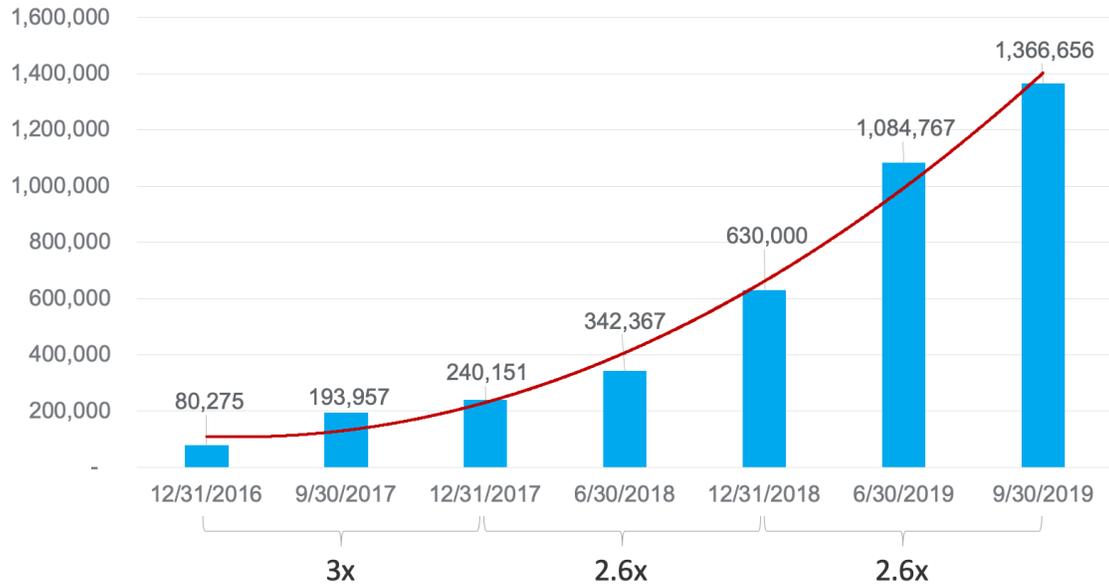
- Cumulative counts confirmed in DNS for the periods ending
- Robust growth
- Nearly doubled in 2H2018 alone
- Excluding 5MM suspicious records created in 4Q2018

Raw Data: Farsight Security
Analysis: DMARC.org



Provided to a colleague for a session at M3AAWG 44 in San Francisco, February 2019

Active DMARC Record Growth



Common Problems with DMARC Records

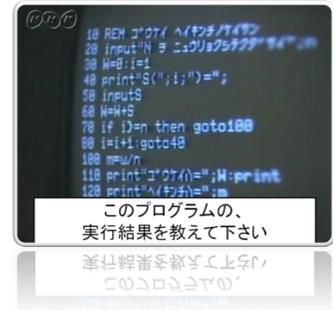
Problems with DMARC Records

- 2012-2016: 489,000 bad TXT records (`_dmarc...`)
- 2017-2019: 446,000 bad TXT records
- Many are non-DMARC “wildcard” records
 - 76,000 `bio=<base64>`
 - 42,000 `google-site-verification`
 - 25,000 `v=dmarc1` (must be `v=DMARC1`)
 - 11,000 `MS=ms[0-9]*`

2019 Q4 data – 3.8MM data records
representing 96MM DNS responses

Problems with DMARC Records

- Many bad records are formatting issues in rdata
 - `\"v=DMARC1`
 - `v= DMARC1 ...`
 - `V-DMARC`
 - `Value: V=DMARC1; ...`
 - `_dmarc... IN TXT \"v=DMARC1 ...`

