Introduction to the Work of the ASRG and Consent-Based Communications

Presented by Yakov Shafranovich, ASRG Co-chair

The Open Group Messaging Forum

Washington, DC area, October 22rd, 2003

Table of Contents

- Part 1:
  - About the ASRG
- Part 2:
  - Consent Based Communications
- Part 3:
  - Current ASRG Status, Selective Proposals and Activities
  - Questions?

Part 1 of 3: About the ASRG.

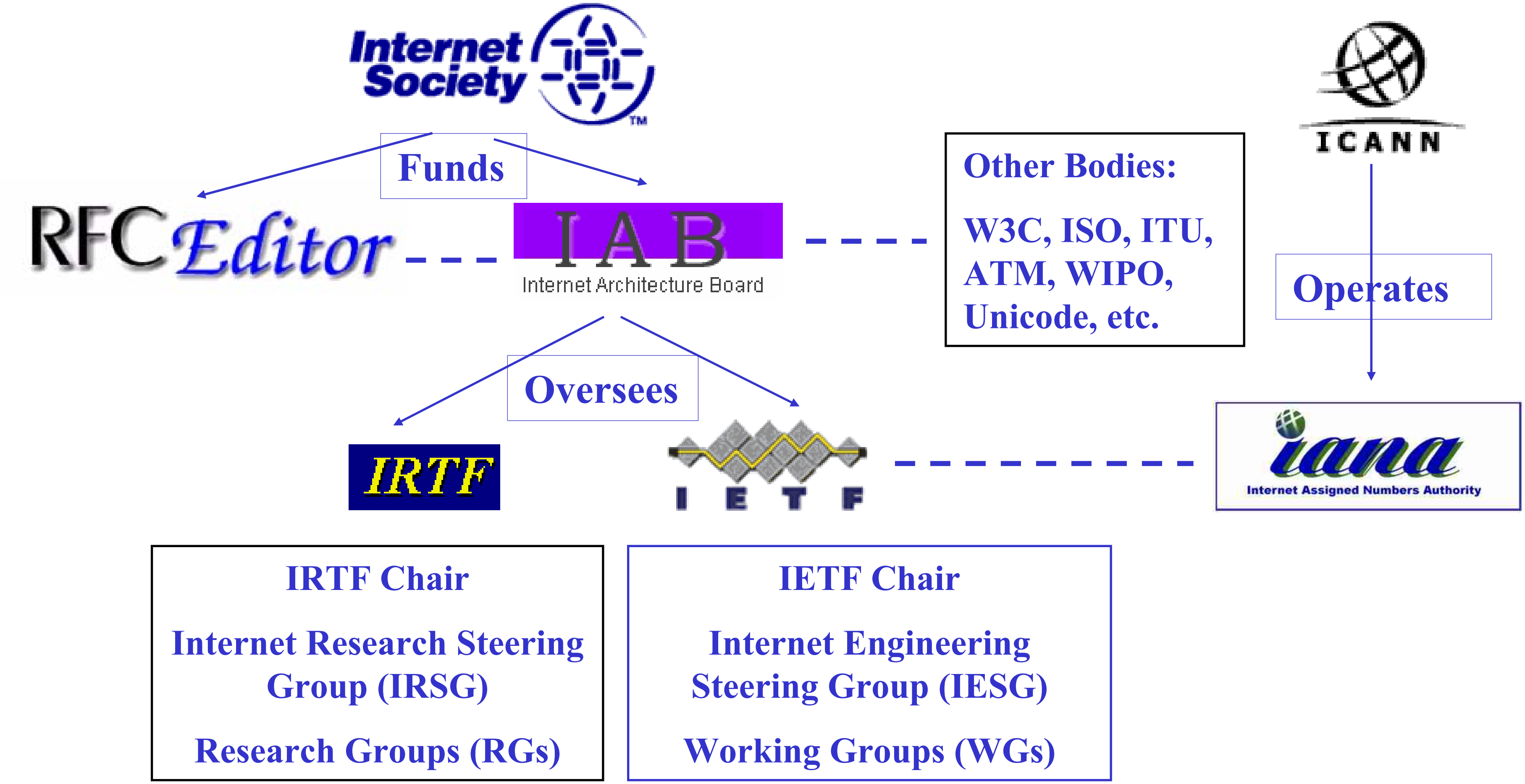
"What's in a name? That which we call a rose
By any other word would smell as sweet."

