iDA SINGAPORE

INFOCOMM DEVELOPMENT AUTHORITY OF SINGAPORE

Catalyst for Change

Curbing Spam via Technical Measures – An Overview

07 Jul 2004

Ho Khee Yoke & Lawrence Tan
Infocomm Development Authority of Singapore

What is spam?

- Definition is still quite controversial - everyone interprets it differently
- One broadly accepted definition: **bulk unsolicited commercial e-mail**
- For most end-users: Mails that I don't want to see in my mailbox

Catalyst for Change

Why do we get spam?

Because sometimes we give out our email addresses unknowingly…

Have you dropped your namecard in a public seminar/exhibition or restaurant?

Have you registered for any (free) Internet services using your email address?

Have you posted any replies in web forums?

Have you opened up unusual email attachments?

Catalyst for Change

Why do we get spam…

… even when we have been careful not to give away our e-mail addresses freely?

Dictionary attack

john@hotmail.com

mary@yahoo.com

meiling@singnet.com.sg

Rogue service providers re-sell e-mail addresses to spammers

Or

Your service provider could have been hacked!

Why do spammers spam?

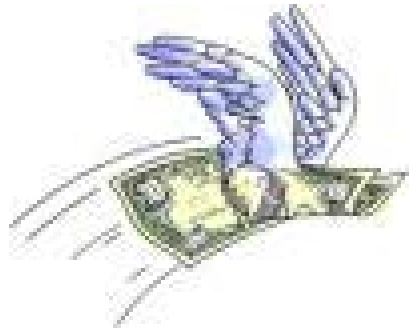
Because it is a low-cost, highly profitable business…

- Start-up cost is low
  - US$150 for a 20 million e-mail list
- Cost of delivering email is low
  - Estimated at 0.05 cents per email as compared to $1.21 per snail mail
- Flaws in Internet e-mail architecture (SMTP) allow mails to be sent anonymously
- Some people actually respond to spam and purchase the advertised products/services

Catalyst for Change

General strategies to stop spam

- Via economic & legal disincentives
  - Increasing the cost of sending spam
  - Reducing the effectiveness of spam



Make spam business less lucrative and spammers will go away…

- Reduce anonymity in email system
  - No false e-mail address/subject
  - No masquerading as someone else

Encourage responsible email marketing so that end-users can choose what they receive…