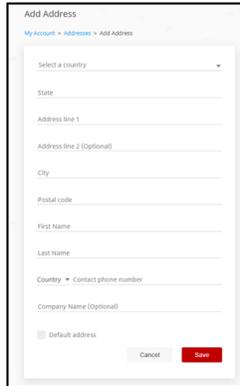


```
10 REM ココイ ハイフン/7470
20 input"N 0 ココイ/5555555555555555"
30 M=0:i=1
40 print"S(i;i)=";
50 inputS
60 M=M+S
70 if i=9 then goto100
80 i=i+1:goto40
100 M=M/n
110 print"ココイ)=";M:print
120 print"ハイフン)=";M
```

このプログラムの、
実行結果を教えてください

実行結果を教えてください
このプログラムの、

Problems (?) with DMARC Records



The image shows a web form titled "Add Address". It contains the following fields: "Select a country" (dropdown), "State" (text), "Address line 1" (text), "Address line 2 (Optional)" (text), "City" (text), "Postal code" (text), "First Name" (text), "Last Name" (text), "Country" (dropdown) with a sub-label "Contact phone number", and "Company Name (Optional)" (text). At the bottom, there is a checkbox for "Default address", a "Cancel" button, and a red "Save" button.

- Policy records with no reporting address
 - “v=DMARC1; p=none”
 - p=reject and no reporting, may be intentional
 - p=none and no reporting...?
- p=none intended to generate reports
- Does this really qualify as deploying DMARC?

Example – NetEase has a “v=DMARC1; p=none” policy published for 126.com and 163.com, not clear why

Problems with DMARC Records

- Bad mailto: URIs in published policy
 - `rua=mailto:devops`
 - `rua=mailto:rua [] example.com`
 - `rua=user@domain` not `rua=mailto:...`
- Not just missing reports, may harass reporter
- Potential privacy violations



- Not Deliverable As Addressed
- Unable To Forward
- Insufficient Address
- Moved, Left No Address
- Unclaimed Refused
- Attempted - Not Known
- Not Such Street Number
- Vacant Illegible
- No Mail Receptacle
- Box Closed - No Order
- Returned For Better Address
- Postage Due _____

Verifying 3rd Party Report Receivers

- Domain owners publish *authorizing records* under RFC 7489 Section 7.1
 - `foo.com` wants DMARC reports sent to `bar.com`
 - `_dmarc.foo.com = "rua=mailto:foo@bar.com"`
 - **`foo.com._report._dmarc.bar.com = "v=DMARC1"`**
- Report generators are not checking
- Big privacy and legal implications

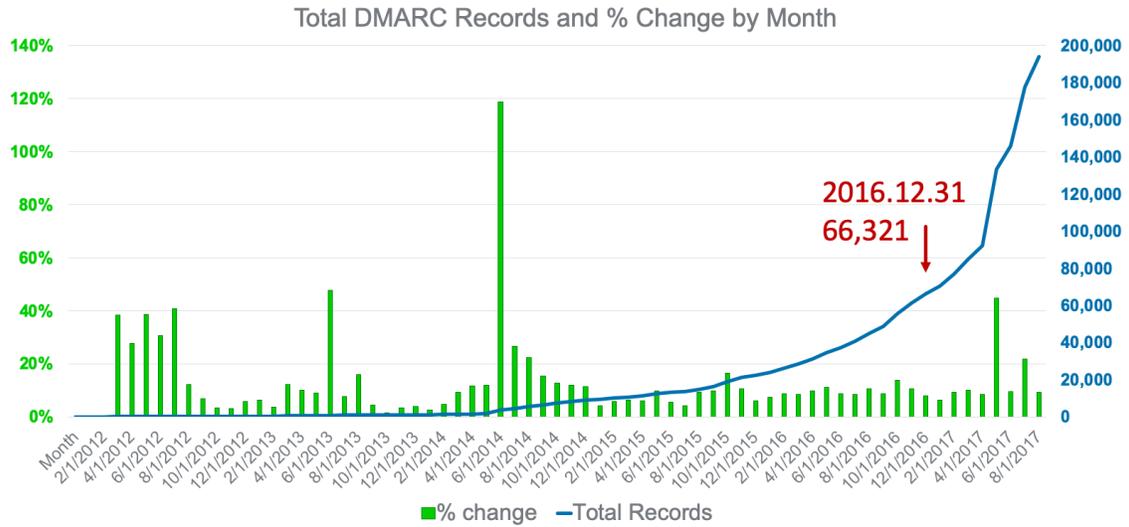
Q & A

—

“Why Do Your Numbers
Change?”

