SPAM

…just can't seem to get enough…

https://www.theguardian.com/environment/2017/feb/13/extraordinary-levels-of-toxic-pollution-found-in-10km-deep-mariana-trench#img-1

Next hour(-ish)

- What is spam?
  - How much is there?
- How is it created/inserted?
- How might it be defeated?
  - Some ways…
- Future…

- And while we're at it, we'll cover a load of stuff about email in general

Materials

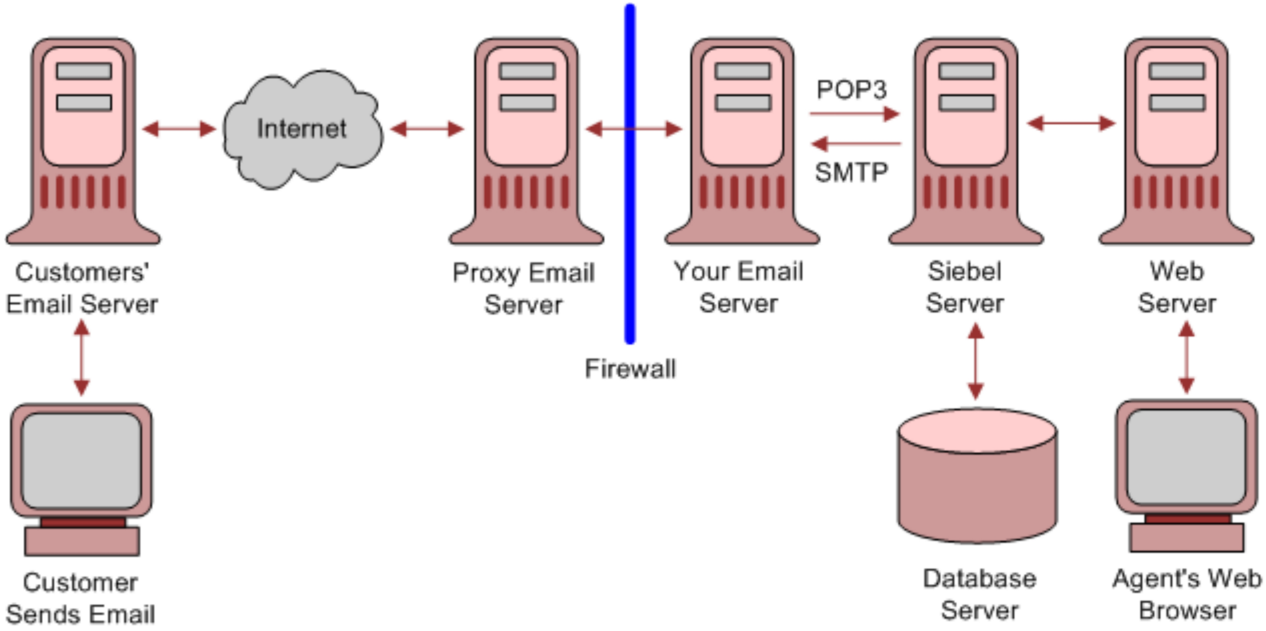
- These slides are partly copied from Jim Fenton's slides
  - https://down.dsg.cs.tcd.ie/cs7053/lectures/fenton.pdf
- DomainKeys
  - http://dkim.org/
  - One anti-spam technique
    - Note: I helped out there
- Misc silliness:
    - http://www.spam.com/ (*Still* no https, sorry:-)
    - https://down.dsg.cs.tcd.ie/ubc206/spam/spam-song.wav

Drive-by Terminology

- MUA, MTA, MSA, MDA, MS …
- Message envelope
- Forwarder, exploder, …
- Message headers:
  - From:, Sender:, Resent-From:
- 2821/2822
  - EHLO/HELO

Email Architecture

- See RFC 5598



https://docs.oracle.com/cd/E63029_01/books/SecurHarden/img/architecture_email_v.gif

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 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 hosts
- Trojans
- 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…

Recent Anti-spam techniques

- Content filtering (Bayesian, etc.)
- DNS Black Lists (SORBS, DNSBL)
- Register of known spam operators (ROKSO)
- Greylisting
  - All in widespread use today
- Sender Policy Framework (SPF)
- SenderID
- Purported Responsible Address (PRA)
- Domain Keys Identified Mail (DKIM)
- Certified Server Validation (CSV)
  - 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:

http://www.pcworld.com/article/2141120/yahoo-email-antispoofing-policy-breaks-mailing-lists.html

http://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
- “Solution” for DMARC downsides being developed (ARC) involving DKIM signing by sender’s MTA and a 2nd DKIM(-like) signature from list agent, but that is complex, and interpreting ARC protections on a message is even more complex so not clear if a technical solution will get deployment
- Work on ARC being done in IETF DMARC WG:
  https://tools.ietf.org/wg/dmarc/
  https://tools.ietf.org/wg/dmarc/draft-ietf-dmarc-arc-protocol/

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