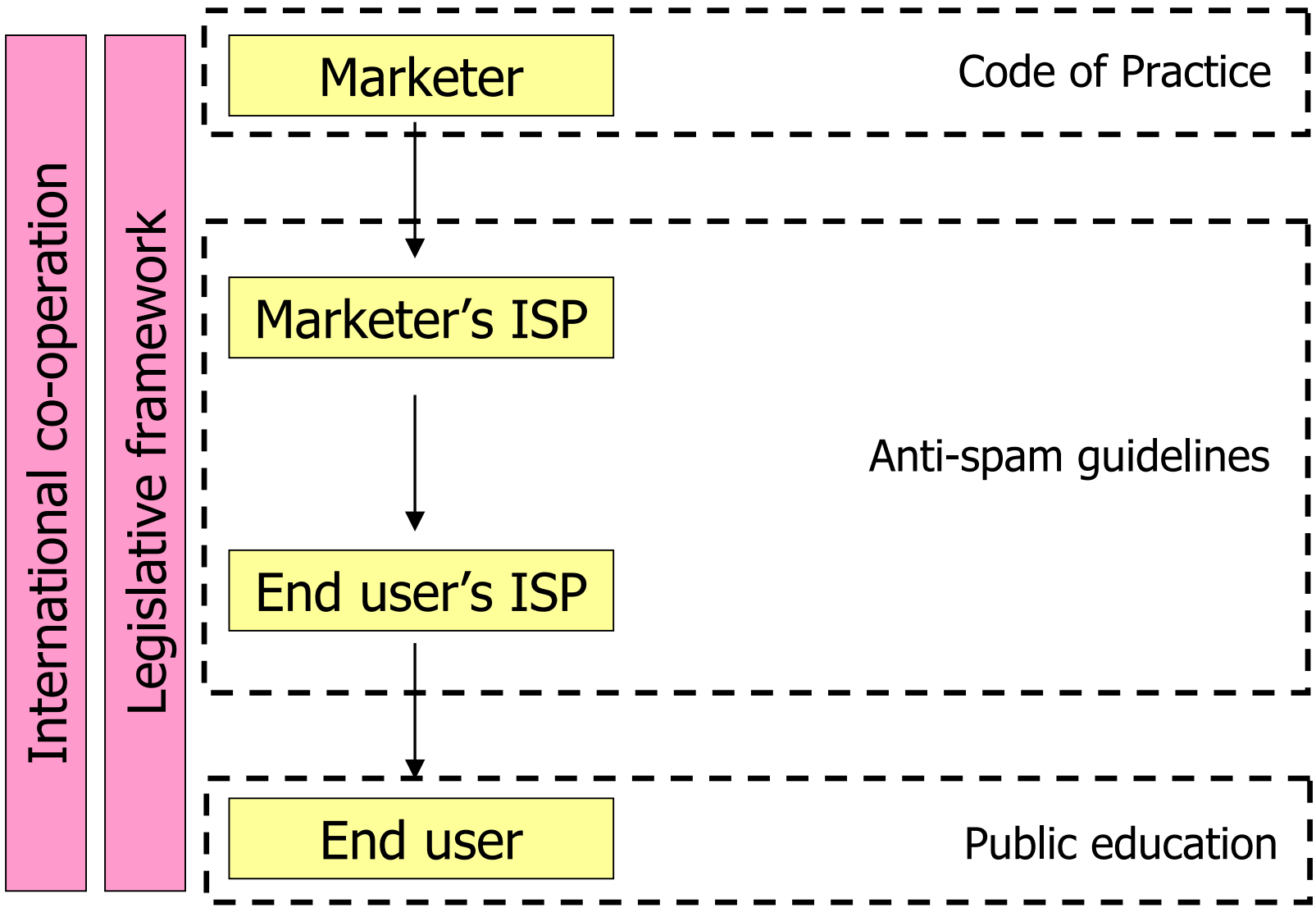
Catalyst for Change

The reality…

- There is no silver bullet to curb spam
- On their own, every measure has been circumvented by spammers, eg:
  - Keyword filters → Cleverly mis-spelled words
  - Adaptive filters → Append legitimate tokens to confuse the filters
  - Blacklisting → Domain spoofing / register with new ISPs
  - Legislation → Legal loopholes / move operation to spam haven
- A multi-pronged approach comprising **legislation**, **technical measures**, **public education**, **industry self-regulation** & **international co-operation** is the best way of tackling the problem