***"Romeo and Juliet"* by William Shakespeare**

Part 1 of 3: About the ASRG

- **Table of Contents**
  1. Internet Standards Bodies and Their Roles
  2. Differences Between the IRTF and IETF
  3. What is the ASRG?
  4. Why was the ASRG Created?
  5. Goals of the ASRG
  6. ASRG Research Agenda
  7. ASRG Organizational Structure

1.1. Internet Standards Bodies and Related Organizations


Internet Society
RFC Editor
IAB
Internet Architecture Board
ICANN
IRTF
I E T F
iana
Internet Assigned Numbers Authority

Funds

Oversees

Operates

Other Bodies:

W3C, ISO, ITU, ATM, WIPO, Unicode, etc.

IRTF Chair

Internet Research Steering Group (IRSG)

Research Groups (RGs)

IETF Chair

Internet Engineering Steering Group (IESG)

Working Groups (WGs)

1.1. Roles of Internet Standards Bodies and Related Organizations

- **Internet Society (ISOC)**
  - Professional membership organization of Internet experts
  - Funds and oversees IAB, IRTF, IETF and RFC Editor
- **Internet Architecture Board (IAB)**
  - A committee of 13 Internet experts chosen by the IETF
  - Provides oversight of Internet architecture, IETF and IRTF
- **The RFC Editor**
  - Edits and publishes Request for Comments (RFC) documents
  - Independent of the IETF and IRTF
- **Internet Assigned Numbers Authority (IANA)**
  - Operated by ICANN on behalf of the IETF
  - Maintains unique parameters for Internet protocols and standards
- **Internet Corporation for Assigned Names and Numbers (ICANN)**
  - Operates the Domain Name System (DNS) under contract with the US Government

1.2. Differences Between IETF and IRTF

- **Internet Engineering Task Force (IETF)**
  - focuses on the ***short-term*** issues of engineering and standards making
  - Operates ***more formally***
  - Consists of 100+ working groups ***working on Internet standards***
- **Internet Research Task Force (IRTF)**
  - focuses on ***long-term*** research issues related to the Internet
  - Operates ***more informally***
  - Consists of 12 research groups ***doing research*** on Internet related issues

1.3. What is the ASRG?

- A *Research* Group (RG) of the IRTF
- An open membership RG, *possible spammer members*
- Formed in March of 2003, founded by Paul Judge
- Membership
  - ***Over 650+*** list ***subscribers*** in addition to website visitors
  - ***Over 6,000+*** mailing list ***messages*** in archive
  - Membership on ***individual basis***, not organizational (RFC 2014)
- Co-Chairs:
  - Dr. Paul Q. Judge
  - Yakov Shafranovich

CipherTrust®
ENTERPRISE EMAIL SECURITY

SolidMatrix
TECHNOLOGIES

1.4. Why was the ASRG Created?

- *Scale, growth, and effect* of spam on the Internet have generated considerable interest in addressing this problem
- Once considered a nuisance, spam has grown to account for a *large percentage of the mail volume* on the Internet.
- This unwanted traffic *stands to affect* local networks, the infrastructure, and the way that people use email.

1.5. Goals of the ASRG

- *Understand* the problem and collectively *propose* and *evaluate* solutions
- *Investigate* the feasibility of *consent-based architecture or framework* to allow individuals and organizations to express consent or lack of consent, and enforce their decisions
- *Will not* pursue research into *legal* issues of spam, other than the extent to which these issues affect, support, or constrain the technology

1.6. ASRG Research Agenda

- **Understanding phase**
- **Proposal Phase**
- **Evaluation Phase**

1.6. ASRG Research Agenda

- **The *understanding phase* includes:**
  - Inventory of problems
  - Analysis and characterization:
    - Analysis of Actual Spam Data
    - Public Trace Data

