Neither Snow Nor Rain Nor MITM…
Real World Email Delivery Security

Zakir Durumeric

University of Michigan

How is your everyday email protected?

Neither Snow Nor Rain Nor MITM...
An Empirical Analysis of Email Delivery Security

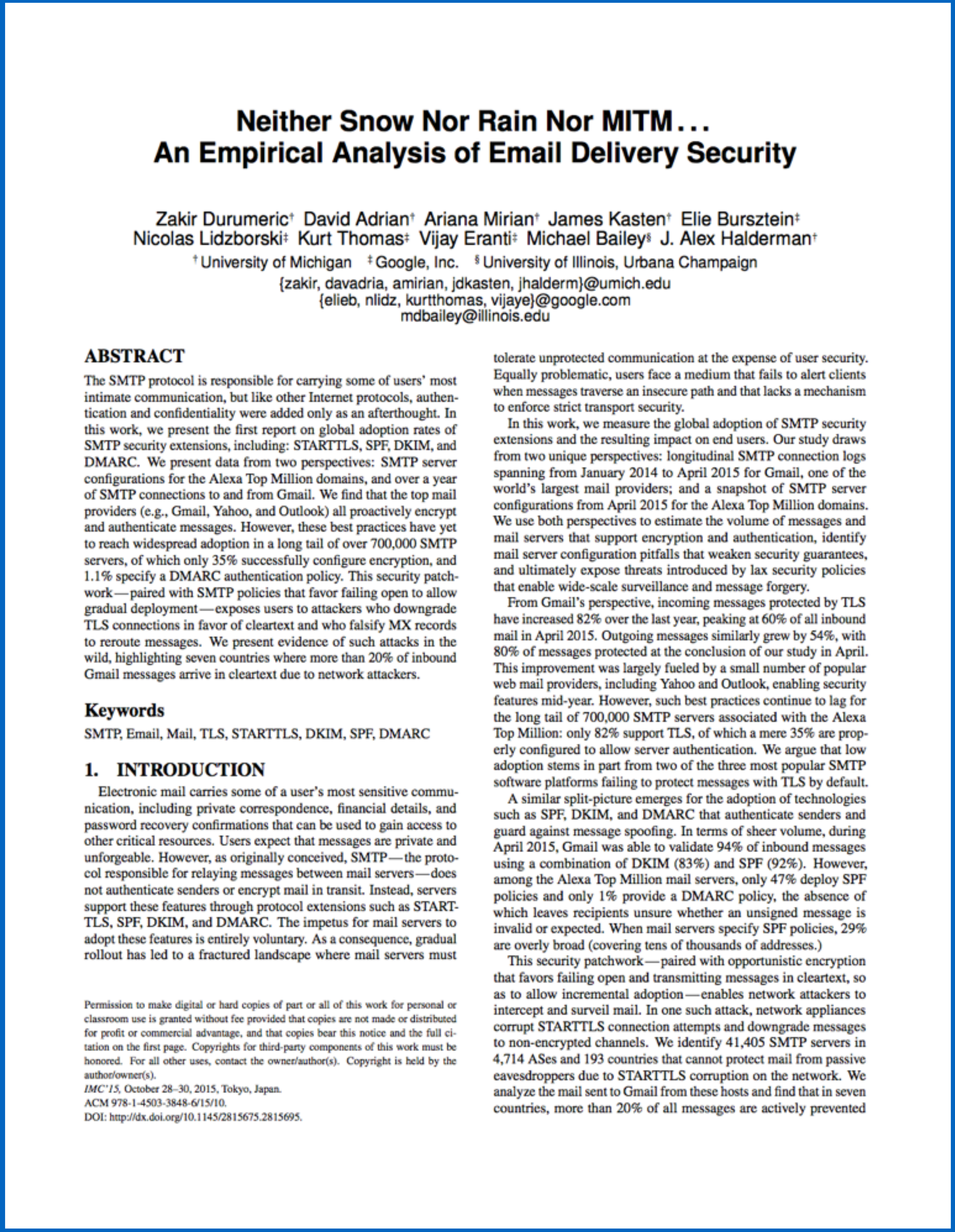
Zakir Durumeric, David Adrian, Ariana Mirian,
James Kasten, Michael Bailey, J. Alex Halderman

University of Michigan, Illinois

Kurt Thomas, Vijay Eranti, Nicholas Lidzborski,
Elie Bursztein

Google

ACM Internet Measurement Conference (IMC’15)



Neither Snow Nor Rain Nor MITM . . .
An Empirical Analysis of Email Delivery Security

Zakir Durumeric[†] David Adrian[†] Ariana Mirian[†] James Kasten[†] Elie Bursztein[‡]
Nicolas Lidzborski[‡] Kurt Thomas[‡] Vijay Eranti[‡] Michael Bailey[§] J. Alex Halderman[†]
[†] University of Michigan [‡] Google, Inc. [§] University of Illinois, Urbana Champaign
{zakir, davadria, amirian, jdkasten, jhalderm}@umich.edu
{elieb, nlidz, kurtthomas, vijaye}@google.com
mdbailey@illinois.edu

ABSTRACT

The SMTP protocol is responsible for carrying some of users’ most intimate communication, but like other Internet protocols, authentication and confidentiality were added only as an afterthought. In this work, we present the first report on global adoption rates of SMTP security extensions, including: STARTTLS, SPF, DKIM, and DMARC. We present data from two perspectives: SMTP server configurations for the Alexa Top Million domains, and over a year of SMTP connections to and from Gmail. We find that the top mail providers (e.g., Gmail, Yahoo, and Outlook) all proactively encrypt and authenticate messages. However, these best practices have yet to reach widespread adoption in a long tail of over 700,000 SMTP servers, of which only 35% successfully configure encryption, and 1.1% specify a DMARC authentication policy. This security patchwork — paired with SMTP policies that favor failing open to allow gradual deployment — exposes users to attackers who downgrade TLS connections in favor of cleartext and who falsify MX records to reroute messages. We present evidence of such attacks in the wild, highlighting seven countries where more than 20% of inbound Gmail messages arrive in cleartext due to network attackers.

Keywords

SMTP, Email, Mail, TLS, STARTTLS, DKIM, SPF, DMARC

1. INTRODUCTION

Electronic mail carries some of a user's most sensitive communication, including private correspondence, financial details, and password recovery confirmations that can be used to gain access to other critical resources. Users expect that messages are private and unforgeable. However, as originally conceived, SMTP — the protocol responsible for relaying messages between mail servers — does not authenticate senders or encrypt mail in transit. Instead, servers support these features through protocol extensions such as STARTTLS, SPF, DKIM, and DMARC. The impetus for mail servers to adopt these features is entirely voluntary. As a consequence, gradual rollout has led to a fractured landscape where mail servers must tolerate unprotected communication at the expense of user security. Equally problematic, users face a medium that fails to alert clients when messages traverse an insecure path and that lacks a mechanism to enforce strict transport security.

In this work, we measure the global adoption of SMTP security extensions and the resulting impact on end users. Our study draws from two unique perspectives: longitudinal SMTP connection logs spanning from January 2014 to April 2015 for Gmail, one of the world’s largest mail providers; and a snapshot of SMTP server configurations from April 2015 for the Alexa Top Million domains. We use both perspectives to estimate the volume of messages and mail servers that support encryption and authentication, identify mail server configuration pitfalls that weaken security guarantees, and ultimately expose threats introduced by lax security policies that enable wide-scale surveillance and message forgery.

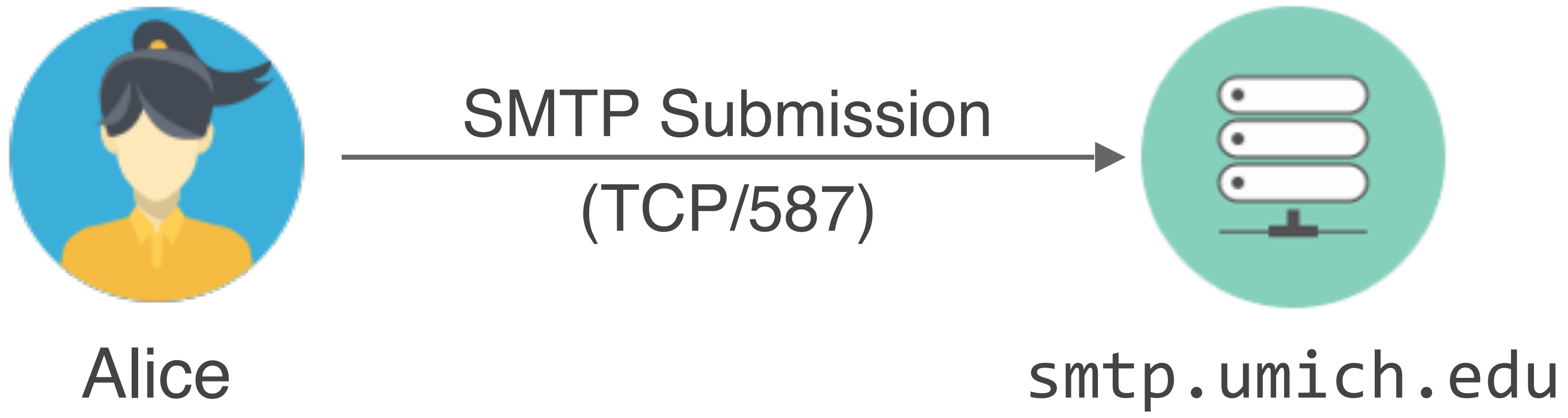
From Gmail’s perspective, incoming messages protected by TLS have increased 82% over the last year, peaking at 60% of all inbound mail in April 2015. Outgoing messages similarly grew by 54%, with 80% of messages protected at the conclusion of our study in April. This improvement was largely fueled by a small number of popular web mail providers, including Yahoo and Outlook, enabling security features mid-year. However, such best practices continue to lag for the long tail of 700,000 SMTP servers associated with the Alexa Top Million: only 82% support TLS, of which a mere 35% are properly configured to allow server authentication. We argue that low adoption stems in part from two of the three most popular SMTP software platforms failing to protect messages with TLS by default.

A similar split-picture emerges for the adoption of technologies such as SPF, DKIM, and DMARC that authenticate senders and guard against message spoofing. In terms of sheer volume, during April 2015, Gmail was able to validate 94% of inbound messages using a combination of DKIM (83%) and SPF (92%). However, among the Alexa Top Million mail servers, only 47% deploy SPF policies and only 1% provide a DMARC policy, the absence of which leaves recipients unsure whether an unsigned message is invalid or expected. When mail servers specify SPF policies, 29% are overly broad (covering tens of thousands of addresses.)

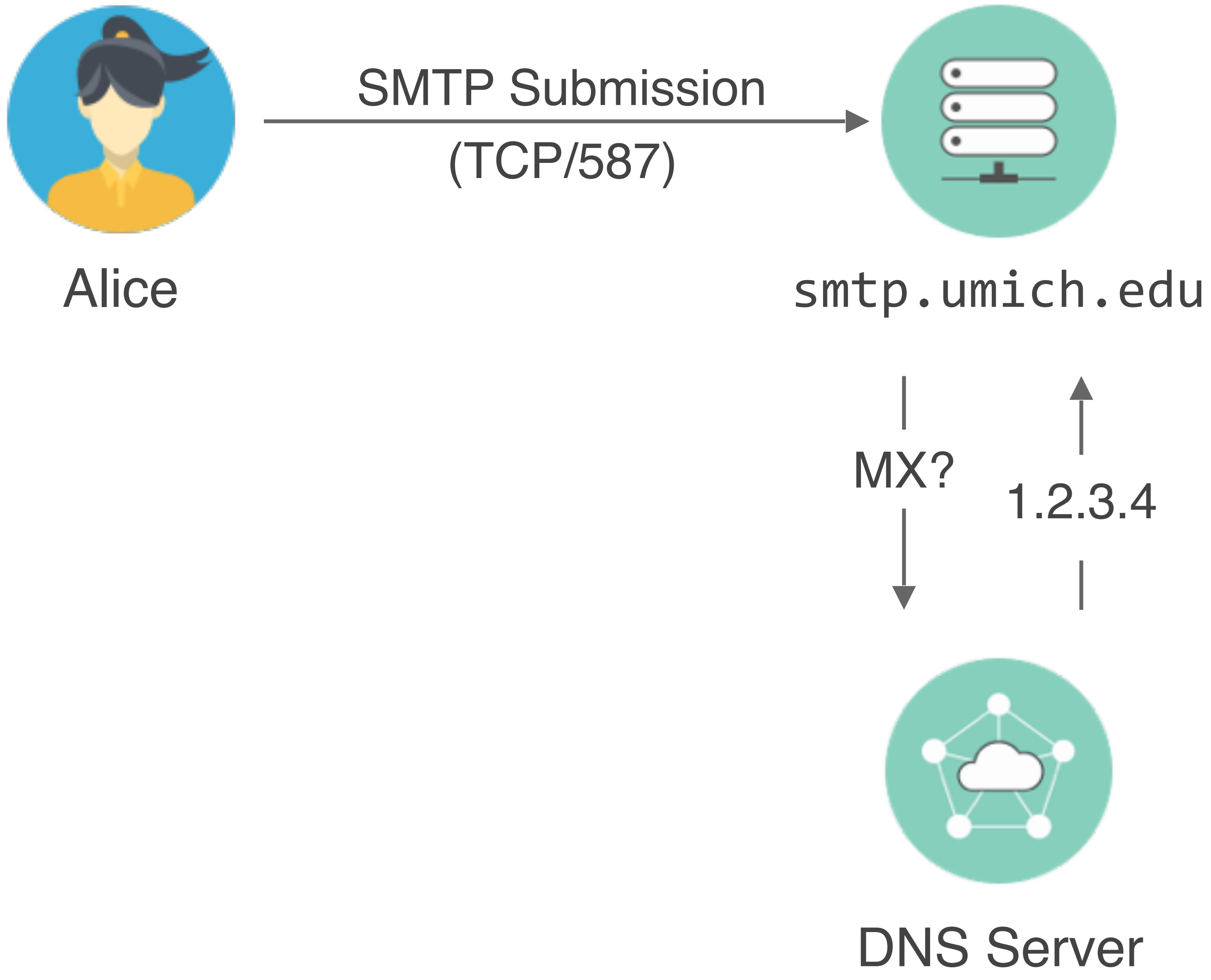
This security patchwork — paired with opportunistic encryption that favors failing open and transmitting messages in cleartext, so as to allow incremental adoption — enables network attackers to intercept and surveil mail. In one such attack, network appliances corrupt STARTTLS connection attempts and downgrade messages to non-encrypted channels. We identify 41,405 SMTP servers in 4,714 ASes and 193 countries that cannot protect mail from passive eavesdroppers due to STARTTLS corruption on the network. We analyze the mail sent to Gmail from these hosts and find that in seven countries, more than 20% of all messages are actively prevented