Catalyst for Change

A multi-pronged approach against spam



Catalyst for Change

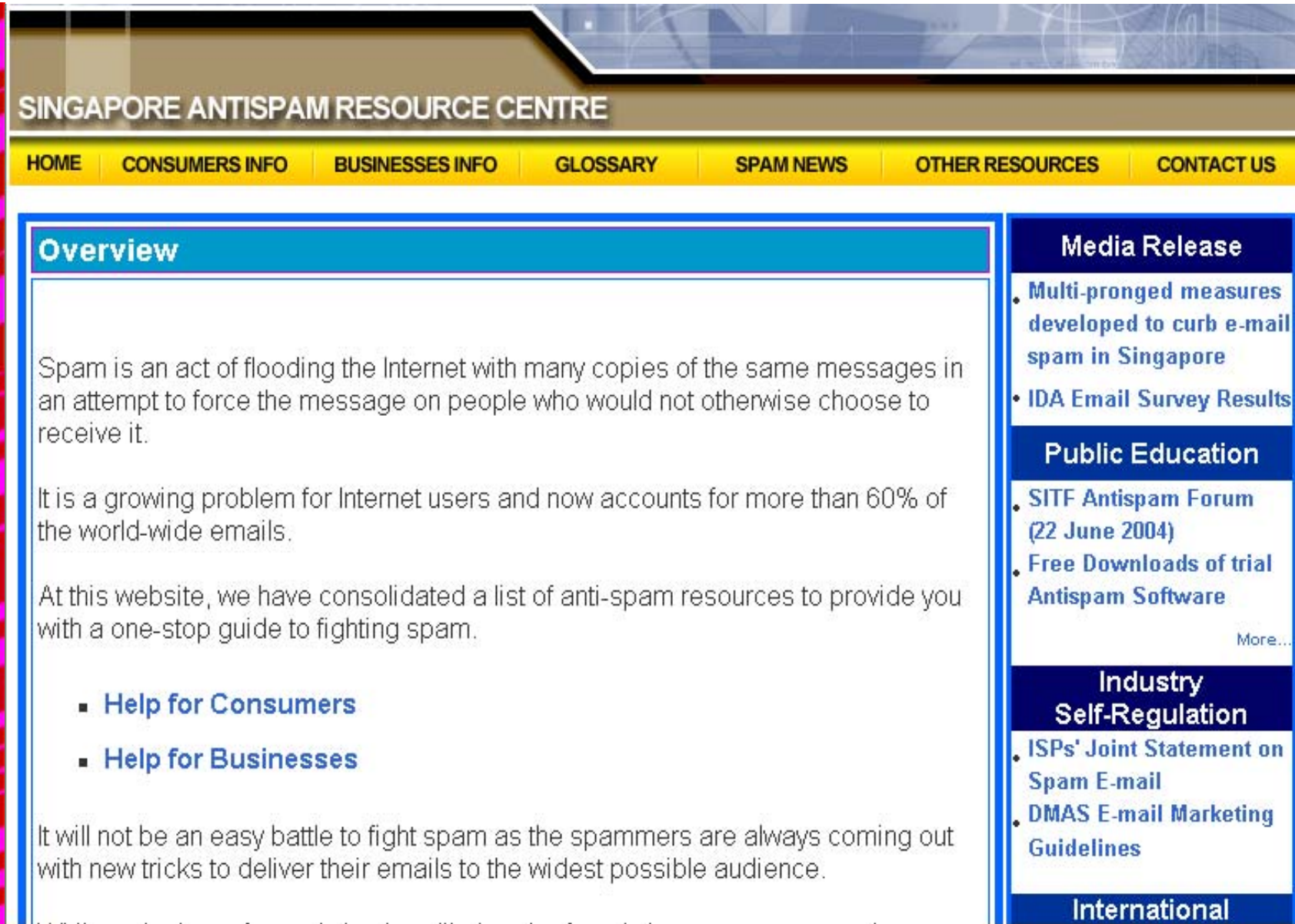
Anti-spam in Singapore

- Infocomm Development Authority of Singapore
- Attorney-General's Chambers
- 3 major ISPs
  - Pacific Internet
  - SingNet
  - StarHub Internet
- Consumer Association of Singapore
- Singapore infocomm Technology Federation
- Singapore Business Federation
- Direct Marketing Association of Singapore

