Spam e Virus

Alberto Cammozzo mmzz -at- pluto.it
Mauro Luzi mauro -at- pluto.it

5 maggio 2004

serate a tema PLUTO Padova

Premessa

- Evidentemente SPAM e Virus funzionano e servono a qualcosa
- Sicuramente servono all'industria antiSPAM e antivirus che non ha interesse a che spariscano
- Ma serve anche a chi invia i virus (soddisfazione personale?)
- E chi invia SPAM: evidentemente lo spam viene letto da qualcuno e rende...
- ...Mah!!!

Soluzione presentata: Postfix+procmail+spamassassin

- Conviene beccare spam e virus dall'MTA con sistemi furbi prima di caricare le CPU con antispam e antivirus.
- **Mail Transfer Agent**: Postfix: modulare, sicuro, affidabile, veloce, robusto, manutenibile e configurabile abbastanza facilmente: sta quasi tutto in `/etc/postfix/main.cf`
- **Mail delivery agent**: Procmail: consente all'utente di impostare i **propri** filtri in `.procmailrc`

Postfix: che controlli si possono fare?

- il Modulo Smtpd consente il controllo di:
  - **Sender address** checks: sull'indirizzo mittente
  - **Recipient address** checks: sull'indirizzo destinatario
  - **Helo** checks: sulla dichiarazione di HELO
  - **Header** checks: controlli sul contenuto dell'header
- Modulo Cleanup
  - **Body** checks: sul contenuto del corpo del *messaggio*

Semplici controlli generali

/etc/postfix/main.cf

```
smtpd_sender_restrictions =
    reject_non_fqdn_sender,
    reject_invalid_hostname,
    reject_unknown_sender_domain

smtpd_helo_required = yes
smtpd_helo_restrictions =
    permit_mynetworks,
    reject_invalid_hostname

smtpd_recipient_restrictions =
    permit_mynetworks,
    reject_unauth_destination,

reject_unknown_recipient_domain
```

Restrizioni sul mittente

Restrizioni sul nome di HELO

Restrizioni sul destinatario

Altri controlli piu' cattivelli

```
header_checks =
   regexp:/etc/postfix/header_checks
```

Controllo sugli header

```
#ATTACHMENT
/.+name=.+\.(hta|com|pif|vbs|vbe|js|jse|exe|bat|cmd|vxd|scr|shm
   |eml|hlp|spl|swf|shb|vba|dll|reg|ocx|wsf|wsh|lnk)"?$/ REJECT
#SUBJECT
/^Subject: retire early.*$/ REJECT
/^Subject: Nuovo Documento Microsoft Word (3).*$/ REJECT
/^Subject: Tami thought you'd be interested in this!!!!.*$/ REJECT
/^Subject: Got Debt? [7xkeh].*/ REJECT
```

- Allegati eseguibili o affini
- Subject noti
- mittenti noti

Altri controlli 'furbi'

```
body_checks =
   regexp:/etc/postfix/body_checks
```

Controllo sul contenuto del messaggio

```
/^TV[nopqr]....[AB]..A.A....*AAAA...*AAAA/  REJECT
/^M35[GHIJK].`..`..*`````/ REJECT
/^UEsDBAoAAAAAA/ REJECT
```

- Gli allegati base64 iniziano in un dato modo se il file che codificano inizia con una data stringa; e gli eseguibili (anche Windows) iniziano sempre nello stesso modo (prima e seconda riga)
- Volendo si può fare lo stesso anche con uuencode, ma non lo usa nessuno
- Anche se sono zippati (terza riga)
- Credit: *Hobbit* <hobbit-at-avian-dot-orb>

SPAM (UCE)

- Programmi antispam: euristici, probabilistici: i pesi dei filtri possono essere adattati (addestrati) in base allo spam effettivamente ricevuto.
- Devono fare un'analisi del contenuto di ogni email: pesano molto sulla CPU
- E' bene sceglierne uno indipendente dal Mail User Agent
- Spamassassin: uno dei più diffusi: www.spamassassin.org.

Esempio di SPAMASSASSIN

```
Spam detection software, running on the system "stavromula.stat.unipd.it", has
identified this incoming email as possible spam.  The original message
has been attached to this so you can view it (if it isn't spam) or block
similar future email.  If you have any questions, see
the administrator of that system for details.

Content preview:  Reed,/ Save 95% for all V+iagr+a/C+iali+s/L+evitr+a.
  http://www.EEEPPP.BIZ/ES001/?affiliate_id#3635&campaign_id@4 [...]

Content analysis details:   (12.2 points, 5.0 required)

 pts rule name              description
---- ---------------------- --------------------------------------------------
 5.4 BAYES_99               BODY: Bayesian spam probability is 99 to 100%
                            [score: 1.0000]
 0.1 BIZ_TLD                URI: Contains a URL in the BIZ top-level domain
 4.1 FORGED_RCVD_NET_HELO   Host HELO'd using the wrong IP network
 0.1 RCVD_IN_RFCI           RBL: Sent via a relay in ipwhois.rfc-ignorant.org
                            [Inaccurate or missing WHOIS data]
 0.7 RCVD_IN_DSBL           RBL: Received via a relay in list.dsbl.org
                            [<http://dsbl.org/listing?ip=211.61.82.23>]
 0.3 DNS_FROM_RFCI_DSN      RBL: From: sender listed in dsn.rfc-ignorant.org
 1.5 RCVD_IN_BL_SPAMCOP_NET RBL: Received via a relay in bl.spamcop.net
                [Blocked - see <http://www.spamcop.net/bl.shtml?211.61.82.23>]
...
```