IMC'15, October 28–30, 2015, Tokyo, Japan.
ACM 978-1-4503-3848-6/15/10.
DOI: http://dx.doi.org/10.1145/2815675.2815695.

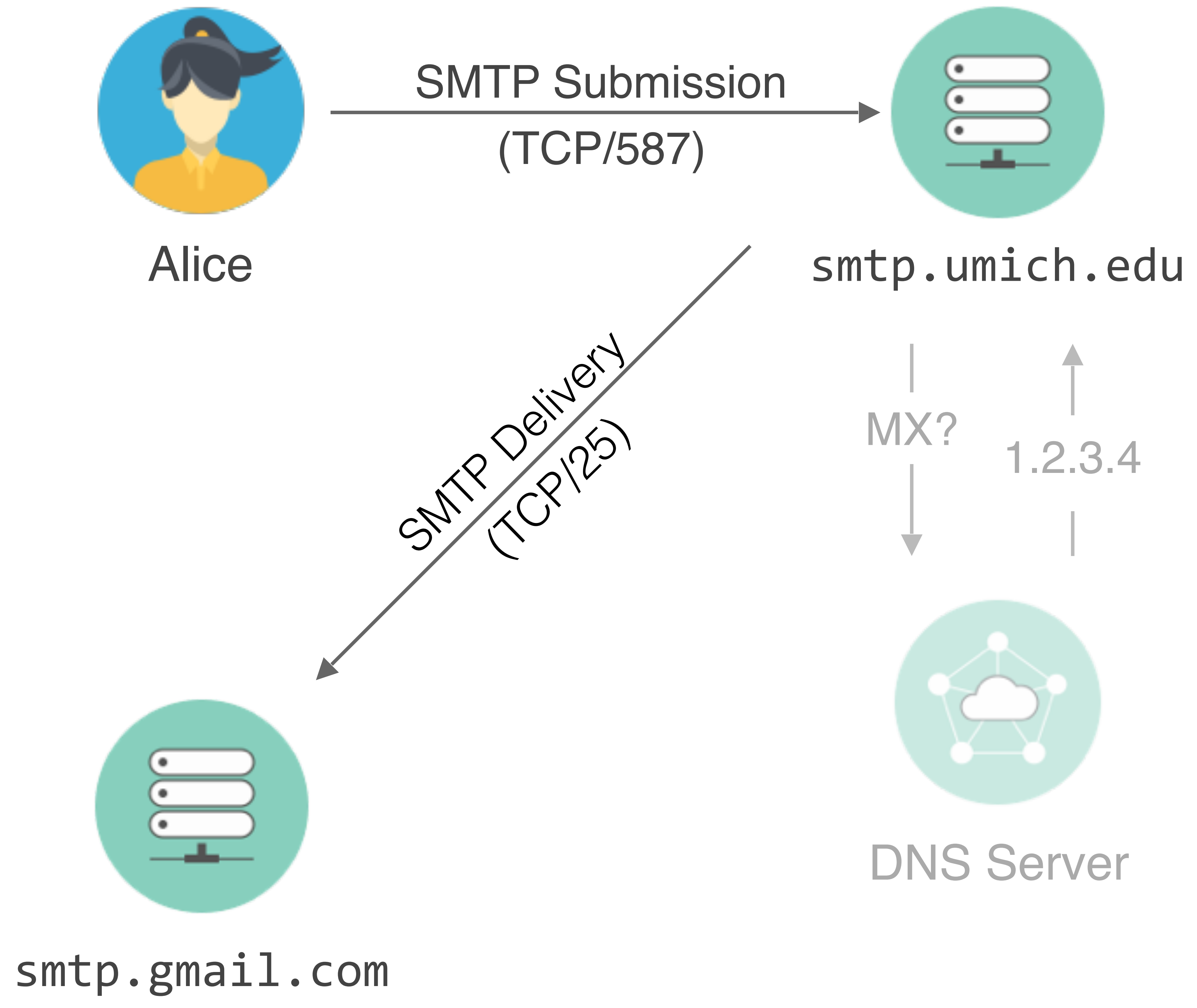
Email Delivery



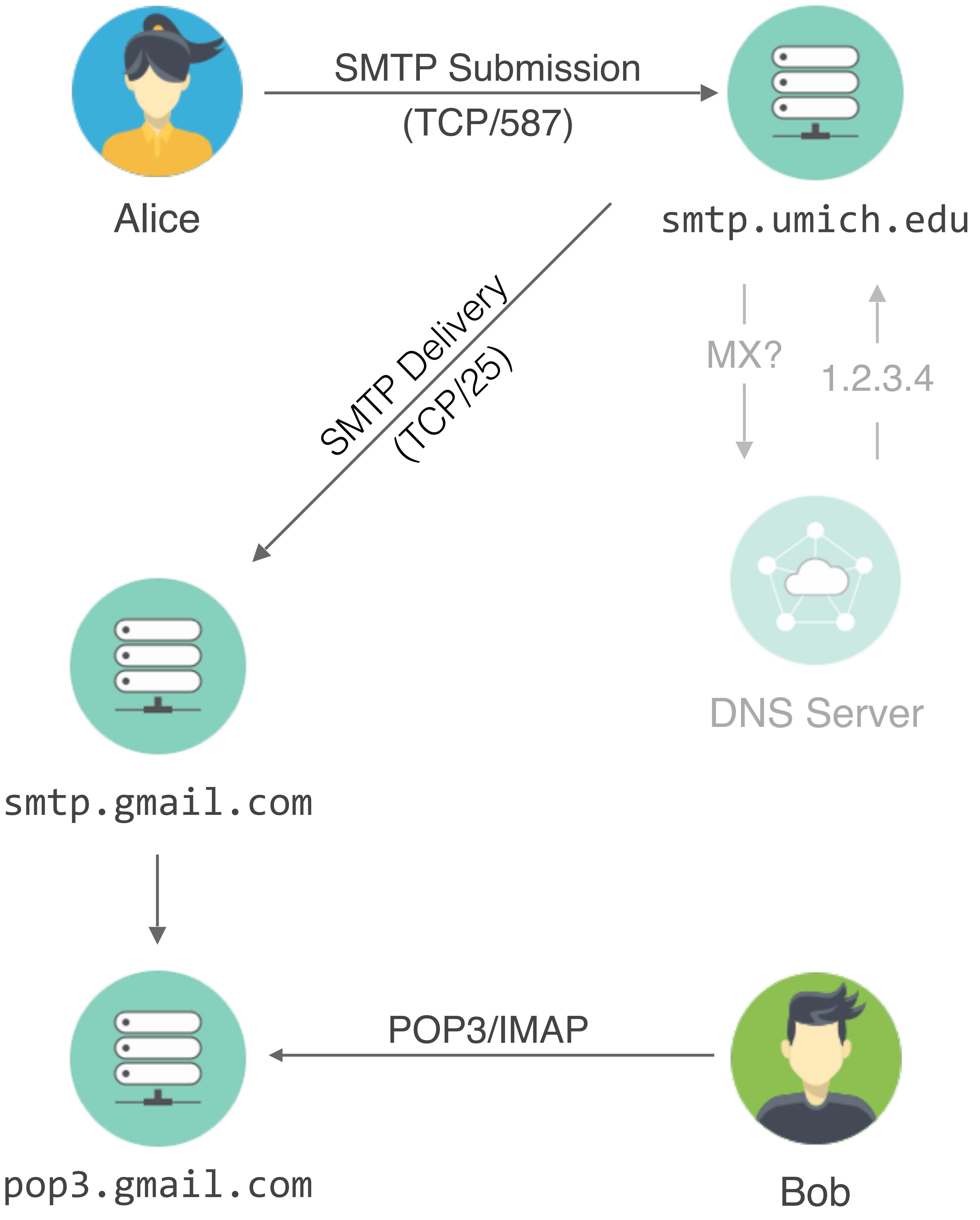
Email Delivery



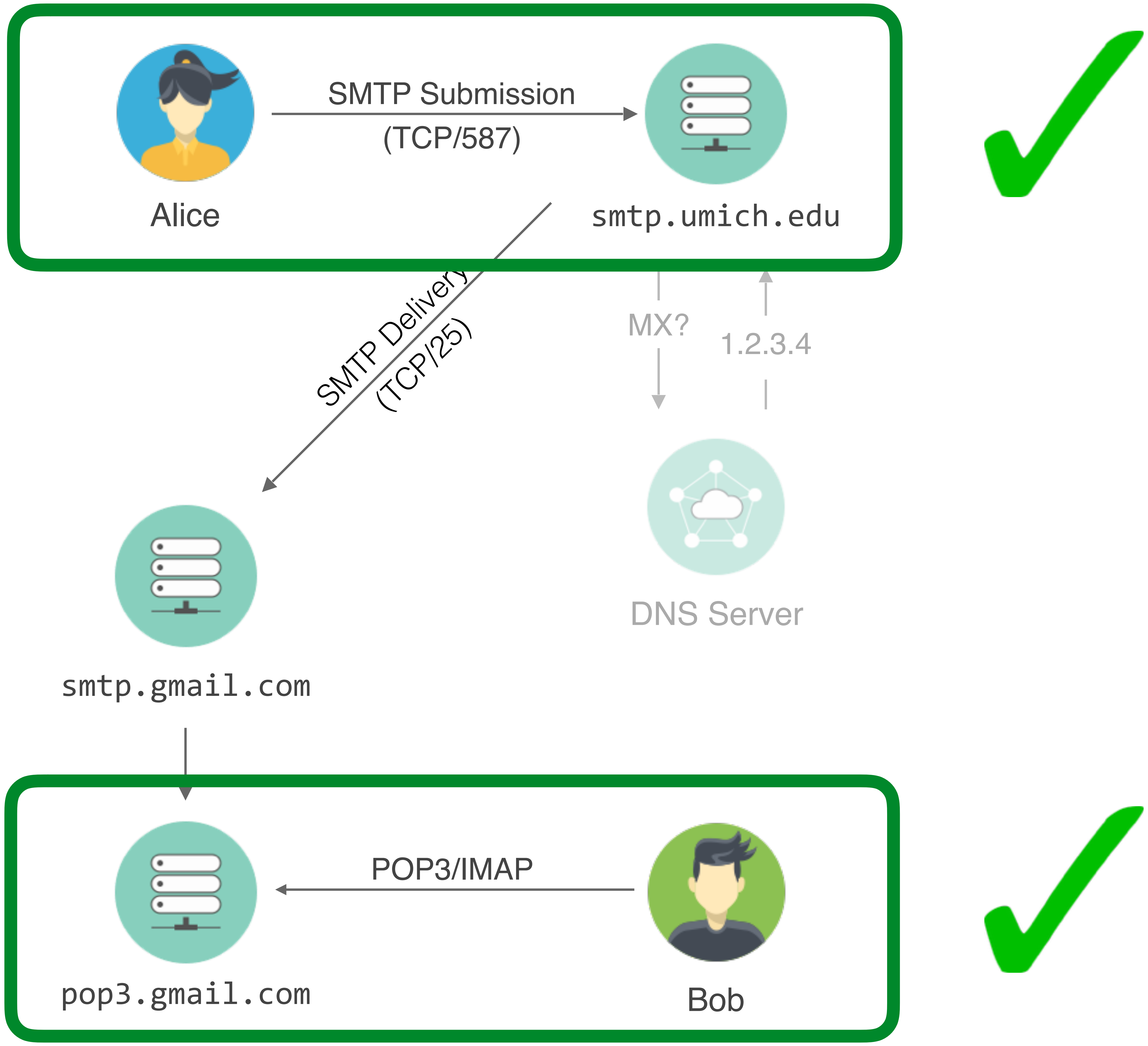
Email Delivery



Email Delivery



Email Delivery



Alice

SMTP Submission
(TCP/587)

smtp.umich.edu

SMTP Delivery
(TCP/25)

MX?