1.6. ASRG Research Agenda

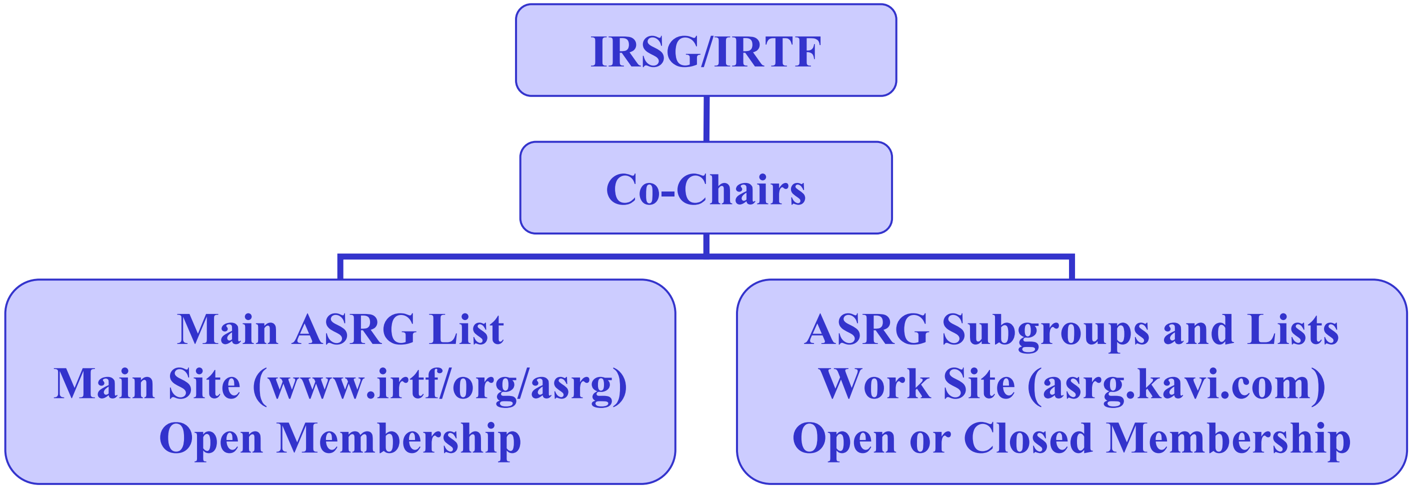
- **The _proposal phase_ includes:**
  - Requirements document
  - Survey of Solutions
    - Taxonomy of solutions
    - Bibliography of spam-related research
    - Consent Framework and related work
  - Identifying standardization requirements
    - Possible later transfer to the IETF
  - Proposals
  - Best Current Practices

1.6. ASRG Research Agenda

- **The *evaluation phase* includes:**
  - Creating an evaluation model
    - Technical Considerations document
    - Requirements document
    - Consent framework
  - Evaluation of Solutions
    - Overall survey
    - Individual proposal by proposal evaluation

1.7. ASRG Organizational Structure.

- Governed by RFC 2014
  - Informal
  - No consensus required
  - Individual not organizational membership



IRSG/IRTF

Co-Chairs

Main ASRG List
Main Site (www.irtf/org/asrg)
Open Membership

ASRG Subgroups and Lists
Work Site (asrg.kavi.com)
Open or Closed Membership

Part 2 of 3:
Consent Based Communications.

“Thou shalt not consent unto him, nor hearken unto him; neither shall thine eye pity him, neither shalt thou spare, neither shalt thou conceal him. But thou shalt surely kill him;”

Deuteronomy 13:8-9 (KJV)

Part 2 of 3:
Consent Based Communications.

- **Table of Contents**
  1. The Many Definitions of Spam
  2. ASRG's Definition of Spam
  3. Why Consent?
  4. Defining Consent
  5. Consent Framework: Goals, Process, Components, and Examples.
  6. Advantages
  7. Disadvantages
  8. Challenges and Future Work
  9. Questions?

2.1. The Many Definitions of Spam

- **Spamhaus:** *"The word "Spam" as applied to Email means Unsolicited Bulk Email ("UBE")"*
  - Unsolicited means that the Recipient has not granted verifiable permission for the message to be sent.
  - Bulk means that the message is sent as part of a larger collection of messages, all having substantively identical content.
- **American Heritage Dictionary:** *"Unsolicited e-mail, often of a commercial nature, sent indiscriminately to multiple mailing lists, individuals, or newsgroups; junk e-mail."*

2.1. The Many Definitions of Spam

□ **Spamhaus and MAPS Technical Definition:**

1. The recipient's personal identity and context are irrelevant because the message is equally applicable to many other potential recipients; (BULK)
2. The recipient has not verifiably granted deliberate, explicit, and still-revocable permission for it to be sent; (UNSOLICITED)
3. The transmission and reception of the message appears to the recipient to give a disproportionate benefit to the sender. (BULK)

2.1. The Many Definitions of Spam

- **Direct Marketing Association (DMA):** *"e-mail that misrepresents an offer or misrepresents the originator--or in some way attempts to confuse or defraud people" (from News.com story)*
- **FTC and CAUCE:** *"Unsolicited Commercial Email"*
- **Others:** *Unsolicited Email or Bulk Email*

2.1. The Many Definitions of Spam

- Unsolicited
- Commercial
- Bulk
- Fraudulent
- Unsolicited + Bulk
- Unsolicited + Commercial
- Unsolicited + Bulk + Commercial
- Other combinations, etc.

