zentraler informatik dienst

universität wien

EHLO

April 2012

Wolfgang Breyha

zentraler informatik dienst

universität wien

SPAM Workshop Teil 1

Vorstellung:

- Wolfgang Breyha
- root am ZID der Universität Wien
- Verantwortlich für Entwicklung und Betrieb des Linux Mailsystems
- Entwicklung und Betrieb von Mailsystemen und Spamfiltern seit 1999

Was kommt nun auf Sie/Dich zu?

Themenüberblick

- Teil 1
  - greylisting
  - SMTP Protokoll/Setup Fehler ausnutzen
  - stottern
  - SPF
  - DNSBLs
  - fake MX
  - exim Beispiele

Grundsätze

- den 100% Spamfilter gibt es nur im Marketing
- Erlaubt ist alles was
    - nicht im Widerspruch zu RFCs steht
    - keine False Positives erzeugt
- keine Angst vor Experimenten
- legitimen Mailverkehr möglichst wenig stören
- “die beste” Antispam-Technik gibt es nicht
    - durch kombinieren ans Ziel
- size matters

size matters

- kleine Server
  - stottern/teergrubing
  - user filter (serverside bayes)
  - MSA auf port 587 wichtig für Trennung von MX
- mehrere MX Server
  - dedizierte MX hosts
  - retry check (whitelist, zB. qmail)
  - große Datenmengen als Entscheidungs-grundlage

Ziele für Spamfilter

- Schutz der eigenen Maschinen
- Regulierung des Mailverkehrs
- Schutz für User
    - direkter Schutz durch blocken
    - indirekter Schutz durch markieren
    - sinnvolle defaults für unerfahrene User

Stufen der Spamabwehr

- Daten erheben
  - Client: IP, hostname, PTR RR, HELO, p0f
  - HELO/EHLO
  - Pipelining
  - Envelope Sender & Recipients
  - RBLs
  - Body Analyse (DCC, Viren-, Spamscan, URIBL...)
- so spät wie sinnvoll eine Entscheidung treffen ob sofort ablehnen, greylisten oder annehmen.

The Good, the Bad and the Ugly

- SMTP erlaubt 2xx, 4xx, 5xx.

  Versuch der Klassifizierung ankommender Verbindungen in
  - legitime Systeme – ungehinderte Zustellung
  - suspekte Systeme – greylisting
  - unerwünschte Systeme - ablehnen

greylisting

- SMTP bietet die Möglichkeit von temporären Fehlercodes (4xx).
- Spammer, insbesondere trojanisierte Rechner in Bot-Netzen, verwenden (noch? wieder?) keine Queues und reagieren daher auf temporäre Fehler wie auf permanente Fehler.
- Greylisting merkt sich Absender, Empfänger und überbringende Host-IP und lehnt Mails für bestimmte Zeit (im Minutenbereich) mit temporärem Fehler ab.
- Normale Mailsysteme probieren es in regelmäßigen Abständen wieder und stellen die Mail nach Ablauf der Embargozeit normal zu.

greylisting

- Vorteil des Zeitgewinns bleibt auch bei wiederkehrenden BOTs
- normalerweise temporäre Fehler nach RCPT TO
    - spießt sich mit sender verification callouts
- oder auch nach DATA
    - führt mit manchen MTAs zu Problemen, welche die Antwort als host Status interpretieren (sendmail)
    - bei selektivem greylisting ideal

Steuerung ankommender Verbindungen

- Verzögerung bei Begrüßung auf MX (zB. 1 Sekunde)
  Chance für Client Fehler zu machen (pipelining)
  Erhebliche Bremse für bots die neu verbinden.
- simple Entscheidungen auf Basis der IP
  - Max Connections/IP
  - Max Connections/Zeitintervall
  - Max Connections/Pool (zB. PTRless hosts)
  - blacklists

Client Setup Analyse

- IP -> PTR -> A -> IP match
- HELO parameter == PTR
- wenn ja => unverdächtig
- wenn nein HELO genauer betrachten
    - check “best of” Liste (dsldevice.lan,...)
    - check kein PTR und kein FQDN
    - check dialup/cable mit hostname patterns
- HELO statt EHLO verwendet?
  ziemlich sicher qmail. Alle bots verwenden EHLO

PTR/HELO auswerten

- Hostnamen abstrahieren
  - Domain bleibt erhalten
  - IP Adressen verschiedener Formate => #
  - unwesentliche Namensteile => !

```
201-67-144-218.bnut3703.dsl.brasiltelecom.net.br =>
#-#-#-#.!.dsl.brasiltelecom.net.br
```

- wesentliche Namensteile in Wortliste definiert
  - böse: adsl, pool, broadband, cable, ...
  - gut: mail, mx, ...
- regex Liste erstellen

p0f, ident

- p0f – passive OS fingerprinting
  - http://lcamtuf.coredump.cx/p0f.shtml
  - analysiert tcp traffic um anhand von Signaturen das Betriebssystem zu erkennen
- ident
  - RFC 1413
  - würde Informationen über den Client liefern
  - kaum aktiviert -> erzeugt hauptsächlich timeouts

zentraler informatik dienst

universität wien

Pause?!