1.2.3.4

DNS Server

smtp.gmail.com

pop3.gmail.com

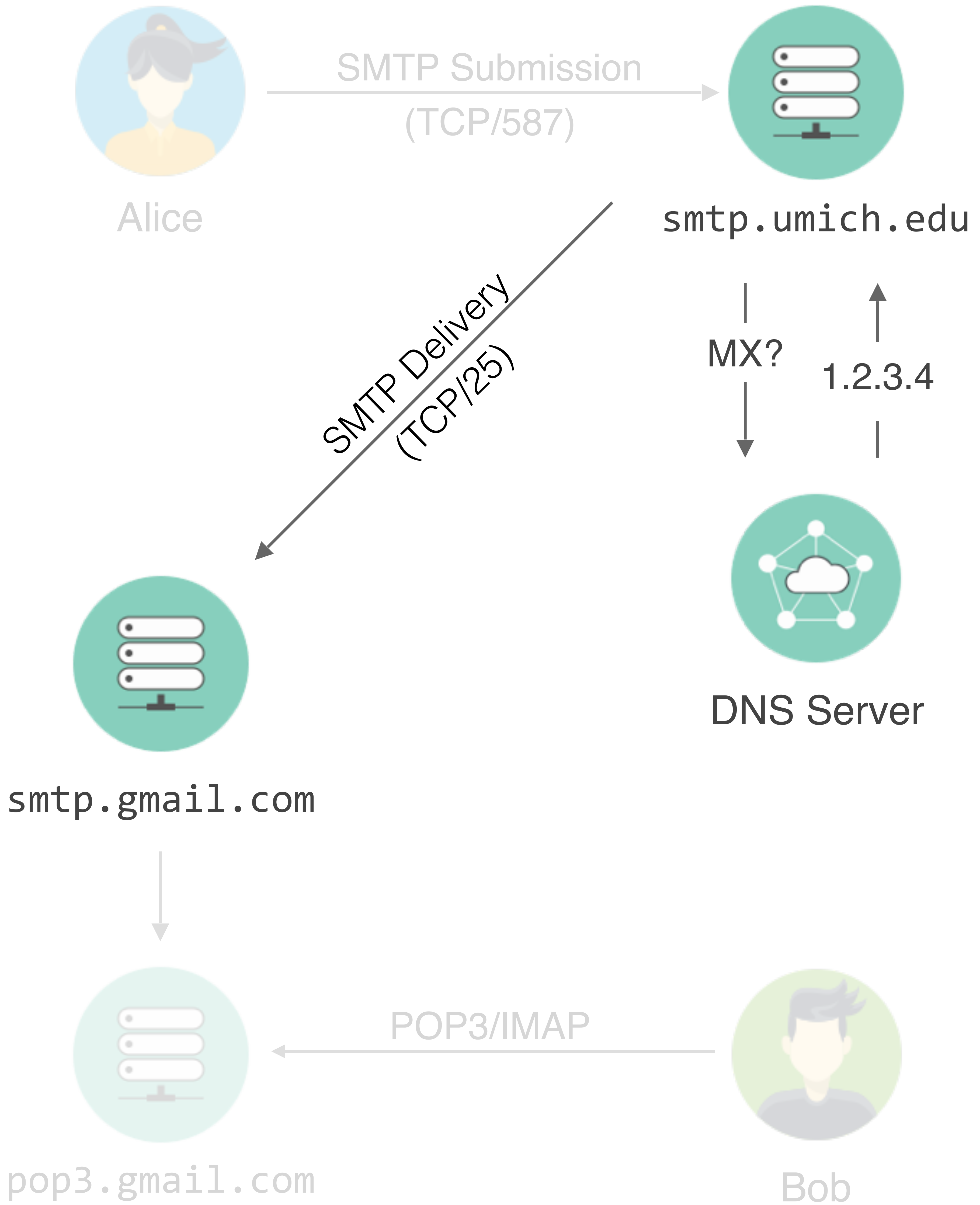
POP3/IMAP

Bob

✓

✓

Email Delivery



SMTP has no built-in security

We've added SMTP extensions to:

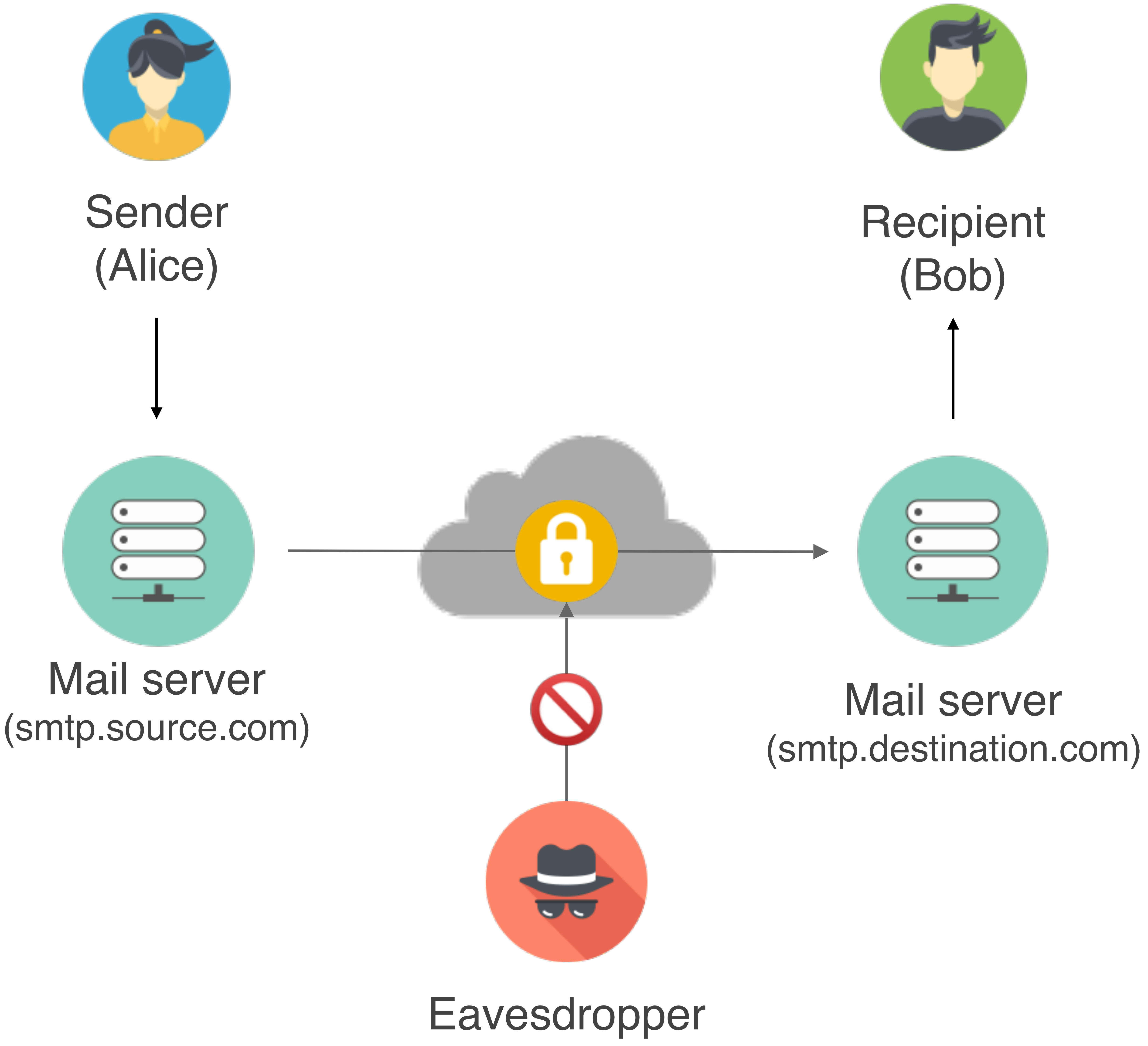
1. Encrypt email in transit
2. Authenticate email on receipt

