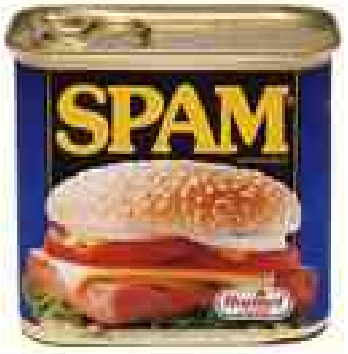
21C3: SPAM Workshop

13:00 Uhr

SPAM- und Zombie-Abwehr bei hohem Mail-Volumen am Beispiel von qmail

Referent: Sirko Zidlewitz <sirko@bytecamp.net>



Die Probleme

* Quellen von SPAM: Zombies ,Kunden großer Provider, Open Relays
* Urheber benutzen Mittelsmänner im Ausland
* Rechtliche Situation weltweit unterschiedlich
* Die eigenen Kunden spammen
* Oft landen Provider auf RBLs
* Dummy-Accounts für Spam-Versand
* Falsche Absender
* HTML-Mails: Werbebotschaft als Bild getarnt
* Adressverifikation: <img
src="http://foo.bar.com/foo/de237718768.gif">
* Text variiert: „asdsk qasa aswwewe", Zitate aus Literatur
* Als NDN getarnt
* Spammer testen selbst mit Spamassassin
* Viren
* Trojaner
* Phishing
* Wordlist Attacks
* DDOS-Attacken

Rechtslage in Deutschland

http://www.recht-im-internet.de

* §7 Gesetz gegen den unlauteren Wettbewerb UWG:
  (1) Unlauter im Sinne von § 3 handelt, wer einen Marktteilnehmer in unzumutbarer Weise belästigt.
  (2) Eine unzumutbare Belästigung ist insbesondere anzunehmen
  ...
  Abs. 3. bei einer Werbung unter Verwendung von automatischen Anrufmaschinen, Faxgeräten oder elektronischer Post, ohne dass eine Einwilligung der Adressaten vorliegt;
* unzulässige und unterlassungsfähige Belästigung im Sinne von §§ 823, 1004 BGB
* Bei beruflich genutzen Accounts: Eingriff in den eingerichteten und ausgeübten Gewerbebetrieb gemäß § 823 Abs. 1 BGB
* Bei privat genutzten Accounts: Eingriff das allgemeine Persönlichkeitsrecht des Betroffenen gemäß § 823 Abs. 1 BGB

Die Lösungsansätze

* qmail-scanner
* Spamassassin/spamd und seine Module
* RBL
* DCC
* SPF
* Bayes Filter
* Syntax- und MX-Checks von Email-Adressen
* ClamAV
* Der SPAMCONTROL-Patch
* Vor- und Nachteile von greylisting
* qgreylistrbl, Erfahrungen, zukünftige Entwicklung
* HELO/EHLO-Spoofing versus RFC 821 in der Praxis
* Die eigene Verantwortung: Tarpitting, SMTP-Auth

Die eigene Verantwortung als Provider

* Kunden werden spammen
* Leider regelmäßig
* Einige der Spammer werden mit Klage drohen.

Gegen-Maßnahmen:

* Rechtsschutz-Versicherung
* AGB
* Tarpitting
* SMTP-Auth
* Reverse Split-Horizon

„qmail big picture“ für eingehende E-Mails

softlimit	(limitiert Verbindungen, Speicher, etc.)
↓	
tcpserver	(setzt $TCPREMOTEHOST u.a.)
↓	
qmail-smtpd	($RELAYCLIENT, rcpthosts)
↓	
qmail-queue	(triggert qmail-send)
↓	
qmail-send	(recipient in locals?)
↓	
qmail-lspawn	(checkt account, uid, übereignet
↓	die Datei, startet qmail-local)
qmail-local	(testet nach .qmail, Zustellung)

Wir ergänzen das „qmail big picture“

softlimit
↓
tcpserver
↓
qgreylist / qgreylistrbl / rblsmtpd
↓
qmail-smtpd mit SPAMCONTROL / isp.patch
↓
qmail-scanner-queue.pl ↔ Virenscan, Spamassassin
↓
qmail-queue
↓
qmail-send
↓
qmail-lspawn
↓
qmail-local .qmail-Datei → maildrop /mailfilter / Virenscan, Spamassassin

qmail-scanner-queue.pl

http://qmail-scanner.sourceforge.net/

* In Perl geschrieben
* benötigt den QMAILQUEUE-Patch
* wird mit der Environment-Variable `$QMAILQUEUE` gesetzt:
* unterschiedliche qmail-scanner-queue.pl für unterschiedliche Remote-Hosts definierbar (tcp.smtp)
* arbeitet mit fast allen Unix-Commandline Virus-Scannern
* arbeitet mit Spamassassin zusammen
* besitzt internen Virus-Scanner
* kann rekursiv Archive entpacken
* kann noch 100.000 Sachen mehr
* Konfiguration erfolgt bei Installation und danach durch Setzen von