2.1. The Many Definitions of Spam.

- Definition varies from "unsolicited commercial email" to "any email the recipient does not want“
- Often there are no technical differences between spam and "acceptable" email
- Format, content and even aggregate traffic patterns may be identical
- "Bulk" is usually very difficult for an individual recipient to prove, but almost always easy to recognize in practice.
- More detailed discussion must, of course, be precise in the definition of "unsolicited“

2.2. ASRG's Definition of Spam.

- We all agree that we disagree
- We want to leave the definition of spam to be defined by each end-user and ISP as they want
- We do not have an official definition and are not seeking for one
- For most working discussions, the term "Unsolicited Bulk Email" is sufficient

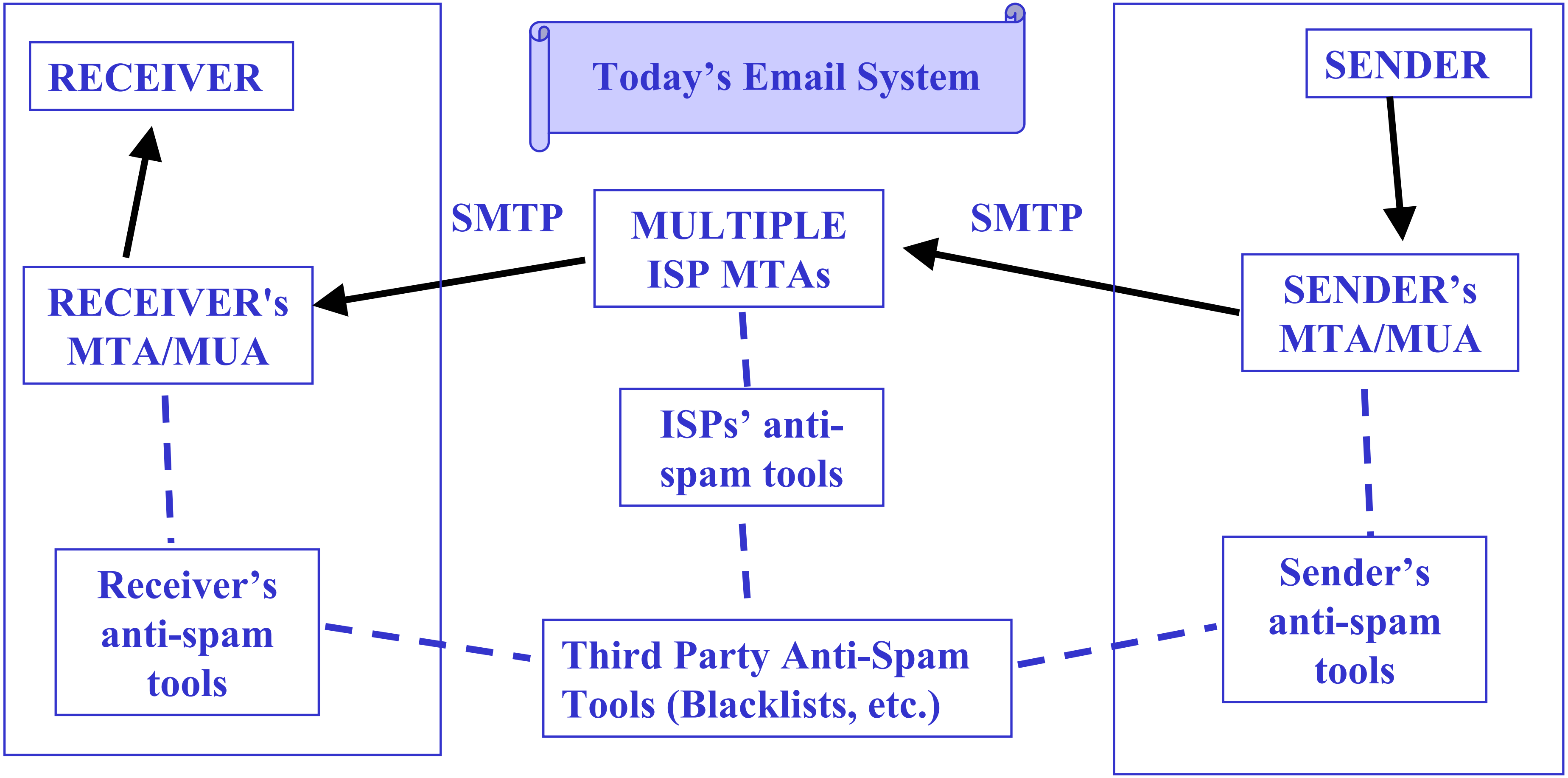
2.3. Why Consent?