However… deployment is voluntary and invisible to end users

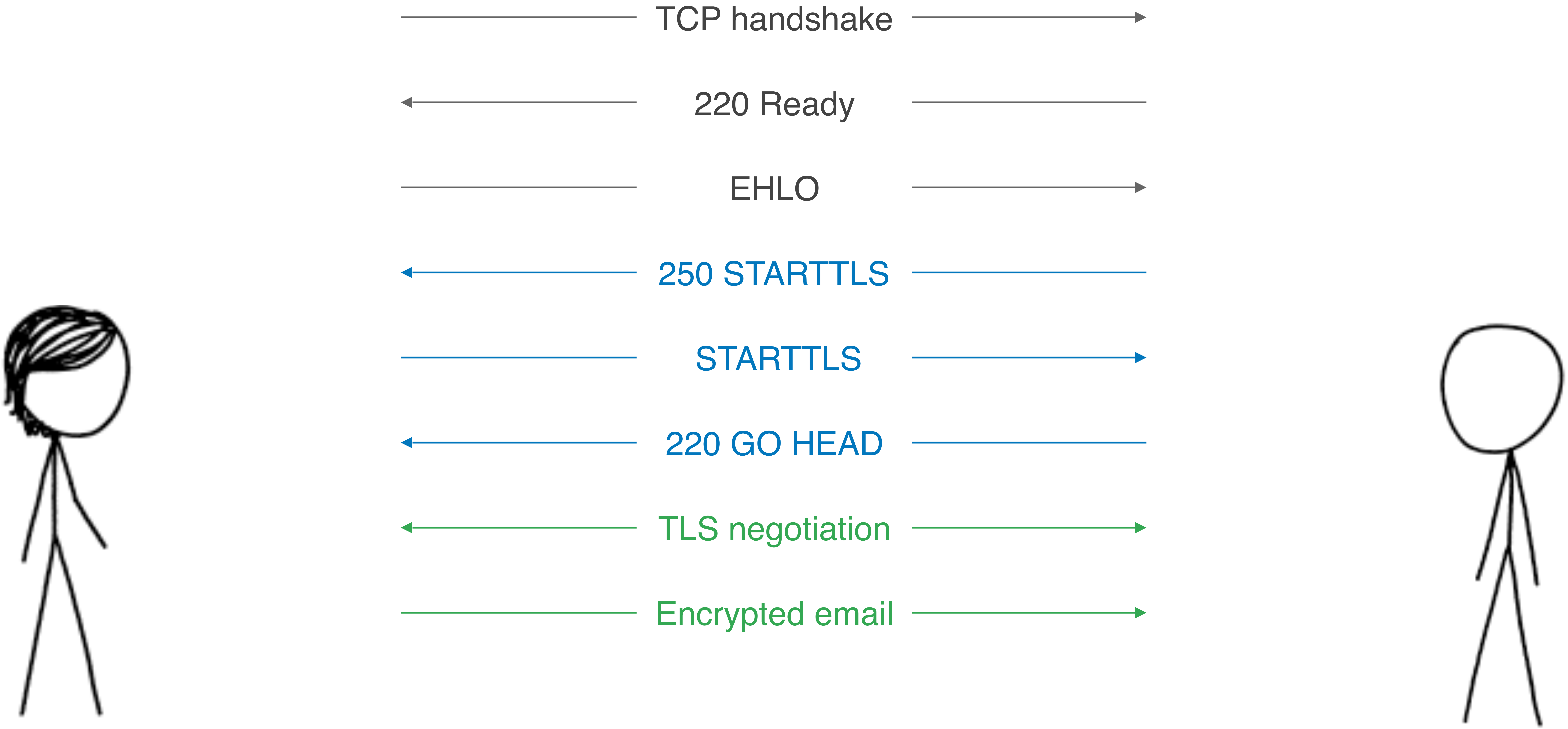
STARTTLS: TLS for SMTP

Allow TLS session to be started during an SMTP connection

Mail is transferred over the encrypted session



STARTTLS Protocol



TCP handshake

220 Ready

EHLO

250 STARTTLS

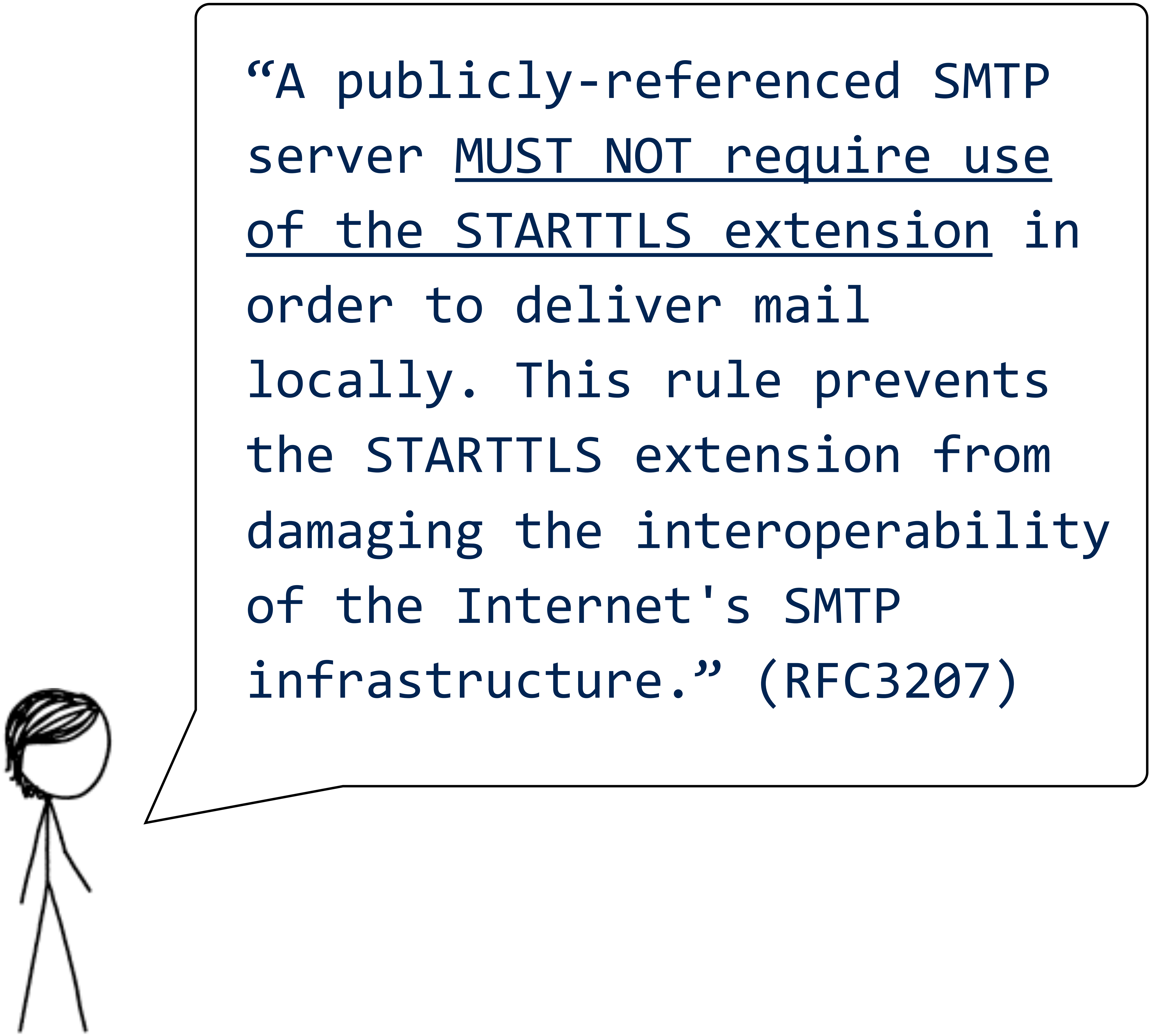
STARTTLS

220 GO HEAD

TLS negotiation

Encrypted email

Opportunistic Encryption Only



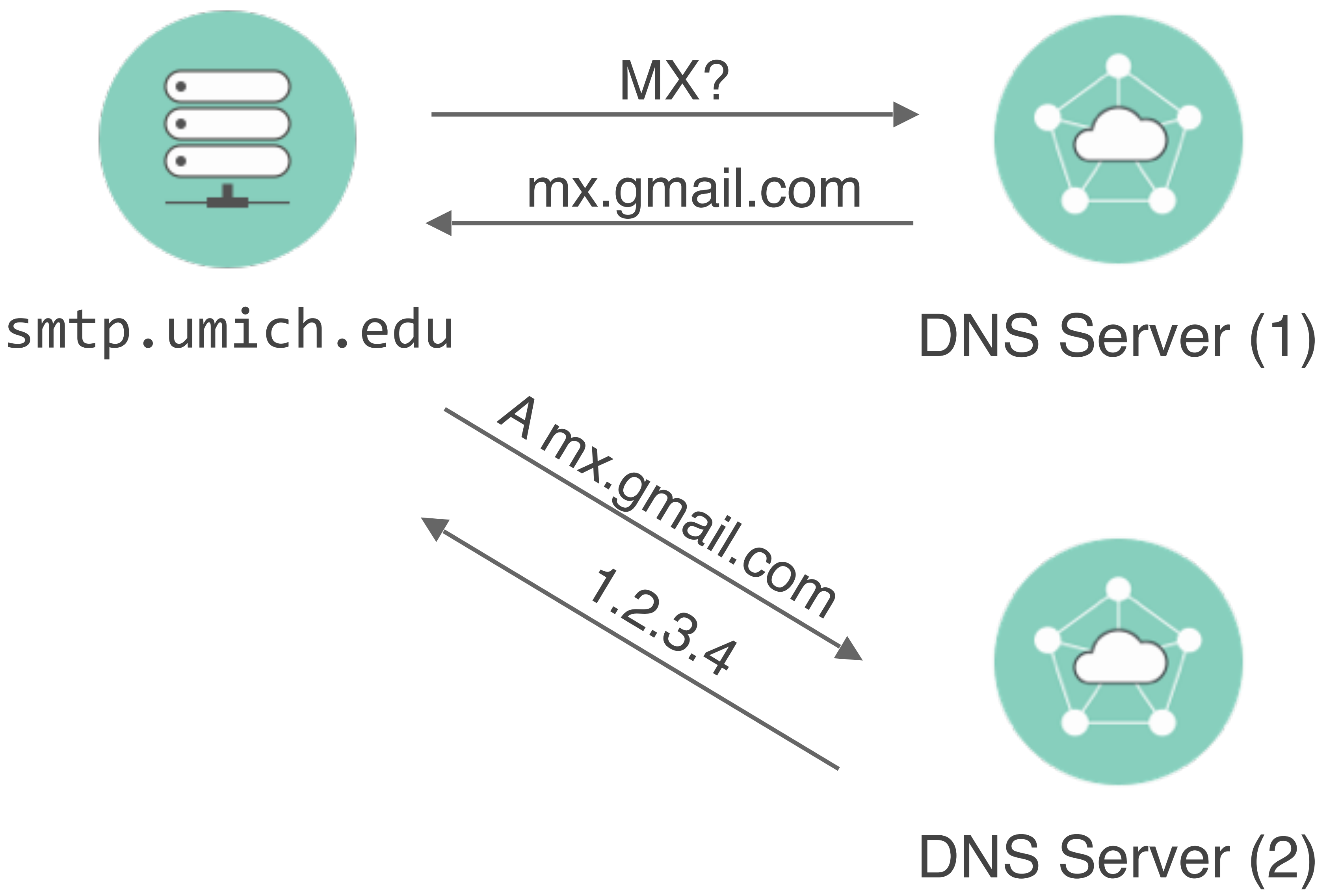
> “A publicly-referenced SMTP server MUST NOT require use of the STARTTLS extension in order to deliver mail locally. This rule prevents the STARTTLS extension from damaging the interoperability of the Internet's SMTP infrastructure.” (RFC3207)

Unlike HTTPS, STARTTLS is used opportunistically

Senders do not validate destination servers — the alternative is cleartext

Many servers do not support STARTTLS

What name to validate?



Two Step DNS Resolution

Unlike HTTPS, unclear what name should go on the certificate

MX Server (e.g., smtp.gmail.com)

- No real security added
- MITM returns bad MX record

Domain (e.g., gmail.com)

- No clear solution for large cloud providers

What name to validate?

Cloud Provider	% Top 1Mil
Gmail	16%
GoDaddy	5%
Yandex	2%
QQ	1%
OVH	1%

Unlike HTTPS, unclear what name should go on the certificate

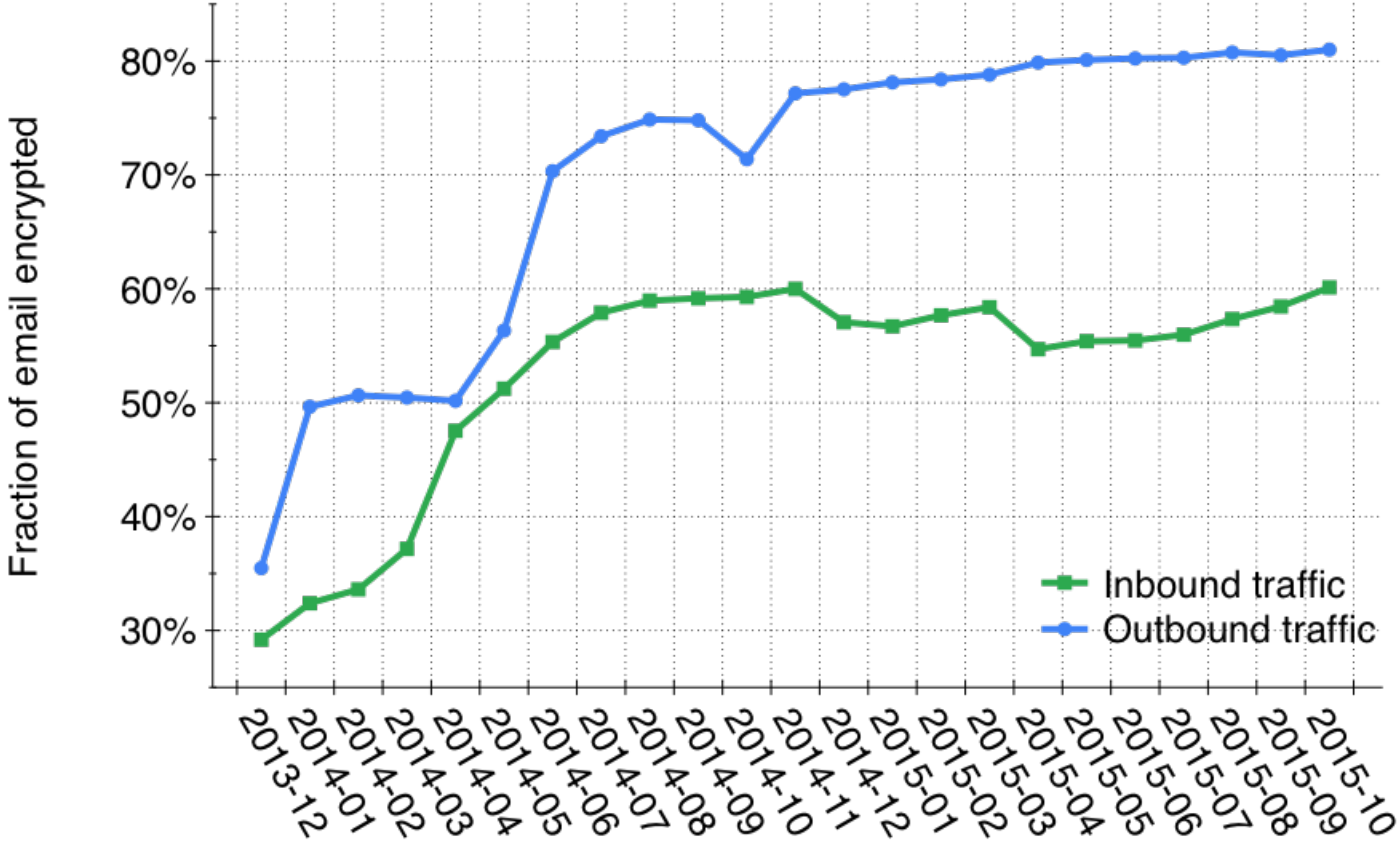
MX Server (e.g., smtp.gmail.com)