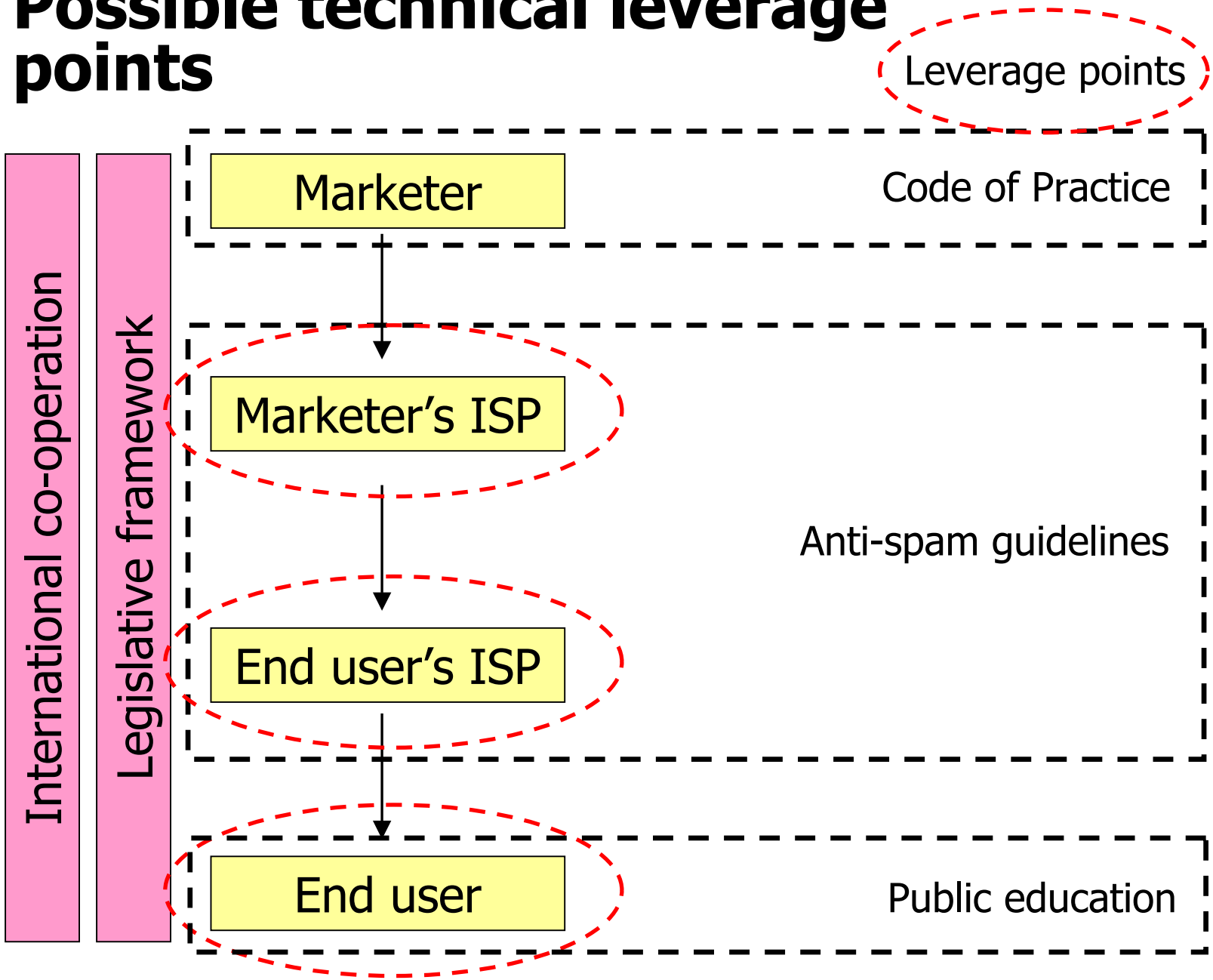
Committed to can spam: (from left) Ms Cheryl Kong, deputy division director of Singapore Business Federation, Mr Saw Ken Wye, chairman of Singapore infocomm Technology Federation, Ms Lisa Watson, chairman of Direct Marketing Association of Singapore, Mr Charles Lim, principal senior state counsel, AGC, Mr Leong Keng Thai, deputy chief executive of IDA, Mr Seah Seng Choon, executive director, Consumers Association of Singapore, Mr Walter Lee, Wireless Broadband/IP VAS, StarHub Interactive, Ms Ooi Lay Yong, chief executive officer of SingNet, and Ms Marian Phuah, vice-president of Consumer Lifestyle Group, Pacific Internet.