- ASRG Charter:
  - *"The definition of spam messages is not clear and is not consistent across different individuals or organizations"*
  - *Therefore, we generalize the problem into "consent-based communication"*
  - *This means that an individual or organization should be able to express consent or lack of consent for certain communication and have the architecture support those desires."*
- Spam may be a permanent part of the Internet like war, taxes, death and cockroaches

2.4. Defining Consent.

- Consent:
  - Expression of wanting to receive specific communications
- Lack of Consent:
  - Expression of not wanting to receive specific communications
- Consent need not necessarily be in advance
- Not the same as the legal concept of Consent
- Computer approximation of person's wishes, does not always correspond exactly to what the person desires
- Exists in a crude form in Instant Messaging systems

2.5. Consent Framework.



2.5. Consent Framework - Goals.

- Provide a system of systems to tie in all anti-spam tools into one cohesive whole
- Leverage existing protocols and email infrastructure
- Allows users and organizations of use their own definitions of spam
- Allows to components to be plugged in as necessary
- Define a set of standard protocols and formats for expressing and denying consent, and for anti-spam tools to communicate
- Allows users to grant and revoke consent, and make the decision known to the sender

2.5. Consent Framework – Process.

- 1. Users and Organizations <u>Define</u> Consent Rules and Policies
  - User's policy may be shared with the ISP or organization
  - ISP's or organization's policy may override the user's policy (possible privacy and anonymity issues)
- 2. MTAs/MUAs <u>Enforce</u> Consent Policies
  - Information from third parties maybe used for enforcement (Blacklists, e-postage, DCC, etc.)
- 3. Some Information May Be <u>Shared</u> with Sender
  - Requests for additional information (C/R, e-postage, etc.)
  - Grant or revocation of consent (opt-in/opt-out)

2.5. Consent Framework – Examples.

- 1. Consent Expression:
  - GUI tools and configuration files to set settings for anti-spam tools
- 2. Enforcement:
  - Filtering tools and anti-spam tools (SpamAssasin, etc.)
  - Third Party Sources:
    - Blacklists and DNSRBLs (Senderbase, MAPS, Spamhaus, SPEWS, etc.)
    - Coordinated detection systems (SpamCop, DCC, Razor, etc.)
    - Marks/Tags (E-postage, Hashcash, TrustedSender, digital certificates, Habeas, etc.)
- 3. Sharing with Sender:
  - Challenge / Response (MailBlocks, etc.)
  - E-Postage requests (TipJar.com, etc.)

2.5. Consent Framework – Components.

- Standard formats and protocols for defining and sharing of consent policies
- Standard protocols and formats for obtaining information from third parties (such as blacklists)
- Standard protocols and formats for consent and revocation of consent, and for sharing consent decisions with the sender
- Best Current Practices
- Extensibility provided in every protocol and format

2.6. Advantages.

- Allows organizations to choose and integrate multiple anti-spam tools easier, providing a united and coordinated response to spam
- Allows each user and organization to define spam as they see fit
- Allows for automatic processing of challenge/response, opt-in and opt-out requests
- Provides a standard format for an opt-in audit trail
- Allows for easier comparison of different anti-spam proposals and solutions
- Edge solution not requiring changes at the network core

2.6. Disadvantages.

- Puts an additional burden on anti-spam tool vendors
- Requires cooperation from anti-spam tools
- Has significant privacy and anonymity issues
- Scalability is unknown
- Effect on spammers unknown
- Deployment issues need to be studied further

2.7. Challenges and Future Work.

□ Investigate the feasibility of consent framework

□ Define consent framework further

□ Define protocols and formats for consent

□ Investigate scalability and deployment issues

□ Analyze possible effect on spammers

Part 3 of 3: Current ASRG Status, Selective Proposals and Activities.

“Hostile armies may face each other for years, striving for the victory which is decided in a single day”

“Art of War”, Sun Tzu

Part 3 of 3: Current ASRG Status, Selective Proposals and Activities.

- **Table of Contents**
  1. ASRG Status.
  2. Foundational Documents.
  3. Analysis and Characterization Subgroup.
  4. Proposals - DNS-based Authentication Methods.
  5. Proposals – Replacing SMTP.
  6. Proposals – E-postage.
  7. Challenge / Response Internetworking (CRI).
  8. Best Current Practices.
  9. Questions?