- No real security added
- MITM returns bad MX record

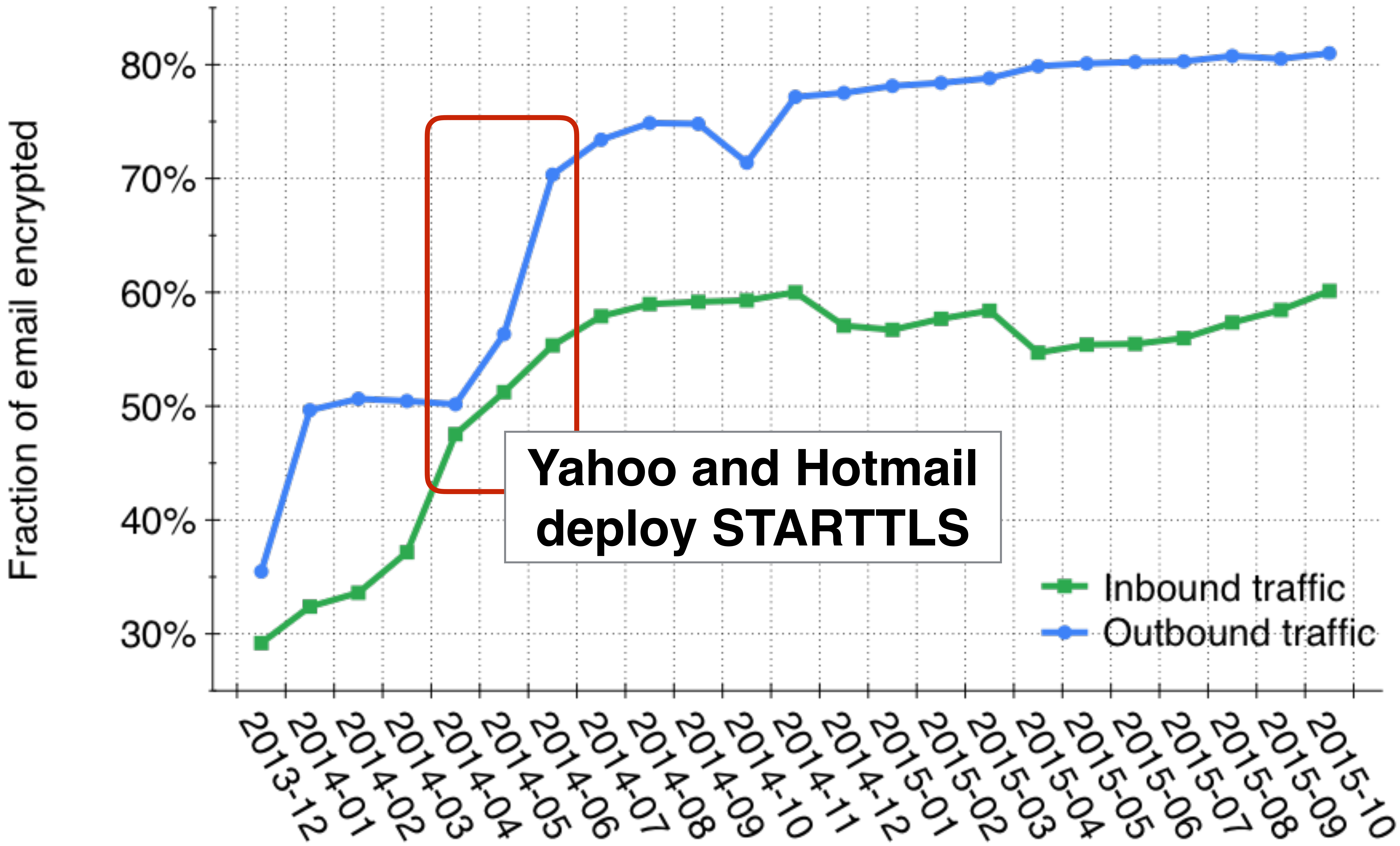
Domain (e.g., gmail.com)

- No clear solution for large cloud providers

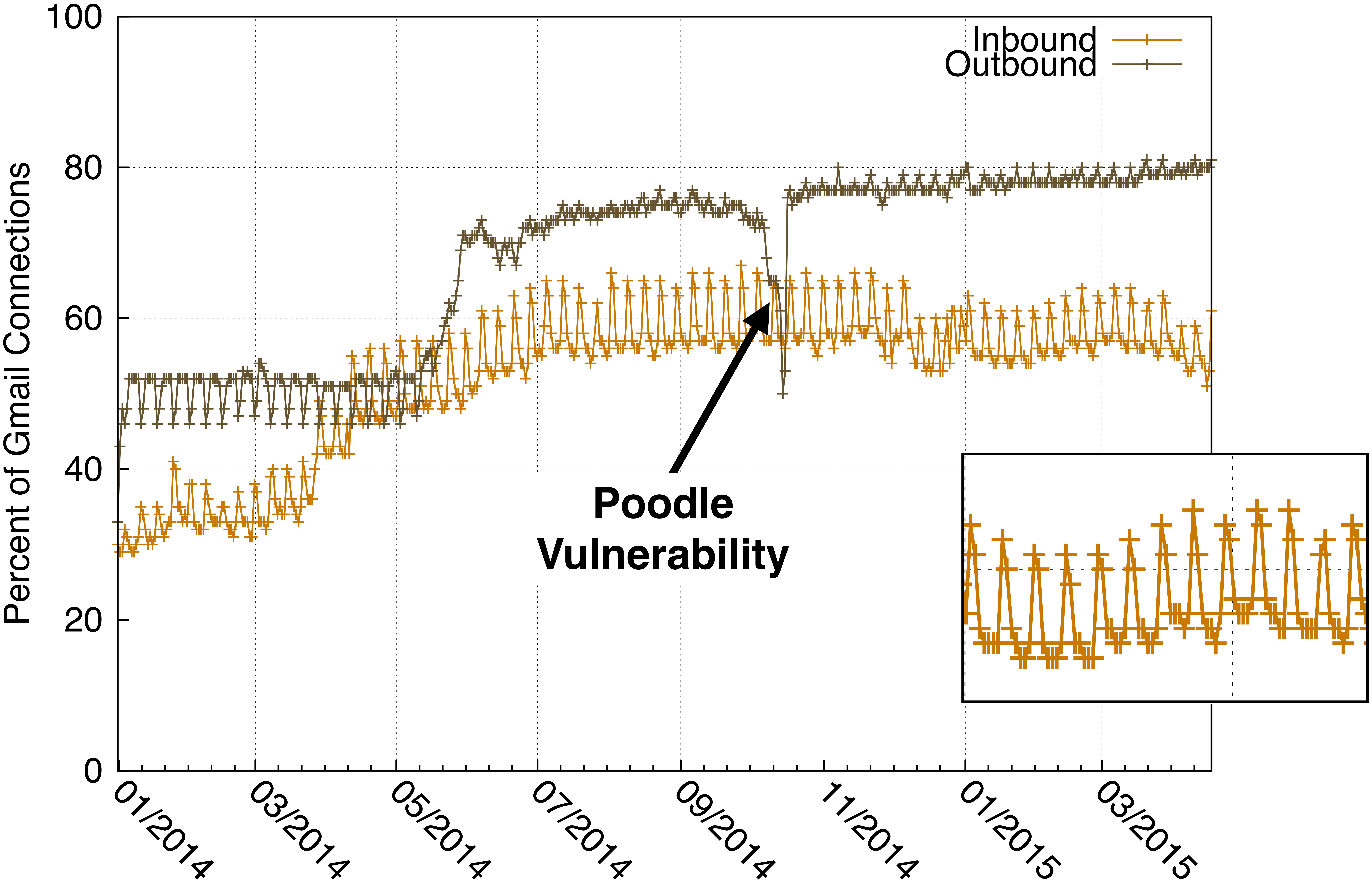
STARTTLS Usage as seen by Gmail



STARTTLS Usage as seen by Gmail



Yahoo and Hotmail deploy STARTTLS


Percent of Gmail Connections
100
80
60
40
20
0
01/2014
03/2014
05/2014
07/2014
09/2014
11/2014
01/2015
03/2015
Inbound
Outbound
Poodle
Vulnerability

Cipher Selection

Provider	Incoming Key Exchange	Incoming Cipher	Certificate Name	Outgoing Key exchange	Outgoing Cipher
Gmail	ECDHE	AES128-GCM	match	ECDHE	AES128-GCM
Yahoo	ECDHE	AES128-GCM	match	ECDHE	**RC4-128**
Microsoft	ECDHE	AES256-CBC	match	ECDHE	AES256
Apple iCloud	ECDHE	AES128-GCM	match	**DHE**	AES128-GCM
Facebook mail	**RSA**	AES128-CBC	**mismatch**	ECDHE	AES128-CBC
Comcast	**RSA**	**RC4-128**	match	**DHE**	AES128-CBC
AT&T	ECDHE	AES128-GCM	match	ECDHE	**RC4-128**

Long Tail of Mail Operators

These numbers are dominated by a few large providers

Of the Alexa Top 1M Domains with Mail Servers:

- 81.8% support STARTTLS
- 34% have certificates that match MX server
- 0.6% have certificates that match domain

Long Tail of Mail Operators

These numbers are dominated by a few large providers

Of the Alexa Top 1M Domains with Mail Servers:

- 81.8% support STARTTLS
- 34% have certificates that match MX server
- 0.6% have certificates that match domain

Needed to verify valid destination!

Common Mail Software

Software	Top Million Market Share	Public IPv4 Market Share	Default Incoming	Default Outgoing
Exim	34%	24%	❌	✓
Postfix	18%	21%	✓	❌
qmail	6%	1%	❌	❌
Sendmail	5%	4%	❌	✓
MS Exchange	4%	12%	✓	✓
Other/Unknown	33%	38%	?	?

Common Mail Software

Software	Top Million Market Share	Public IPv4 Market Share	Default Incoming	Default Outgoing
Exim	34%	24%	❌	✓
Postfix	18%	21%	✓	❌
qmail	6%	1%	❌	❌
Sendmail	5%	4%	❌	✓
MS Exchange	4%	12%	✓	✓
Other/Unknown	33%	38%	?	?

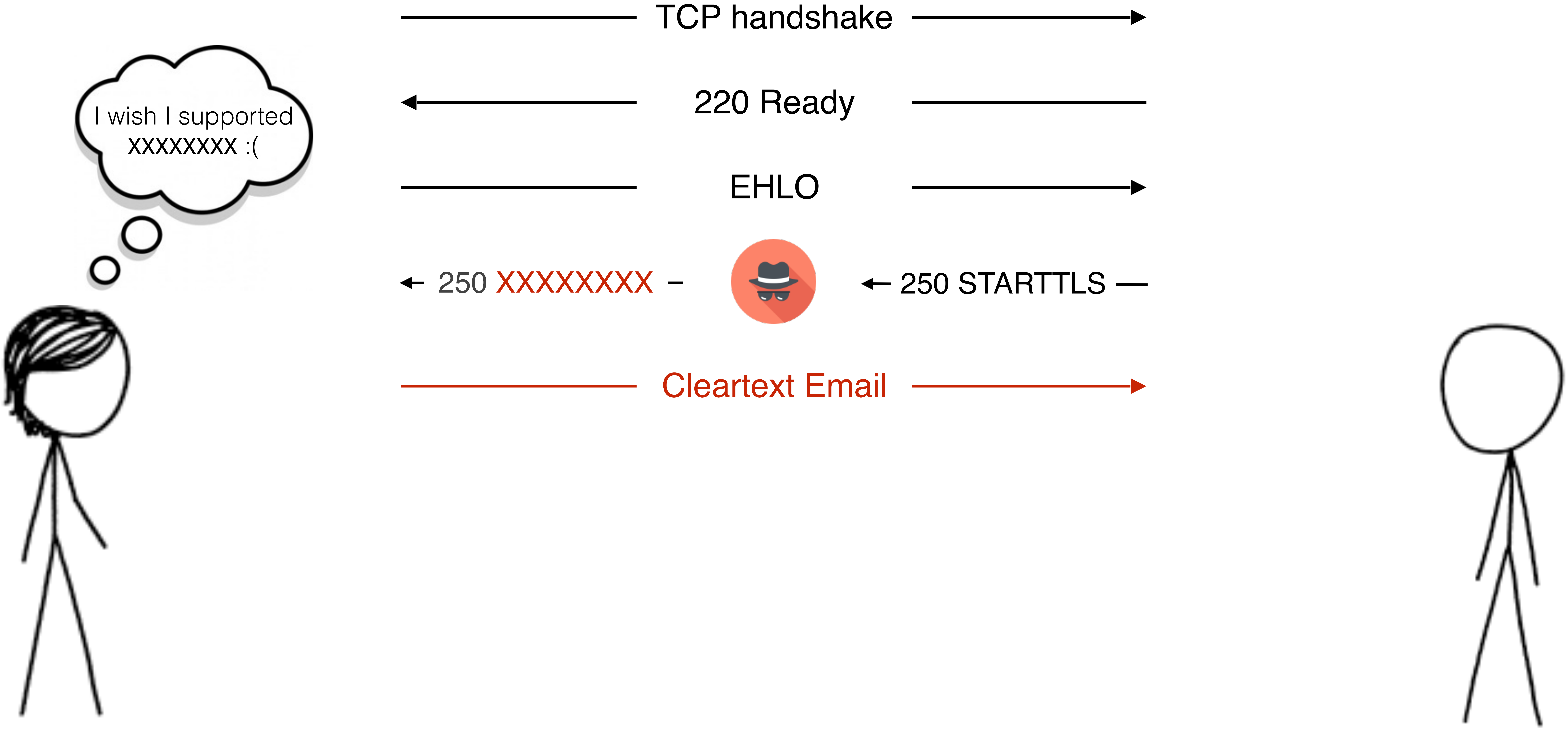
Common Mail Software

Software	Top Million Market Share	Public IPv4 Market Share	Default Incoming	Default Outgoing
Exim	34%	24%	✗	✓
Postfix	18%	21%	✓	✗
qmail	6%	1%	✗	✗
Sendmail	5%	4%	✗	✓
MS Exchange	4%	12%	✓	✓
Other/Unknown	33%	38%	?	?