Procmail + Spamassassin

$HOME/.procmailrc

```
:0fw: spamassassin.lock
| /usr/bin/spamassassin

:0:
* ^X-Spam-Status: Yes
Mail/SPAM
```

- La prima riga invoca spamassassin
- La seconda archivia automaticamente la mail etichettata come SPAM da spamassassin nel file $HOME/Mail/SPAM

• L'archiviazione automatica può essere attivata quanto non si hanno piu' *falsi positivi* (mail erroneamente etichettata come SPAM)
• Lo può fare l'utente stesso
• L'amministratore deve agire cautamente: se attiva l'archiviazione senza il consenso dell'utente può commettere il reato di *intercettazione, impedimento o interruzione illecita di comunicazioni informatiche o telematiche'* (617 quater C.P.: a 1 a 4 anni)

Calibrare Spamassassin

- L'utente può addestrare l'assassino secondo i propri gusti
- Basta salvare la mail SPAM non intercettata in una mailbox [SPAM] , e quella non SPAM in un'altra [HAM]

```
sa-learn --showdots --mbox --spam [SPAM]
sa-learn --showdots --mbox --ham [HAM]
```

La seconda operazione (quella sulla mail buona, con '--ham') va fatta solo se spamcop sbaglia e genera falsi positivi

Configurare Spamassassin

.spamassassin/user_prefs

```
required_hits              5
whitelist_from             amico@bravagente.orgg
```

- **sensibilita':** al posto del 5 può essere messo un numero più alto per renderlo meno suscettibile, o più basso per fare in modo che uccida per molto meno.
- **whitelist**: inserire indirizzi di corrispondenti dai quali sicuramente non riceviamo SPAM.

Anti Virus

- **Centralizzato**: Amavis (antivirus daemon) + un programma di antivirus a scelta (sotto Linux):
  - Pesante sulla CPU: ogni allegato deve essere controllato.
  - Soggetto a Denial Of Service (?)
- **Periferico**: l'antivirus solo dove serve: sui PC con Windows
  - Non carica il server con controlli inutili su mail che verra' letta da Linux.
  - L'antivirus su Windows serve comunque anche per virus e worm di origine non-email.

Una piccola analisi s'ti[s]tica

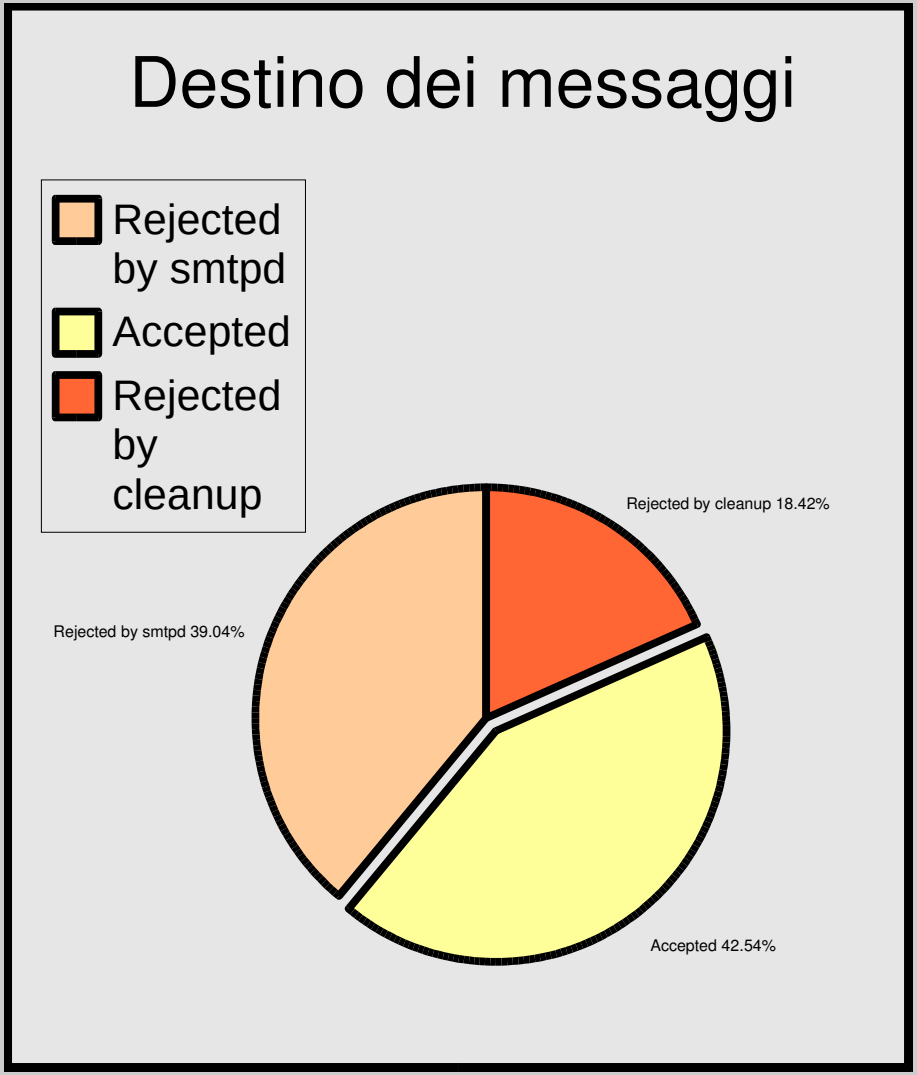
- Postfix su un server email con circa 250 utenti
  - uso intenso della mail
  - alcuni indirizzi pesantemente diffusi da anni
- Log di postfix della settimana dal 25 Aprile al 2 Maggio 2004:
  - **34221** connessioni di in ingresso
  - 3,4 contatti al minuto.
  - (Poca roba).