3.1. ASRG Status

- Working on Foundational Documents
- Beginning Analysis of Spam work
- Analyzing some proposals
- Organizing existing anti-spam data
- Working on Consent framework
- Many additional efforts
- More Volunteers Needed!!!

3.2. Foundational Documents.

- Inventory of Problems
  - Lists problems caused by spam and related problems in the current email system
  - Draft being worked on by a subgroup
- Technical Considerations for Spam Control Mechanisms
  - Outlines high-level considerations for anti-spam tools
  - Discusses possible control points in the email infrastructure
  - Written by John Levine, Dave Crocker and Vernon Shryver, all known anti-spam experts, currently in second version
- Requirements for Anti-Spam Proposals
  - Defines common terminology for anti-spam proposals
  - Outlines requirements for anti-spam proposals
  - Draft submitted as an Internet draft

3.3. Analysis and Characterization Subgroup.

- Applies empirical and quantitative methods to problems and issues surrounding spam:
  - Where it comes from
  - What it looks like
  - Ways to eliminate it
- Headed by a professional statistician
- Areas of interest include (but not limited to):
  - Data acquisition and dissemination
  - Research design
  - Measurement & metrics
  - Data analysis and interpretation

3.4. Proposals – DNS-based Authentication Methods.

- RMX/SPF:
  - Seeks to eliminate MAIL FROM forgery
  - Defines a DNS record that needs to be present for every sending SMTP server for each domain used in MAIL FROM
  - Possibly requires a new DNS record type
- DRIP:
  - Seeks to eliminate HELO forgery
  - Defines a DNS record in the domain used in the HELO command containing the IP address of the sending MTA
- Meta Mark:
  - Uses TXT records to marks whether a specific IP address is an MTA or not
- Currently all DNS-based proposals are being combined by a small subgroup into a single proposal
- Significant deployment and anonymity issues need to be analyzed

3.5. Replacing SMTP.

- Several proposals have been submitted to both the IETF and the ASRG
- Seeks to create an alternative email system not backwards compatible with SMTP
- Variations include:
  - Using digital certificates for server-to-server authentication (AMTP)
  - Using DNS records for server-to-server authentication, similar to RMX/SPF/DRIP (MTP)
  - Charging for email – e-postage
  - Digital signatures for every message and a centralized verification system (GIEIS)
  - Pull instead of push approach (IM2000)
  - Alternative “business class” email system with authentication and guaranteed delivery, similar to today’s Express snail mail

3.5. Replacing SMTP – Issues.

- Installed base the size of the Internet is not likely to make such a change anytime soon
- Can take decades to reach that level of adoption, if it ever does.
- Internet comprises a massive number of independent administrations, what is important and feasible to one might be neither to another
- Replacing SMTP with a protocol that allows strangers to send each other mail would not stop spam any more than SMTP-AUTH stopped spam

3.6. Proposals – E-postage.

- Seeks to add cost to existing email systems similar to postal stamps in snail mail
- Various kinds of schemes:
  - Centralized digital money
  - Anonymous digital money (Digicash)
  - Processing power (Hash Cash)
  - Other mechanisms

3.6. E-postage – Issues.

- Lack of an international infrastructure for micro-payments
- Anonymity (Digicash and Hashcash may solve the problem)
- Hijacked Computers and Accounts
- Viruses and worms causing charges to ring up
- Mailing lists suddenly faced with payment choices
- Spammers can steal or buy high performance computers (for Hash Cash)
- Unknown financial, administrative and social costs
- Deployment and scalability issues
- Maybe suitable best for niche applications

3.7. Challenge / Response Internetworking (CRI).

- What is CRI?
  - A protocol for two C/R systems to automatically communicate
  - Saves the trouble of manually clicking on the response
  - Maybe a starting point for a consent token exchange protocol
- Issues with C/R:
  - Adds an authentication layer to SMTP, significant anonymity and deployment issues
  - Problems with disabled people
  - Unknown effect on spammers

3.8. Best Current Practices.

- Defining best practices for:
  - End users
  - Mail administrators
  - Anti-Spam tools vendors
  - Blacklist operators
  - Email senders
  - Consent framework
- Updating existing documents:
  - Existing RFCs 2505, 2635 and 3098

Introduction to the Work of the ASRG and Consent-Based Communications

ASRG Website:

www.irtf.org/asrg

Questions? Comments?

ASRG Mailing List:

asrg@ietf.org