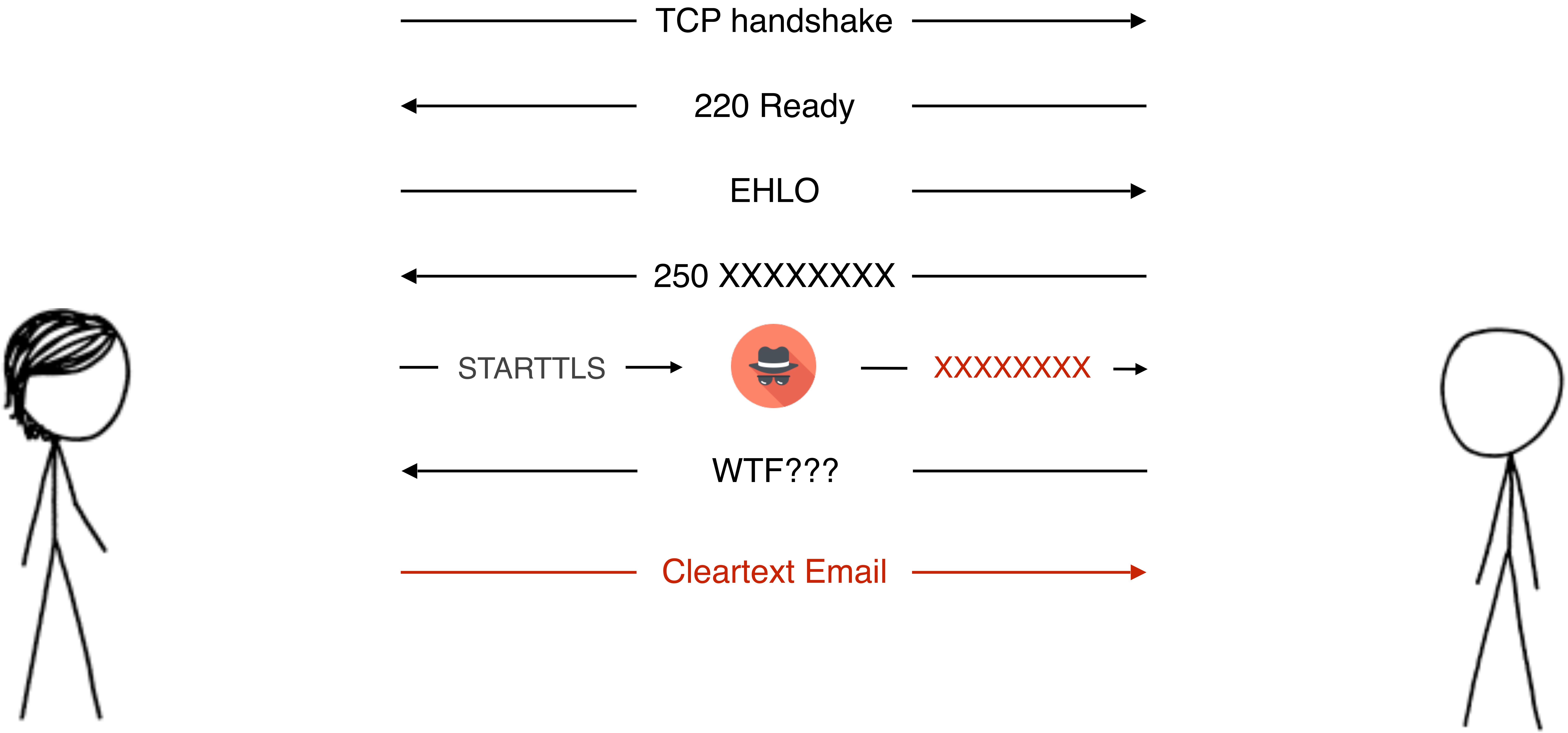
StartTLS protects against passive eavesdropping. Nothing else.

What’s the simplest way to eavesdrop on servers that use StartTLS?

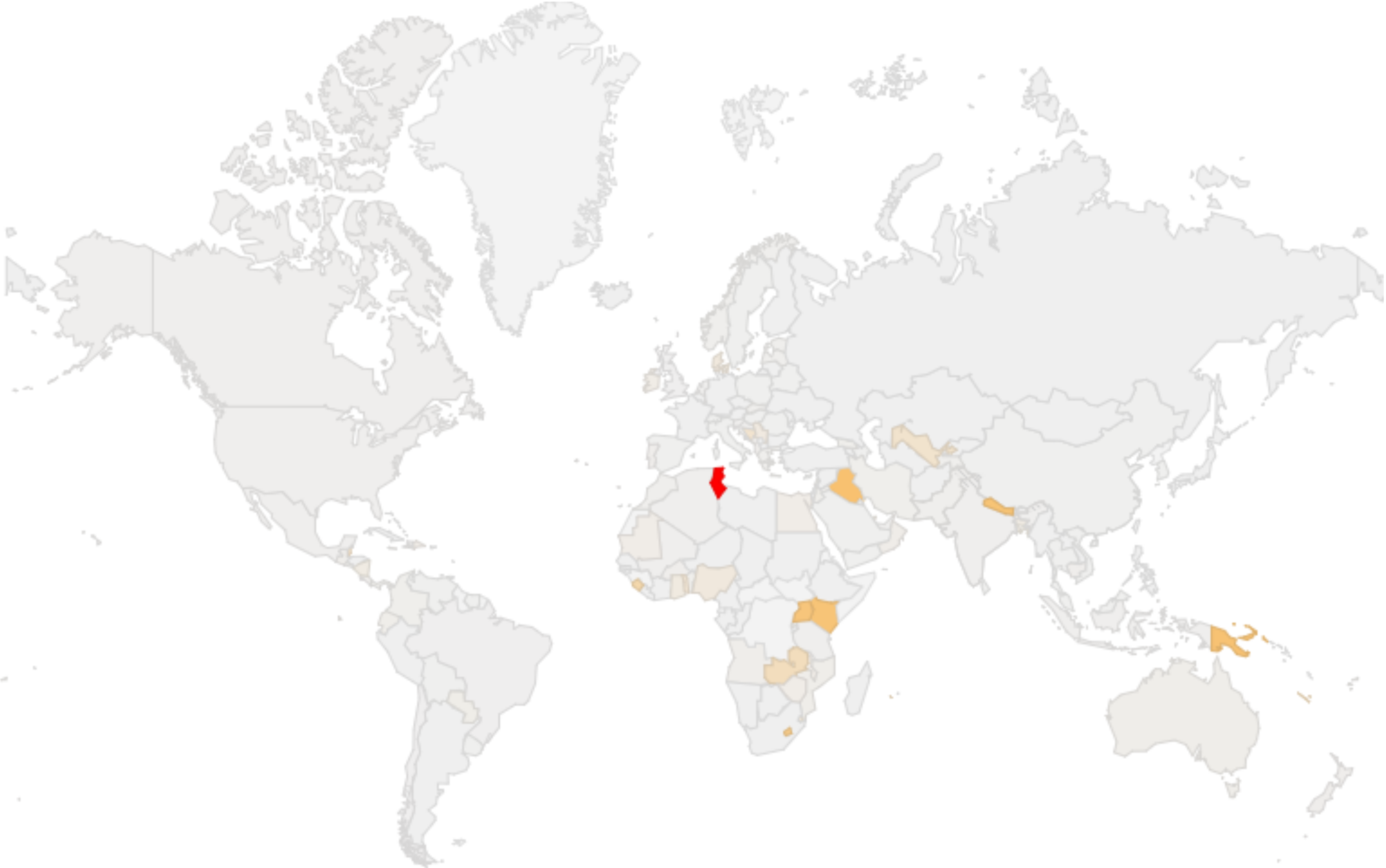
STARTTLS Stripping (1)



STARTTLS Stripping (2)



STARTTLS Stripping in the Wild



Country	
Tunisia	96.1%
Iraq	25.6%
Papua New Guinea	25.0%
Nepal	24.3%
Kenya	24.1%
Uganda	23.3%
Lesotho	20.3%
Sierra Leone	13.4%
New Caledonia	10.1%
Zambia	10.0%

STARTTLS Stripping in the Wild

Country	
Tunisia	96.1%
Iraq	25.6%
Papua New Guinea	25.0%
Nepal	24.3%
Kenya	24.1%
Uganda	23.3%
Lesotho	20.3%
Sierra Leone	13.4%
New Caledonia	10.1%
Zambia	10.0%

Country	
Reunion	9.3%
Belize	7.7%
Uzbekistan	6.9%
Bosnia and Herzegovina	6.5%
Togo	5.5%
Barbados	5.3%
Swaziland	4.6%
Denmark	3.7%
Nigeria	3.6%
Serbia	3.1%

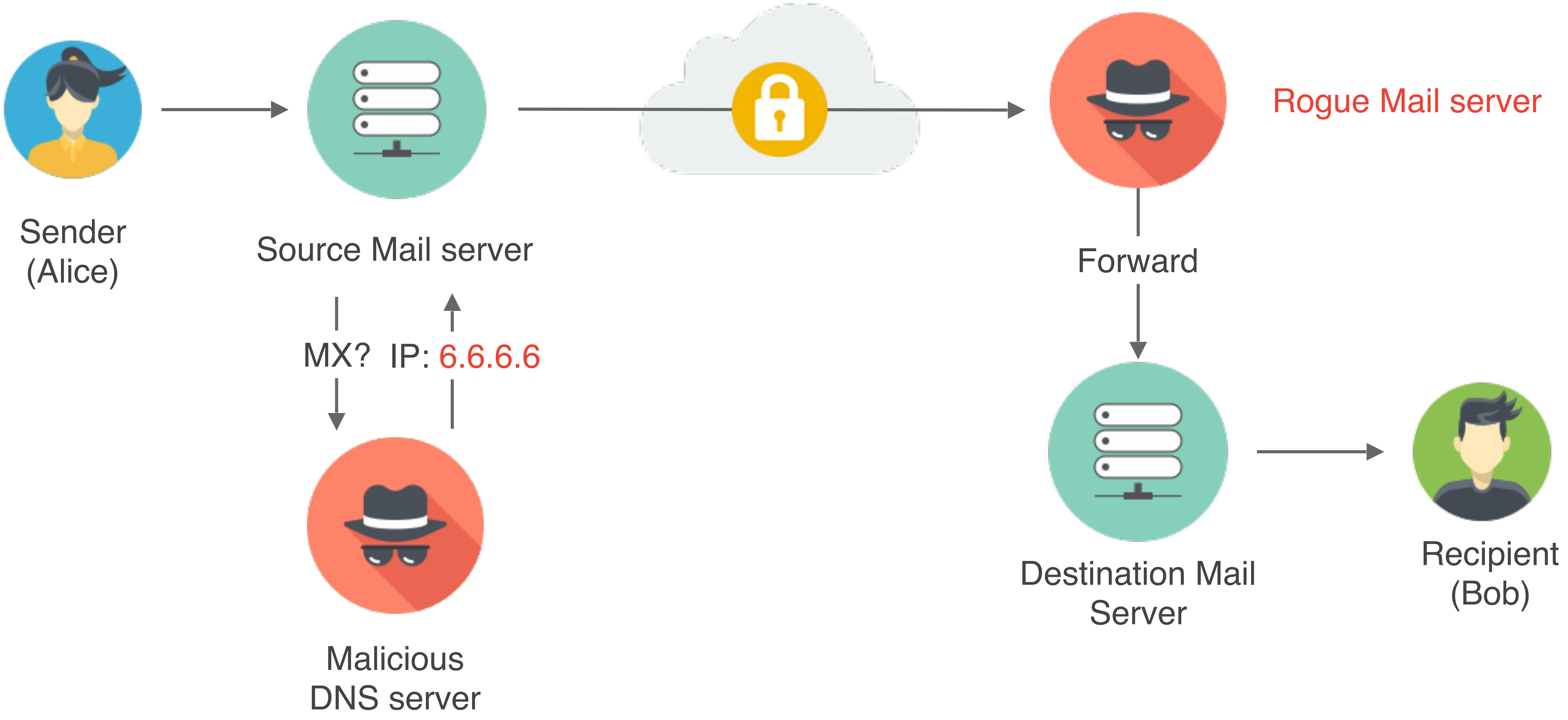
Not Necessarily Malicious…

Organization Type	
Corporation	43%
ISP	18%
Financial Institution	14%
Academic Institution	8%
Healthcare Provider	3%
Unknown	3%
Airport	2%
Hosting Provider	2%
NGO	1%

