

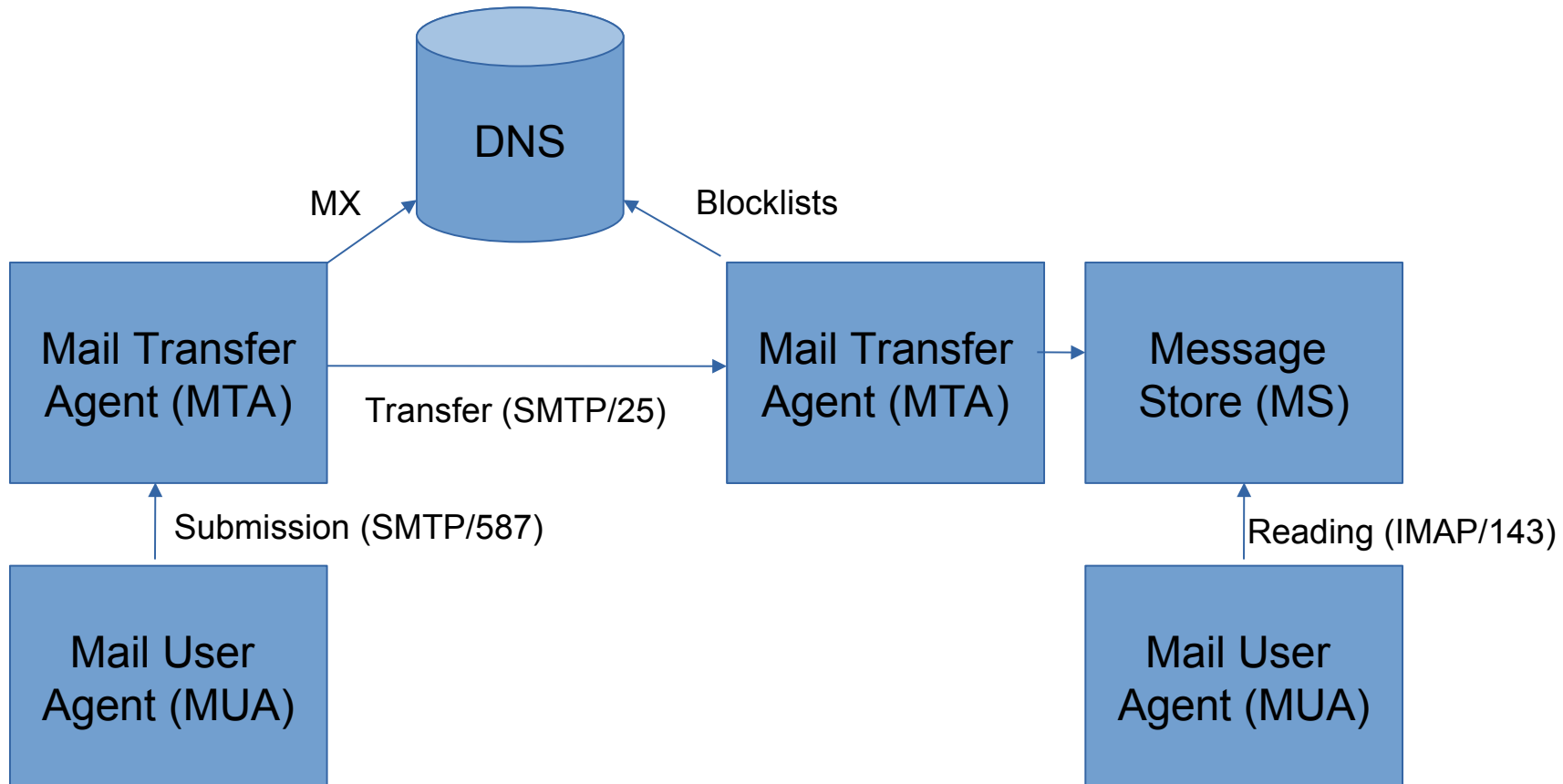
Next hour(-ish)

- Mail architecture & mail transport security
 - What is spam?
 - How much is there?
 - How is it created/inserted?
 - How might it be mitigated?
 - Some ways...
 - Future...
-
- And while we're at it, we'll cover a load of stuff about email in general

Materials

- Additional slides:
 - Jim Fenton's slides
<https://down.dsg.cs.tcd.ie/cs7053/lectures/fenton.pdf>
 - Murray Kucherawy's slides:
<https://down.dsg.cs.tcd.ie/cs7053/lectures/slides-87-dmarc-3.pdf>
- DomainKeys
 - <http://dkim.org/> (still no https!)
 - One anti-spam technique
 - Note: I helped out there
- Misc silliness:
 - <https://www.spam.com/>
 - <https://down.dsg.cs.tcd.ie/ubc206/spam/spam-song.wav>

Another architectural view



Drive-by Terminology

- MUA, MTA, MSA, MDA, MS ...
 - MSA: Mail submission agent
 - MDA: Mail delivery agent
- Message envelope
- Forwarder, exploder, ...
- Message headers:
 - From:, Sender:, Resent-From:
- 2821/2822
 - EHLO/HELO

Mail Transport Security Stats

- Gmail publish “transparency report” stats
 - <https://transparencyreport.google.com/safer-email/overview>
- Other service providers see commensurate numbers
- Recently: 89% of outbound and 93% of inbound are protected via some form of TLS
 - Note: “Some form” == possibly opportunistic!
- Regional variations remain

Opportunistic SMTP/TLS (1)

- Sending MTA looks up MX RR of recipient's DNS domain
- MX name might not be same as recipient's domain, e.g., tcd.ie's MX is msft/outlook
- Sending MTA is TLS client, receiving MTA is TLS server.
- Question: which name to require in TLS server cert? MX name or recipient domain?
- Answer: many receiving domains outsource MX to someone big and good at anti-spam meaning sending MTA has to deal with MX names and can't insist MX be able to serve TLS certificate of recipient domain.