—

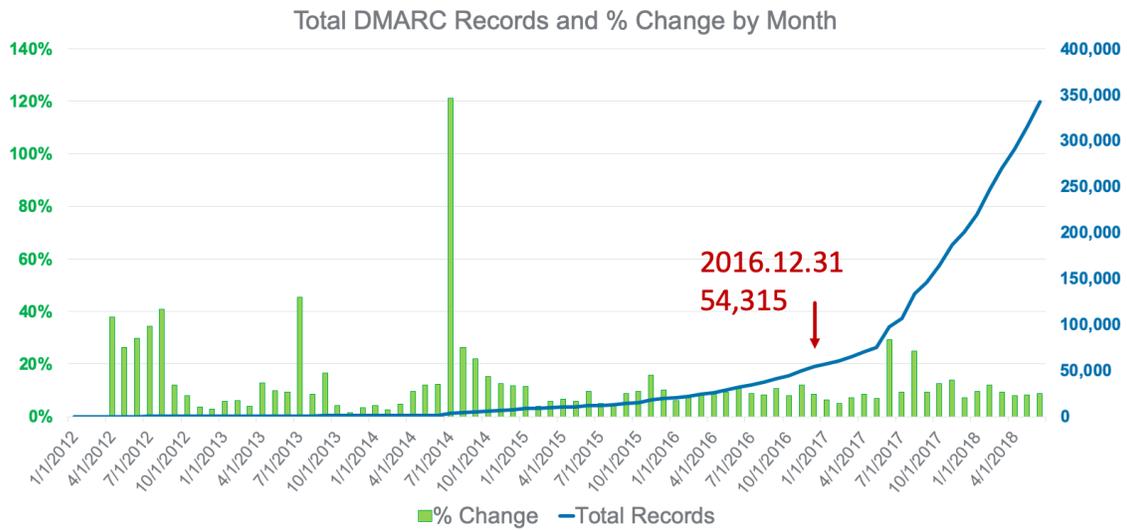
Growth of DMARC Adoption Globally – 3Q 2017



Data provided by Farsight Security
Graph © 2017 Trusted Domain Project

Only 62,000 ending same quarter in 2016

Growth of DMARC Adoption Globally – 2Q 2018



Data provided by Farsight Security
Graph © 2018 Trusted Domain Project

~75% increase from August 2017

Farsight Security DNS Data

- Sensors located at network providers around the world
- Response data – the answer – is timestamped and stored
- Sensors only see records when somebody looks them up
- **DMARC.org only includes valid records still published in DNS, and are tracked by when they were first published**
 - The set of active records changes over time

Why Do The Counts Change Over Years?

- `ichi.com` and `ni.com` publish DMARC records during 2015
- They are both still published as of 2015.12.31, so the total for 2015 as of 2015.12.31 is 2
- During 2016 `ni.com` removes its DMARC record, but `san.com` publishes a DMARC record
- The total for 2015 as of 2016.12.31 is 1, and the count for 2016 as of 2016.12.31 is 1.
- During 2016 `ichi.com` removes its DMARC record
- The count for 2015 as of 2017.12.31 is 0, and the count for 2016 is 1

Concrete Example

- As of 2017.09.30: We reported 66,321 DMARC records for 2016.12.31
- As of 2018.06.30: We reported 54,315 DMARC records for 2016.12.31
- 12,006 records that were active during the 2017.09.30 validation were no longer active during the 2018.06.30 validation
- Since they were no longer in DNS, they are not included in the 2016 total for the 2018 report

ありがとうございました
Thank you

—