Cisco advertises this feature to prevent attacks and catch spam

Unclear if operators know they’re putting users at risk

Lying DNS Servers



DNS Spoofing Seen by Gmail

Country	
Slovakia	0.08%
Romania	0.04%
Bulgaria	0.02%
India	0.01%
Israel	0.01%
Poland	0.01%
Switzerland	0.01%
Ukraine	0.01%
Others	10.1%

Authenticating Email

Authenticating Email

Sender Policy Framework (SPF)

Sender publishes list of IPs authorized to send mail

DomainKeys Identified Mail (DKIM)

Sender signs messages with cryptographic key

Domain Message Authentication, Reporting and Conformance (DMARC)

Sender publishes policy in DNS that specifies what to do if DKIM or SPF validation fails

Sender Policy Framework (SPF)

1. Sender publishes a DNS record that specifies what servers can send mail for the domain:

```
    _spf.example.com.   3599  IN TXT   "v=spf1 ip4:64.18.0.0/20 ~all"
```

2. Recipient looks up sender’s SPF policy and and checks if the message was sent from an allowed host

Domain Keys Identified Mail

1. Sender publishes a cryptographic public key in DNS record

```
20120113._domainkey.gmail.com.    300 IN TXT  "k=rsa\; p=MIIBIjAN…AQAB”
```

2. Sender attaches cryptographic signature in a message’s headers

```
DKIM-Signature:
   v=1;
   a=rsa-sha256;
   c=relaxed/relaxed;
   d=gmail.com;
   s=20120113;
   h=from:date:...:subject:to;
   bh=RjhXzraob5/q4159GO00YE=;
   b=YZmpde8KxvpfX…anUdYxVgc
```

3. Recipient looks up key and checks a message’s signature

Domain Keys Identified Mail

1. Sender publishes a cryptographic public key in DNS record

```
20120113._domainkey.gmail.com.    300 IN TXT   "k=rsa\; p=MIIBIjAN…AQAB”
```

2. Sender att a message’s headers

Impossible to know if a domain uses DKIM a priori.

```
DKIM-Sign
    v=1;
    a=rsa-
    c=relaxed/relaxed;
    d=gmail.com;
    s=20120113;
    h=from:date:...:subject:to;
    bh=RjhXzraob5/q4159GO00YE=;
    b=YZmpde8KxvpfX…anUdYxVgc
```

3. Recipient looks up key and checks a message’s signature

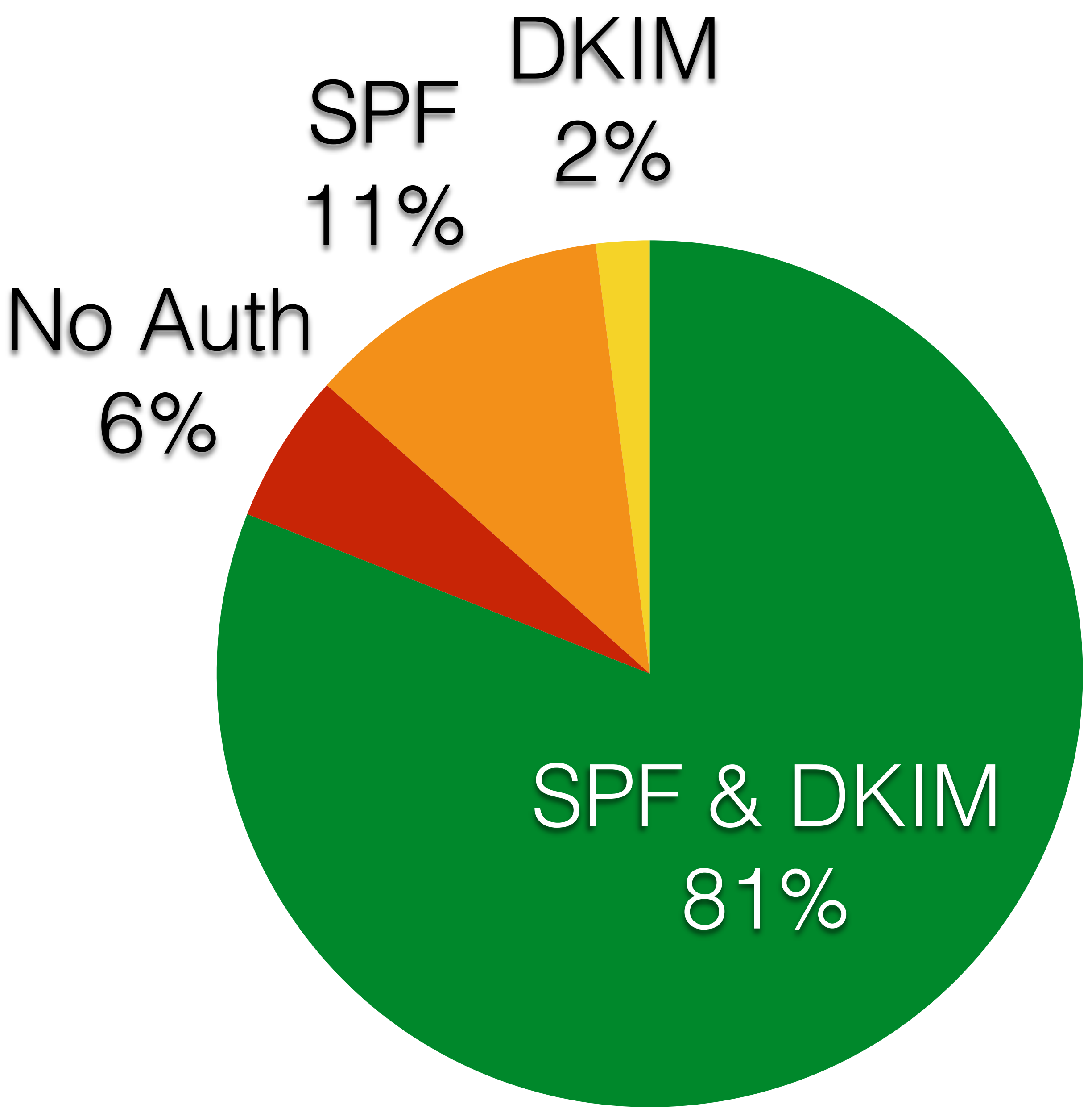
Domain Message Authentication, Reporting and Conformance (DMARC)

1. Sender publishes a mail policy in a DNS record:

```
    _dmarc.yahoo.com.   1800  IN TXT   “v=DMARC1;
                                        p=reject;
                                        pct=100;
                                        rua=mailto:dmarc_y_rua@yahoo.com;"
```

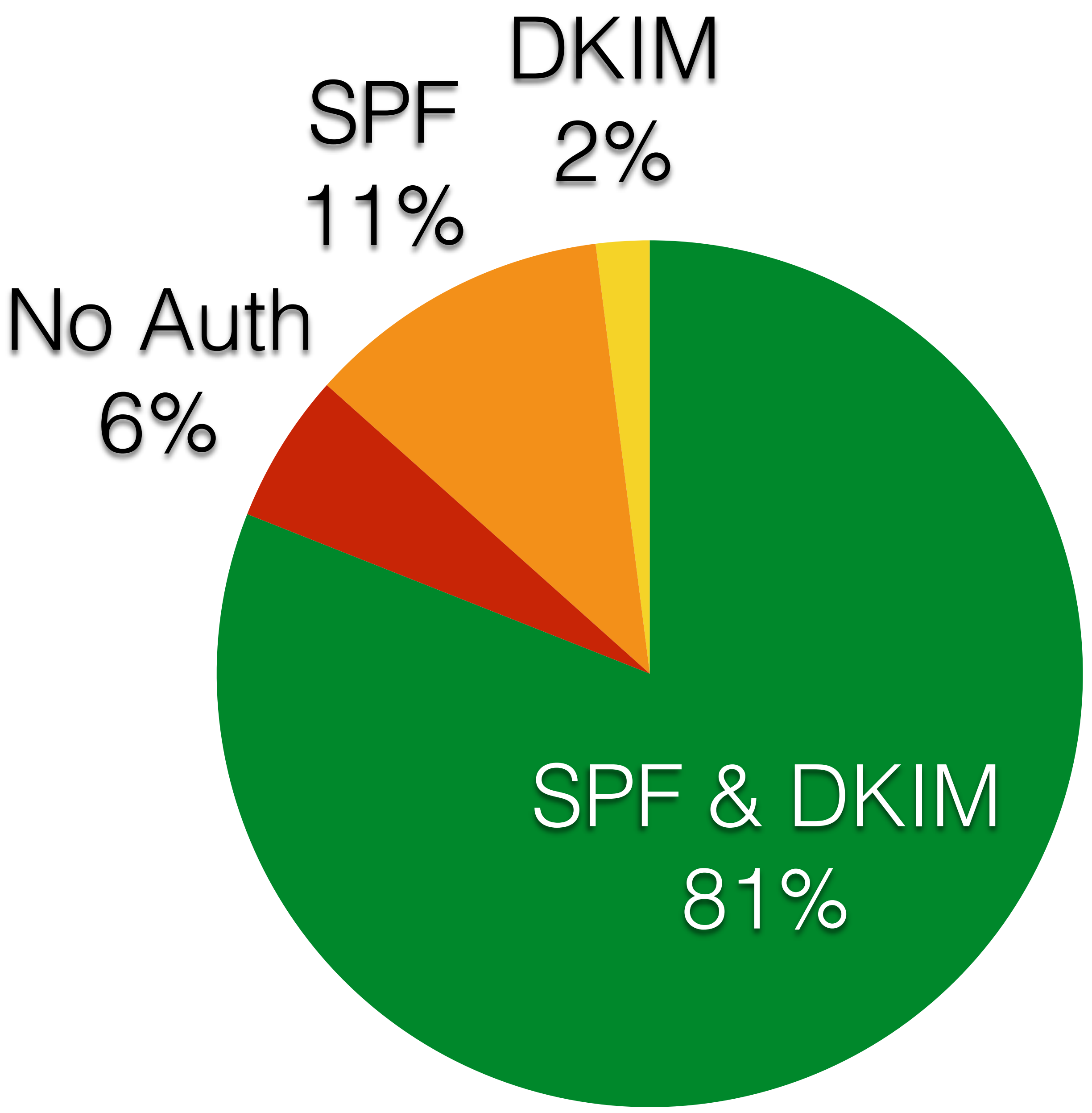
2. Recipient checks for a sender’s policy and if they should reject messages without signatures, and/or report them to the sender

Authentication from Gmail Perspective



Delivered Gmail Messages

Authentication from Gmail Perspective



Delivered Gmail Messages

Technology	Top 1M
SPF Enabled	47%
DMARC Policy	1%

DMARC Policy	Top 1M
Reject	20%
Quarantine	8%
None	72%

Top Million Domains

Moving Forward

Two IETF proposals to solve real world issues:

SMTP Strict Transport Security

Similar to HTTPS HSTS (key pinning)

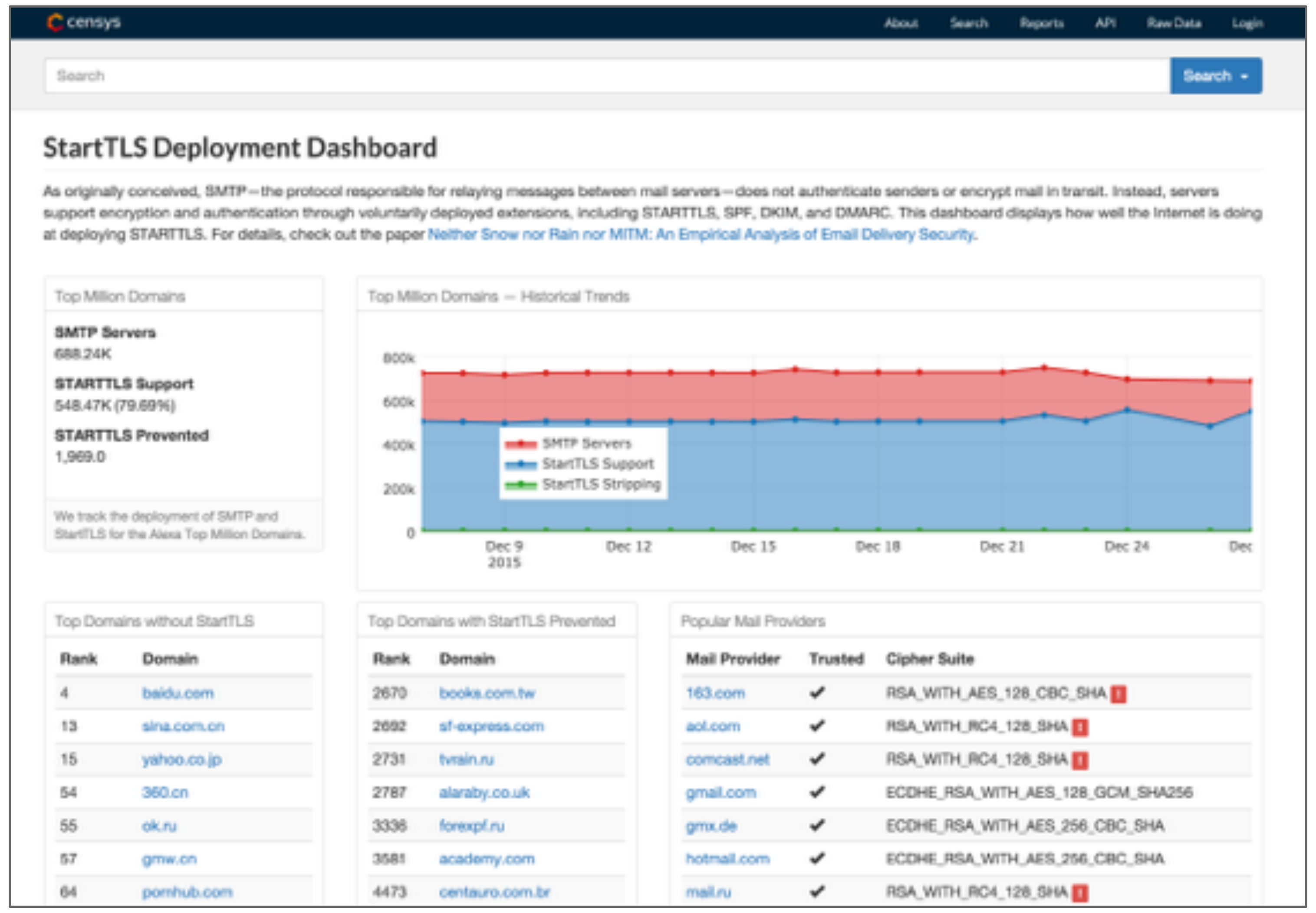
Authenticated Received Chain (ARC)