Opportunistic SMTP/TLS (2)

- Today many MX's still serve self-signed certificates or expired certificates or certificates signed by some CA not trusted by sending MTA (e.g. a corporate CA).
- For all the above reasons, SMTP/TLS is often still “opportunistic” in that MTAs enable it, but don't impose the same level of checks as are done on the web
- Note: on the web, there's often a warm body who can see/react to errors or warnings – there is no human user present when a sending MTA tries to establish SMTP/TLS with a receiving MTA.

MTA-STS

- To improve on opportunistic SMTP/TLS we need receiving MTAs to publish a signal that it's ok for sending MTAs to be more strict
- MTA Strict Transport Security (RFC 8461) defines a way to publish that information in the DNS (so sending MTA can look for it before starting to connect)
 - Modelled on HSTS and DKIM
- Important: has a “testing” phase and defines how sending MTAs can report stats occasionally

What is spam?

- Various acronyms:
 - Unsolicited bulk email (UBE)
 - Unsolicited commercial email (UCE)
- Spam is bad:
 - Resource consumption
 - Filters, scanners etc. cost time & money
 - Malware
 - Phishing attempts

Sometimes hard to know...

HILARY TERM GREETINGS FROM THE
COLLEGE CHAPLAINS The College
Chaplains send best wishes to all, and
would like to bring the following upcoming
events to your attention. They are open to
any students or staff members who wish to
join us. ...

Original spam tricks

- Just send email!
 - Ahh...the naivety of it all!
- Email to list
 - Listservers got better, e.g. subscriber only with controlled subscription
- Forge headers
- Send via open relay
 - Used to be a lot of these, very few now
 - toad.com is (or was) an exception!

More spam tricks...

- Confusion:
 - accounts@paypa1.com
 - support@eboy.com
 - postmaster@boi-support.com
 - About to get worse thanks to I18N
 - security@bigbank.com
 - ^ Unicode 0430 is cryllic small 'a'
- Throwaway domains/addresses
- Zombie'd hosts
- BGP spoofing
- Fake ISPs

Yet more

- HTML messing
 - Colour-related
 - Relay sites
 - Encoded URIs
 - Font size 0: break words with zero width spaces

How much spam is there?

- Lots
 - Hard to get good figures, these are ones I've overheard
- ISP backbones:
 - 70% + of email traffic
- Delivered mail:
 - 40% + delivered
- Increasing or not?
 - Harder to tell if MTAs silently filter

How to solve spam...

- Over to you...
- But beware of FUSSP
 - <https://www.dmath.org/fussp/>

Recent Anti-spam techniques

- Content filtering (Bayesian, etc.)
 - Misc. AI magic
- DNS Black Lists (SORBS, DNSBL)
- Register of known spam operators (ROKSO)
- Greylisting
 - All in widespread use today
- Sender Policy Framework (SPF)
- Domain Keys Identified Mail (DKIM)
 - I'll use Jim's slides for these.
- DMARC!!! Argh!!!
 - I'll use Murray's slides for that.

DMARC Downsides

- So DMARC sounds great but, there are valid and (to some) important uses of 3rd party sending:
 - Mailing lists, alumni addresses
- Mailing lists are the main tool for discussing how Internet standards (incl. Mail, incl. DKIM, incl. DMARC) should work, so breaking those seems dim
- But some mail service providers will prefer to reduce their costs even so:

<https://www.pcworld.com/article/2141120/yahoo-email-antispooing-policy-breaks-mailing-lists.html>

<https://wiki.asrg.sp.am/wiki/>

Mitigating_DMARC_damage_to_third_party_mail

DMARC Downsides (2)

- When sender to list is from a domain that publishes “p=reject” policy list-recipients may bounce the mail, leading to non-delivery, but also, after a few such bounces, leading the mailing list to unsubscribe the sender
 - Result: <user>@yahoo.com can no longer participate in e.g. IETF
- Risk: If gmail.com published “p=reject” that would likely cause havoc for organisations that depend on mailing lists
 - Google have said they won’t until the impact of that would be reduced sufficiently
- Workarounds
 - Use gmail or a vanity sending domain you control
 - mailing list re-writes sender’s address, e.g. to
first.last=40example.com@dmARC.iETF.org
- “Solution” for DMARC downsides (ARC) involves DKIM signing by sender’s MTA and a 2nd DKIM(-like) signature from list agent, but is complex, and interpreting ARC protections on a message is even more complex so not clear if ARC will get deployment
 - Authenticated Received Chain (ARC) is RFC 8617

DKIM vs DNSSEC

- DNSSEC DS requires talking to parent domain, chains up to signed root zone
 - ~1% deployment
- DKIM publishes public key in DNS in TXT record with no additional DNS security
 - Today DKIM is almost mandatory if you want your email delivered
 - Partly, that's due to the power of the major mail service providers
 - Spam and phishing still exist and are not “solved”
- Think about the pros and cons?

SPAM summary

- Mail is a historically and currently important Internet application
- Hop-by-hop/transport security for mail has significantly improved since 2013, end-to-end security for mail is still moribund (and staying so, sadly)
- SPAM/Anti-SPAM is an arms-race, beware the FUSSP!
http://www.circleid.com/posts/20131226_the_naive_arrogance_of_fussps/
- SPF and DKIM are fairly deployable and widely used
- DMARC is good for some but not all