  Variablen im Programm selbst.
* Nach Abarbeitung wird die Nachricht durch Aufruf von qmail-queue zurück ins normale qmail-System geliefert

Spamassassin/spamd und seine Module

http://spamassassin.apache.org/

* Modularer Spam-Check
* Client-Server-Prinzip
* Client in C geschrieben, Server in Perl
* Server-Master-Prozess startet viele Childs
* Client und Server können auf unterschiedlichen Maschinen laufen
* Client bekommt Email an STDIN und gibt sie an STDOUT zurück
* Besonderheit: Hunderte von Tests „benoten“ E-Mail positiv oder negativ
* Summe aller Noten, ergbibt Gesamt-Wertung, z.B. `10.4`
* Einsatz an unterschiedlichen Stellen im Big-Picture möglich:
  über `qmail-scanner-queue.pl` oder
  durch `.qmail`-Datei
* verfügbare Module sind u.a. RBL-Abfragen, SPF, Bayes Filter, DCC, Pyzor

DCC – Distributed Checksum Clearinghouse

http://www.rhyolite.com/anti-spam/dcc/

* Collaborative Filtering Network
* Ermittelt unscharfe Prüfsummen (Fuzzy Checksums) von Headern und Body
* Verteilte Datenbank zählt, wie oft Email weltweit empfangen wurde
* Mehr als 150 Mio. Prüfsummen an werktagen gezählt
* Jeder der rund 250 Server hat im Schnitt 3 Peers
* Spam wird erkannt, indem Vorkommen der Prüfsumme gezählt wird
* Eigene, isolierte Server machen kaum Sinn
* Kaum Traffic
* UDP
* Client sucht sich selbstständig schnellsten DCC-Server
* Einsatz an verschiedenen Stellen im Big-Picture möglich
* Probleme: „fuhuh dfdfdf bndsskd“

RBL – Realtime Blackhole Lists

* Form von einer „schwarzen Liste“
* Datenbank von IP-Adressen von denen Spam verschickt wurde
* wurde über DNS realisiert
* verschiedene Listen für Open Relays und Trojaner
* bekannte RBL-Betreiber sind u.a. spamcop.net, spamhaus.org
* RBL-Anfragen machen an mehreren Stellen des Big-Picture Sinn
* Problem: Provider landen auch regelmäßig auf RBLs

Neu: URIDNSBL

* statt Absender-IPs werden in Emails enthaltene URLs geprüft
* auch über DNS realisiert
* Anbieter einer solchen Liste z.b. SBL von spamhaus.org
* in Spamassassin ab Version 3.0

Bayes Filter

* basiert auf Bayes-Theorem
* Datenbank von Wörtern
* Wahrscheinlichkeitswert für jedes Wort
* Hohe Trefferrate
* muß immer trainiert werden
* in Spamassassin integriert
* Kann global oder vom User selbst trainiert werden (effektiver)
* Probleme:
    - Zitate aus Literatur
    - Werbebotschaft selbst wird in Bildern untergebracht

Bayes Filter - Beispiele

0.034	2654	8877	1104249659	Wochen
0.079	12936	17465	1104249576	unserer
0.018	1781	11505	1104249659	paar
0.040	6851	18925	1104249413	daher
0.958	1	0	1104249440	HX-Mailman-Approved-At:Tue
1.000	761	0	1104249398	HX-Message-Info:sk:RNDUCCH
0.000	22	7479	1104249566	HX-No-Archive:yes
0.003	360	13659	1104249617	erworben
0.988	76267	108	1104249556	Photoshop

SPF – Sender Policy Framework

http://spf.pobox.com/

* TXT-Einträge in den DNS einer Domain
* offizielle Einführung war 1. Oktober 2004
* In Spamassassin integriert
* bei den meisten Providern nur für Rating bei Spamassassin benutzt
* Einsatz an mehreren Stellen des Big-Picture möglich
* auf der Homepage befindet sich ein Wizard für die DNS-Einträge
* Probleme: Weiterleitungen
  Pflege der SPF-Einträge
  Spammer haben oft korrekte SPF-Einträge

Test nach SPF-Einträgen

```
>host -tTXT bytecamp.net
bytecamp.net text "v=spf1 ip4:212.204.60.0/24"

>host -tTXT gmx.de
gmx.de text "v=spf1 ip4:213.165.64.0/23 ?all"

>host -tTXT web.de
>
```

Wer ein `?all` in den Eintrag schreibt, vertraut auf ALLE Hosts, die mit dem
angegebenen Domain-Namen enden.