Source: Singapore Press Holdings - Computer Times

Singapore Antispam Resource Centre (www.antispam.org.sg)

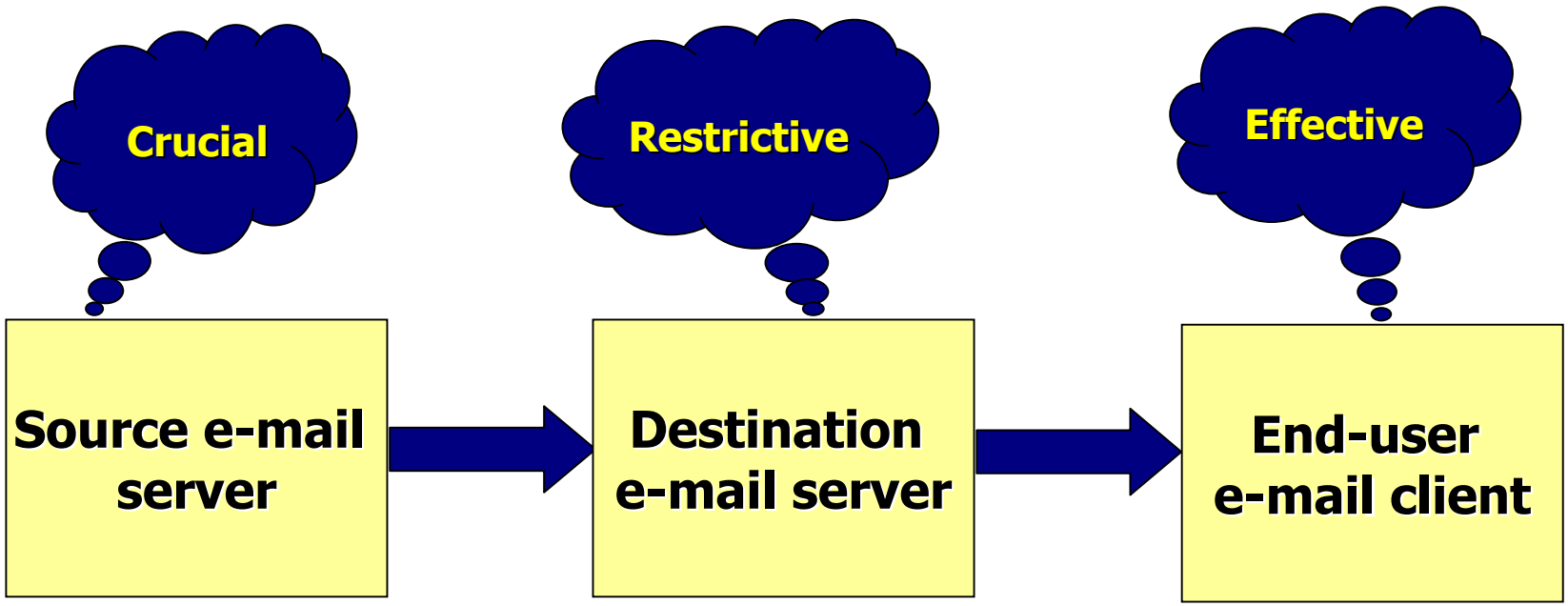


Possible technical leverage points



Catalyst for Change

Overview of technical measures



- Blocking Port 25
- Rate Limiting
- Authentication
- Payment

- Rate Limiting
- Reputation System
- Checksum

- Static Filtering
- Adaptive Filtering
- Reputation System

At the source e-mail server

- Most critical battlefield in the fight against spam
  - Win the war here, and we don't have to bear the cost of handling spam any further down the chain
- Unfortunately, it is a difficult war because
  - Too many e-mail servers
  - Some of them are operated by the spammers
  - Will take a long time for any one solution to be implemented by everyone
- Goal here is to get the "good guys" to make the change first
  - "Bad guys" who don't make the change risk getting blocked by everyone else