Virus e spam intercettati da Postfix

Motivo del rifiuto della mail	Occorrenze	% nella categoria	% sul totale
550 User unknown in local recipients table	12829	**96.03%**	37.49%
554 Relay access denied	242	**1.81%**	0.71%
504 Sender address rejected: need fully-qualified address	228	**1.71%**	0.67%
501 Helo command rejected; Invalid name	56	**0.42%**	0.16%
552 message exceeds fixed limit	2	**0.01%**	0.01%
ETRN from unknown	2	**0.01%**	0.01%
Postfix smtpd reject	13359	100.00%	**39.04%**
regexp ^TV (exec attach)	5029	**79.77%**	14.70%
regexp ^UE (exec attach)	1181	**18.73%**	3.45%
other attachments and header checks (3)	94	**1.49%**	0.27%
Postfix cleanup reject	6304	100.00%	**18.42%**
Totale Rejected	19663		57.46%
Totale **Accepted**	14558		42.54%
Totale connessioni da 14723 client diversi.	34221		100.00%

segnale/rumore = ???

- il 37% dei messaggi in ingresso è per **destinatari inesistenti**,
- un altro 2% è scartato per vari altri motivi ,
- almeno il 18% del traffico in ingresso è costituito da [virus in] **allegati eseguibili**;
- solo il 42% restante viene consegnato,
- di questo probabilmente la metà è SPAM.



Casi particolari

- Mail con from=<>
  - richiesto da RFC 822 per i double-bounce: from=<> è un mittente valido
- Mail da host non FQDN (host con nome e non solo IP)
  - consentito
  - biasimato
  - considerato sospetto

Mail con from=<>

Motivo del rifiuto di messaggi con `from=<>`	Occorrenz	%su mail from <>	% sul totale
Smtpd reject: 550 User unknown in local recipients table	10778	**99.95%**	31.50%
regexp ^TV (exec attach)	4	**0.04%**	0.01%
regexp ^UE (exec attach)	1	**0.01%**	0.00%
Total Rejected (from=<>)	10783	100.00%	**31.51%**
Totale accepted (from=<>)	4		**0.01%**
Totale (from=<>)	10787		**31.52%**
Totale connessioni	34221		

- più del 30% dei messaggi in ingresso sono con con from=<>,
- il 99,92 % dei quali sono per utenti localmente inesistenti.

Mail con sender host unknown (non fqdn)

Motivo del rifiuto da client unknown	Occorrenze	%su unknown	% sul totale
550 User unknown in local recipients table	1863	**25.82%**	5.44%
554 Relay access denied	32	**0.44%**	0.09%
504 Sender address rejected: need fully-qualified addres	78	**1.08%**	0.23%
501 Helo command rejected; Invalid name	34	**0.47%**	0.10%
Postfix smtpd reject	2007	27.81%	5.86%
regexp ^TV (exec attach)	883	**12.24%**	2.58%
regexp ^UE (exec attach)	246	**3.41%**	0.72%
other attachments and header checks (3)	16	**0.22%**	0.05%
Postfix cleanup reject	1145	15.87%	3.35%
Totale Rejected from unknown host	3152	43.68%	**9.21%**
Totale accepted from unknown host	4064	56.32%	**11.88%**
Totale connect from unknown	7216	100.00%	**21.09%**

- Il 20% dei messaggi vengono da host senza un nome registrato,
- solo il 5% dei quali per destinatari locali non esistenti,
- e solo il 3% dei quali contiene sicuramente virus,
- il resto viene da host con regolare FQDN!

Appendice filosofica

- A cosa serve tutto questo SPAM e Virus?
- Soluzione autoreferente:
- lo SPAM serve a addestrare i motori antispam, cioe' a discriminare segnale dal rumore
  - gli antispam filtrano efficacemente senza capire nulla del messaggio che trattano... Straordinario!
- I virus servono a addestrare il sistema immunitario della rete: “che si faccia le difese”.

Grazie