EHLO/HELO-Spoofing

RFC-821 Section 3.5:

The sender-SMTP MUST ensure that the <domain> parameter in a HELO command is a valid principal host domain name for the client host. As a result, the receiver-SMTP will not have to perform MX resolution on this name in order to validate the HELO parameter.

The HELO receiver MAY verify that the HELO parameter really corresponds to the IP address of the sender. However, the receiver MUST NOT refuse to accept a message, even if the sender's HELO command fails verification.

EHLO/HELO-Strings in freier Wildbahn

Router

localhost

Netgear

work

whitehouse#gov

Muhammed Ali

@!§3

Die IP oder der Name des Mailservers, der die E-Mail gerade annimmt.

greylisting

http://projects.puremagic.com/greylisting/links.html

„450 Temporary local problem. Try later.“

* beim ersten Verbindungsversuch temporäre Fehlermeldung
* IP der Gegenstelle kommt auf „Warteliste“
* zweiter Verbindungsaufbau wird akzeptiert
* Auch spätere Verbindungsaufbauten werden akzeptiert
* Sehr wirksam bei Zombies
* Weniger NDNs an falsch angegebene Absender
* Probleme:
  - Einige Mails kommen verspätet an
  - Bayes-Filter wird nicht so gut trainiert
  - Einige Callback/Sender-Verify-Implementationen

qgreylistrbl

http://www.datenklause.de/de/software/qgreylistrbl.html

* Greylisting nur für Dialin-Rechner und RBL-gelistete Hosts
* Replacement für rblsmtpd
* Überprüft Syntax der Sender/Empfängeradressen
* Limit für Anzahl der Empfänger einstellbar
* Test fuer EHLO/HELO Spoofing (Syntax, Mailservername/IP)
* Test auf NDNs an mehrere Empfänger
* Fazit nach einem Jahr:
    -80% weniger Spam/Viren
    -Viele Tests der Version 0.4a wurden durch SPAMCONTROL obsolet
    -Gesamt-Performance stieg

qgreylistrbl - Ausblick

* Im Test gerade Version, die greylist-Einträge nur bei HAM vornimmt
* Ergebnisse von Spamassassin und ClamAV werden ausgewertet
* Nächster Release wird wg. SPAMCONTROL weniger Tests vornehmen
* Test auf Existenz der Accounts wird entfallen
* Nur eine Version, die sowohl mit oder ohne vpopmail laufen wird

Der SPAMCONTROL-Patch

http://www.fehcom.de/qmail/spamcontrol.html

* Sammelpatch für qmail, hauptsächlich für qmail-smtpd
* STMP-Auth für qmail-remote
* (DNS Lookup for the `HELO/EHLO` greeting)
* DNS Lookup für den Domain-Part der Absender-Adresse
* Tarpitting
* Reverse Split-Horizon: `Mail-From:` muß lokale Domain sein
* Limit für `Rcpt To:` pro SMTP-Session
* Smart Rejection bei zu vielen falschen Empfängern
* Doublebouncetrim, Länge von Bounces wird begrenzt
* Nur ein Empfänger pro NDN
* badmimetypes, badloadertypes
* QMAILQUEUE-Patch

Ist das alles wirklich nötig?

Was passiert, wenn man Absender nicht mehr fälschen kann?

Was passiert mit Newsgroups? Selbsthilfegruppen...

Beispiel-Installation

softlimit
↓
tcpserver
↓
qgreylistrbl
↓
qmail-smtpd mit SPAMCONTROL
↓
qmail-scanner-queue.pl ↔ ClamAV, Spamassassin <DCC,Bayes,SPF...>
↓
qmail-queue
↓
qmail-send
↓
qmail-lspawn
↓
qmail-local .qmail → |preline /usr/local/bin/maildrop /etc/mailfilter

Der praktische Teil – benötigte Pakete

* qmail mit QMAILQEUE-Patch
* daemontools
* ucspi-tcp
* qmail-scanner
* maildrop
* checkpassword
* clamav
* spamassassin
* dcc-dccd