Possible measures at the source

Catalyst for Change

Blocking Port 25

- Prevent "zombies" from sending out spam
- Can create problems for those who need to run their own e-mail servers

Rate Limiting

- Limit number of outgoing e-mail (e.g. 100 e-mails per day per account)
- Already a fairly common practice

Authentication

- Prevent spammers from masquerading as you
- Authentication system by itself does not stop spam but complements other anti-spam measures

Payment

- Charge for sending e-mails, for example: 0.1 cent per email – 100 e-mails per day = 10 cents or $3 per month or $36 per year **(Probably costs less than professional antispam services)**
- C.f. cost to spammer of sending 1 million e-mail per day = $1000!
- Difficult to implement – need billing infrastructure, business peering etc.

At the destination e-mail server

- Must do something because
  - Mailbox storage is limited – if mailbox is full, legitimate e-mail may get discarded
  - Bandwidth is limited – Users don't want to download too many e-mails to their client (especially mobile client)
- Yet our options are restricted
  - Tolerance level for spam higher than tolerance for loss of legitimate e-mails
  - One man's spam maybe another man's meat
- Hence we need to adopt a more conservative approach (e.g. tag/quarantine instead of discard spam)

Possible measures at the destination

Rate Limiting

- Limit number of incoming e-mail (e.g. 100 e-mails per day per account)

Reputation System

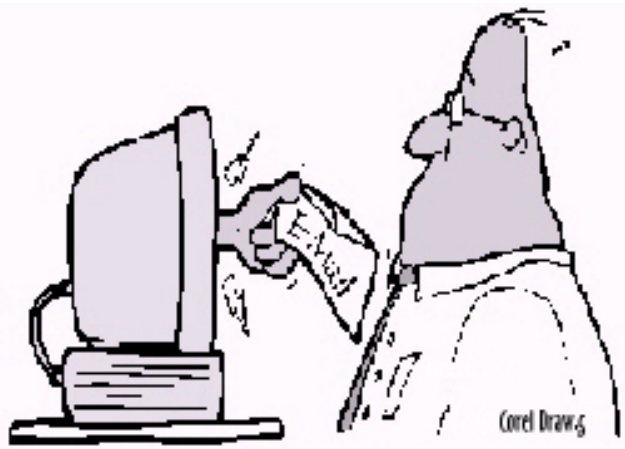
- Decide what to accept based on the known past reputation of the source e-mail server
- Whitelist based on business peering relationship
- Blacklist based on known spam servers, open relay servers
- Blacklist should be use with extreme care – may cause good e-mails to be discarded

Checksum

- Server generally can't tell if an e-mail is spam
- It can suspect an e-mail is spam if it is also sent to many other users on the server
- If similar e-mail is sent to many users in many servers, very high likelihood it is spam
- Compare checksum instead of actual content for privacy reasons

At the end-user client

- Most flexible and effective
- High degree of control over their own e-mails
- Can achieve astonishing results with off-the-shelf software available today



Catalyst for Change

Possible measures for the end-user

Static Filtering

- Filter spam based on incoming email attributes, such as: sender name, subject title, email content etc.
- Need to constantly refine the filtering rules

Adaptive Filtering

- Filter spam based on statistical model
- Ability to learn what is spam and is non-spam based on user preference
- Most popular approach: Bayesian filtering can achieve up to 90% accuracy!!!

Reputation System

- Accept/reject emails based on incoming email address
- Blacklist generally not effective – each spam likely to come from different email address
- Whitelist – allowing only people in your address book to send your email

Conclusions

- None of these measures are perfect
- But they all contribute to the battle against spam
  - Less than 10% of spam makes it to the end-user inbox
  - Blocking of port 25 and server-side authentication will significantly reduce the amount of spam that enters our networks
- Let's be **optimistic!**
- All these measures (technical + legal + other approaches) do not need to add up to 100% effectiveness
- Because spammers will go away once the spam business is no longer commercially viable

Catalyst for Change

THANK YOU