DKIM replacement that handles mailing lists

Tracking Progress

Censys STARTTLS Report

https://censys.io/reports/mail



StartTLS Deployment Dashboard

As originally conceived, SMTP—the protocol responsible for relaying messages between mail servers—does not authenticate senders or encrypt mail in transit. Instead, servers support encryption and authentication through voluntarily deployed extensions, including STARTTLS, SPF, DKIM, and DMARC. This dashboard displays how well the Internet is doing at deploying STARTTLS. For details, check out the paper Neither Snow nor Rain nor MITM: An Empirical Analysis of Email Delivery Security.

Top Million Domains

SMTP Servers
688.24K

STARTTLS Support
548.47K (79.69%)

STARTTLS Prevented
1,969.0

We track the deployment of SMTP and StartTLS for the Alexa Top Million Domains.

Top Domains without StartTLS

Rank	Domain
4	baidu.com
13	sina.com.cn
15	yahoo.co.jp
54	360.cn
55	ok.ru
57	gmw.cn
64	pornhub.com

Top Domains with StartTLS Prevented

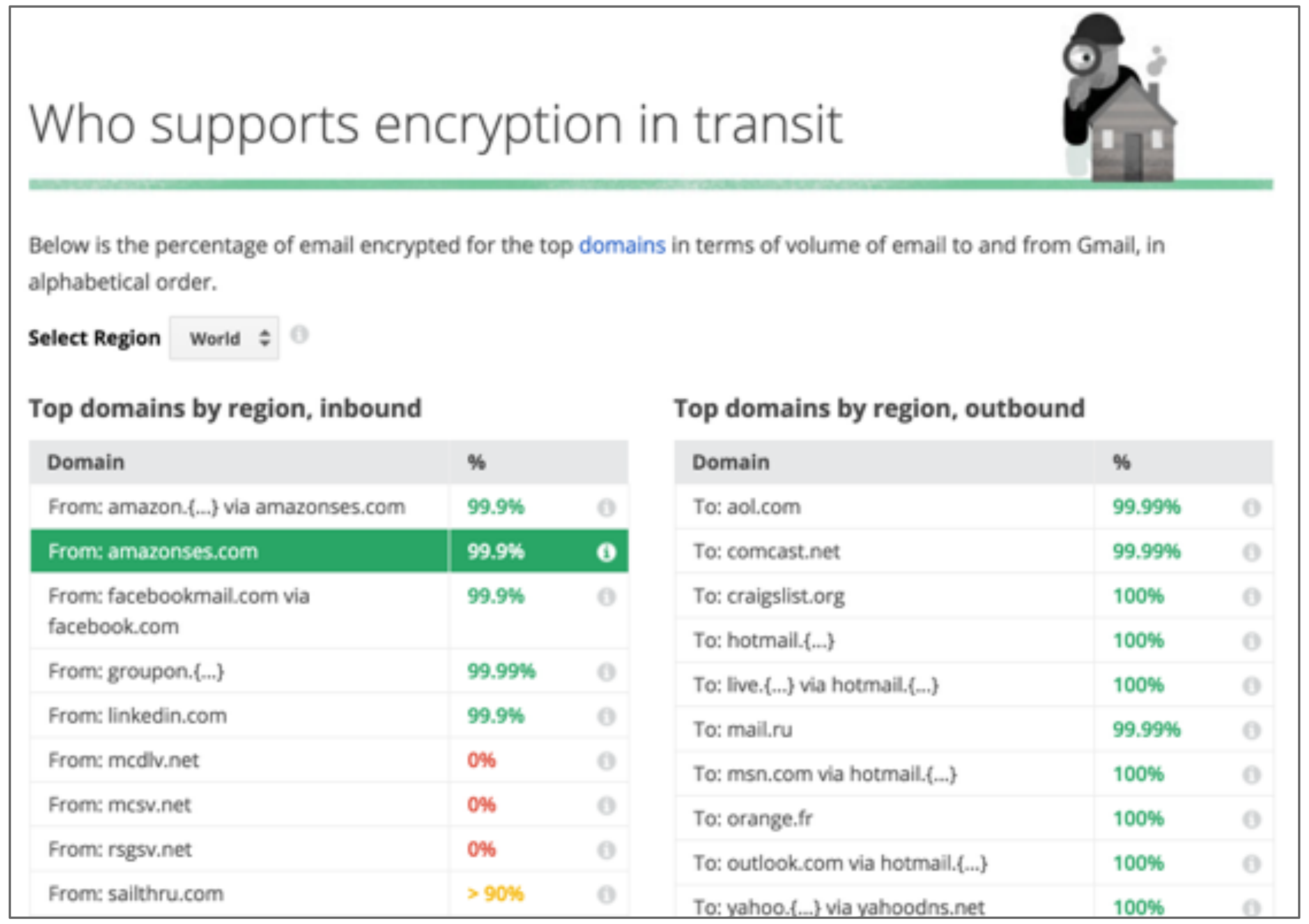
Rank	Domain
2670	books.com.tw
2692	sf-express.com
2731	tvrain.ru
2787	alaraby.co.uk
3336	forexpf.ru
3581	academy.com
4473	centauro.com.br

Popular Mail Providers

Mail Provider	Trusted	Cipher Suite
163.com	✓	RSA_WITH_AES_128_CBC_SHA !
aol.com	✓	RSA_WITH_RC4_128_SHA !
comcast.net	✓	RSA_WITH_RC4_128_SHA !
gmail.com	✓	ECDHE_RSA_WITH_AES_128_GCM_SHA256
gmx.de	✓	ECDHE_RSA_WITH_AES_256_CBC_SHA
hotmail.com	✓	ECDHE_RSA_WITH_AES_256_CBC_SHA
mail.ru	✓	RSA_WITH_RC4_128_SHA !

Google Transparency Report

https://www.google.com/transparencyreport/saferemail



Who supports encryption in transit

Below is the percentage of email encrypted for the top domains in terms of volume of email to and from Gmail, in alphabetical order.

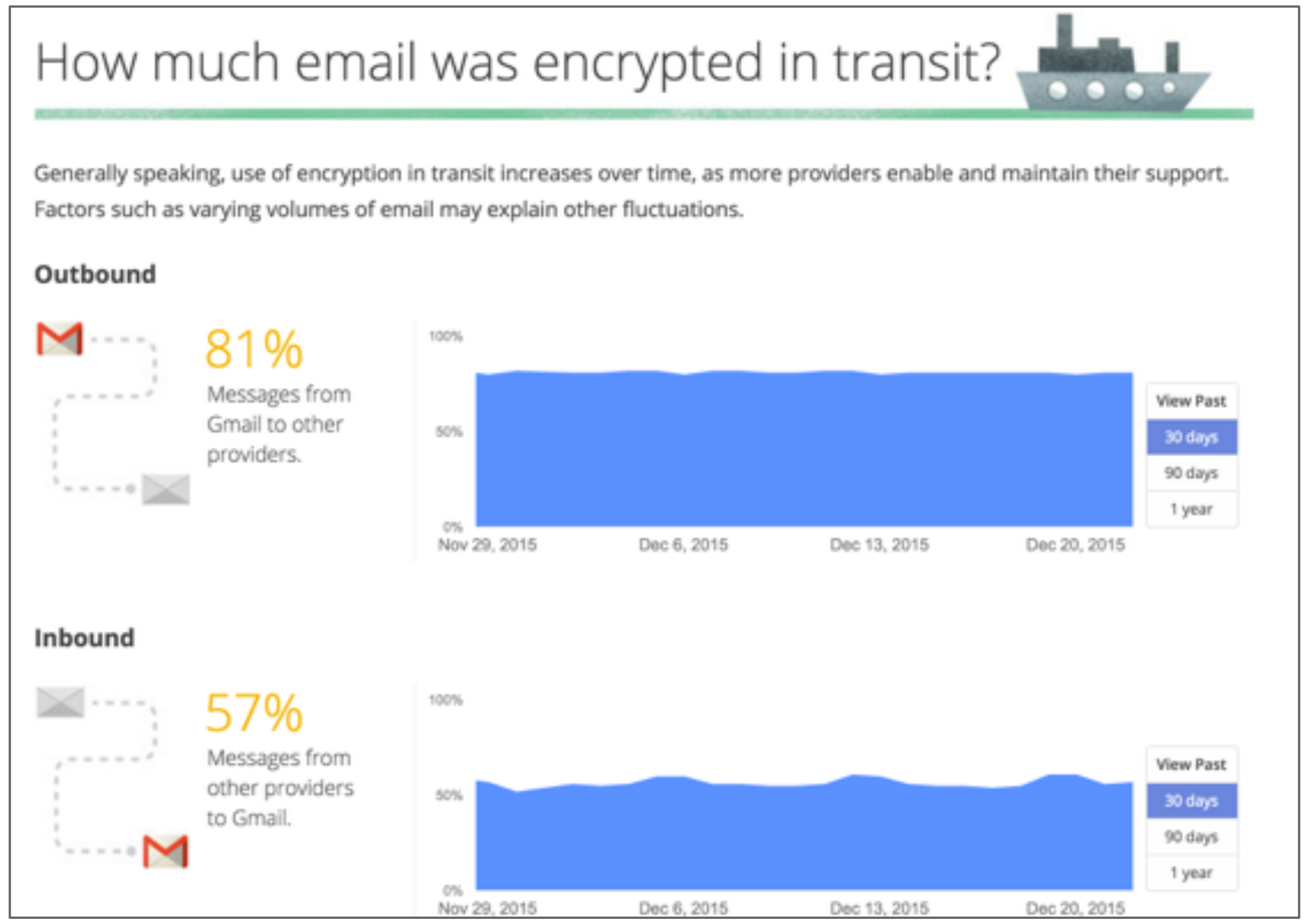
Select Region World

Top domains by region, inbound

Domain	%
From: amazon.{...} via amazonses.com	99.9%
From: amazonses.com	99.9%
From: facebookmail.com via facebook.com	99.9%
From: groupon.{...}	99.99%
From: linkedin.com	99.9%
From: mcdlv.net	0%
From: mcsv.net	0%
From: rsgsv.net	0%
From: sailthru.com	> 90%

Top domains by region, outbound

Domain	%
To: aol.com	99.99%
To: comcast.net	99.99%
To: craigslist.org	100%
To: hotmail.{...}	100%
To: live.{...} via hotmail.{...}	100%
To: mail.ru	99.99%
To: msn.com via hotmail.{...}	100%
To: orange.fr	100%
To: outlook.com via hotmail.{...}	100%
To: yahoo.{...} via yahoodns.net	100%



How much email was encrypted in transit?

Generally speaking, use of encryption in transit increases over time, as more providers enable and maintain their support. Factors such as varying volumes of email may explain other fluctuations.

Outbound

81%
Messages from Gmail to other providers.

Inbound

57%
Messages from other providers to Gmail.

Conclusion

Mail community has started to deploy new security extensions, but progress is slow for many organizations

Unfortunately, until near pervasive deployment, it is unlikely that operators will require encryption

Clear that StartTLS is not a long-term solution—attacks are pervasive in many regions

Both researchers and mail operators have a lot of remaining work to do

Neither Snow Nor Rain Nor MITM… Real World Email Delivery Security

Zakir Durumeric
@zakirbpd