Entscheidungen nach EHLO

- evtl. EHLO ablehnen um
  - Zeit zu gewinnen
  - keine Extensions zuzulassen
- im Unterschied zu Herbst 2009 scheitern neuerdings ettliche bots an abgelehntem EHLO;-) ... 2012 kaum welche.
- delay?

Stottern

- Antworten Byte für Byte schicken.
  - Nicht in MTAs implementiert
  - Aufwand fraglich
- delays vor Antworten
  - in manchen MTAs leicht möglich (exim, ..?)
  - effektiv gegen manche bots (meist 60 Sekunden)

- delays verteilen. Die Summe zählt.
- generell Gefahr von DoS => connectionpools

MAIL FROM: stage

- SPF - RFC 4408
  - DNS TXT Records definieren legitimierte Mailrelays für fragliche Domain
  - $ host -t txt sproing.com
    sproing.com descriptive text "v=spf1 a mx ~all"
  - $ host -t txt utanet.at
    utanet.at descriptive text "v=spf1 ip4:213.90.36.0/25 ... ?all"
  - Nutzen von SPF leider gering
  - Aufwand durch SRS erheblich erhöht
  - SPF result als greylisting trigger

zwischendurch – SRS

- SPF spießt sich mit Forwards
- Beispiel GMX

  gmx.net => "v=spf1 ip4:213.165.64.0/23 ip4:74.208.5.64/26 -all"

  absender@gmx.net via 213.165.64.1 an empfaenger@univie.ac.at

  empfaenger@univie.ac.at hat forward an empfaenger@utanet.at

  utanet.at MX erhält mail via 131.130.3.115 => SPF=fail

- SRS schreibt envelope from um

  absender@gmx.net => SRS0+xxxx=xx=gmx.net=absender@univie.ac.at

  damit auch zuständig für bounce! muß ebenfalls übersetzt werden

  nur umschreiben wenn SPF=pass!

RCPT TO: stage

- alle Entscheidungen die nicht vom Body abhängen möglichst hier treffen.
- Ausnahmen für postmaster, abuse, ...!
- lokale Blocks für Absender, Empfänger, ... bzw. Kombinationen daraus. regex ist besonders praktisch

<mf>from@univie.ac.at<rt>rcpt@gmx.at<fqdn>host.univie.ac.at<ip>2001:62a:4:204::1<helo>host.univie.ac.at

RCPT TO: stage

- HELOs
  - lokale hostnamen
  - IP Adressen
  - localhost
  - blacklist (logs auswerten)
  - DUL blocks (logs auswerten)

RCPT TO: stage

- DNS SERVFAIL
- unqualified hostname && kein PTR
- Bogus IP Networks (sofern nicht auf routern geblockt)
- sender verify
- recipient verify (late bounces vermeiden, callouts)
- RBLs

DNSBLs – RFC 5782

- Informationen im DNS
- IP Adressen im PTR Format

  115.3.130.131.rbl.ispa.at
  5.1.1.0.5.2.0.0.0.0.0.0.0.0.0.0.5.2.0.0.4.0.0.0.a.2.6.0.1.0.0.2.rbl.i...
- Antworten als A Record, meist Zusatz als TXT Record
  - einfachste Variante 127.0.0.2 wenn eingetragen
  - auch Mehrfachantworten oder Bitmuster möglich
- Auch beliebige andere Kriterien möglich
  - example.tld.rbl.ispa.at (sender domain, uribl)
  - sender.x-at-x.domain.tld.wanted.univie.ac.at
- rbldnsd als Server (Ausnahme rbl.ispa.at)

Auswahl von DNSBLs

- rfc-ignorant.org
    - dsn, bogusmx
- spamhaus
    - zen, pbl, dbl
- uceprotect.net
- ix.dnsbl.manitu.net (heise)
- dnsbl.dronebl.org
- multi.surbl.org
- spameatingmonkey.com (mit IPv6 DNSBL)

fake MX

- hosts die
  - die Verbindung komplett ablehnen
  - immer temporäre Fehler bei RCPT TO: liefern
- meist secondary MX, aber auch “wildere” Mischungen
- Vorteile
  - entlastet echte MX hosts
  - führt dumme bots in die Irre
- Nachteile
  - Probleme mit qmail (zB. gmx)
  - Logging++ (Loganalyse, Statistiken)

Vorschau

- Teil 2
  - DCC/Razor/Pyzor/cloudmark/eXpurgate
  - Spamtraps
  - SpamAssassin (Config, Module, Rules)
  - DKIM
  - DMARC
- Teil 3
  - Logfileanalyse
  - Feedbackschleifen
  - Maßnahmen gegen ausgehenden Spam

zentraler informatik dienst

universität wien

Fragen?

April 2012

Wolfgang Breyha