# 垃圾信戰爭

2009/09/02 14:00~17:00

王英鼎

# 大綱

- 垃圾信戰爭歷史
- 過濾系統介紹
- Amavis + SpamAssassin + ClamAV
- Dspam的聯結方式
- 大串聯!!
- Dspam的過濾原理與訓練方式

# 垃圾信戰爭歷史

- 網路消費協會(http://www.nca.org.tw/)的調查
  - 從2001年的7%，上昇到2003年的50%以上，且仍舊繼續不斷攀升
- 郵件及網路安全設備廠商Barracuda Networks「2007年度垃圾郵件調查報告」
  - 2007年垃圾郵件量已達到總郵件量的90~95%

# 垃圾信戰爭歷史

- SpamAssassin從1997年開始發展到現在
- Amavis + AntiSpam + AntiVirus
- greylist, RBL(Real-time Blackhole List)
- DKIM, SenderID
- Dspam
- 各種技術都有其可能誤擋正常信的狀況
- Botnet發送廣告信
- 信件標題與內容隨機插入一些符號以使得無法以規則判別
- 根據來源或內容判斷是否為廣告信變得非常困難
- 甚至是規則設越多, 垃圾信沒減少, 誤判攔截卻變嚴重

# 過濾系統介紹

- DNS反查
  - 拒絕沒有反查的連線來源
  - 對內: 可避免盜用IP的發信端連線, 造成追蹤困難
  - 對外: 沒有反查的來源可視為該IP被盜用
  - 缺點: 對岸的網路尚未齊備, 可能會誤擋對岸公司發出的信件
    - 例外允許即可
  - 可設定於hosts.allow或postfix的main.cf中

# 過濾系統介紹

- 灰名單機制(greylist)
  - 第一次連線
    - 回應450 暫時性錯誤, 要求五分鐘以後再送一次
    - 紀錄其IP與連線時間點
  - 第二次連線
    - 距離上一次連線 少於24小時?
      - 是: 距離第一次連線 超過5分鐘?
        - 是:接受送信 否: 拒絕, 重設連線時間
      - 否: 當作 第一次連線 處理

# 過濾系統介紹

- 灰名單機制(greylist)
  - 正常的SMTP server在收到4XX的錯誤代碼,會在設定的延遲時間後重送信件
  - 一般預設為30分鐘
    - 有效期最短不能低於30分鐘
    - 要求的延遲不能高於30分鐘

# 過濾系統介紹

- 灰名單機制(greylist)
  - 有效阻止大部分的botnet與發廣告信軟體的連線
    - Botnet一般不會做到MTA應有的queue信與重送功能
    - 發廣告信軟體僅會做一次性連線, 也不會有重送功能
  - 對於以下狀況無效
    - 透過正常的MTA server發的廣告信
    - Botnet或發廣告信軟體一日內發送多次廣告信
      - 縮短有效期, 加長延遲時間, 可減少此問題
  - 不能用在提供校內relay的SMTP server
    - 不能要求使用者5分鐘之後重送信件
  - 發信端使用分散式MTA -> 總是被擋
    - 另設定白名單直接通過

# 過濾系統介紹

- RBL(Real-time Blackhole List)
  - 參與者可詢問RBL server目前的連線來源是否要拒絕連線
  - 清單由RBL中心控管
  - 可同時詢問多個RBL server
  - 優點: 由他人幫您建立黑名單, 不需自行管理
  - 缺點: 很可能意外擋了您要收信的來源端
  - 現今botnet spam盛行, 此清單效用已低落很多

# 過濾系統介紹

- DKIM(DomainKeys Identified Mail), SenderID
  - 前者由yahoo與cisco研發而成
  - 後者由微軟的hotmail系統研發而成
  - 兩者做法大同小異
    - 在DNS設定一個公鑰提供目的端MTA可下載
    - 在本地MTA對發送的信使用私鑰產生hash key加載於信件header中
    - 目的端可用公鑰確認hash key是否正確, 正確即為此封信確實由本地MTA發送的信件

# 過濾系統介紹

- DKIM, SenderID
  - 優點: 有效作為判別廣告信的權重之一
  - 缺點: 仍無法確認發信者的身分
    - 發信者身分僅能用PGP之類的Email電子簽名技術做判別
  - 金鑰由發信MTA自行產生, 較無公信力, 發廣告信的MTA也可加載同樣的資訊

# 過濾系統介紹

- SpamAssassin
  - 老字號過濾軟體
  - Perl語言, 可編譯成中介碼加速運作速度
  - 結合多種判別方式並做加權
    - Header完整性, 圖文比例, 文字編碼錯誤率, 特定字串, RBL, 貝式演算, razor線上spam資料庫, 連線路徑正反查
  - 僅為一獨立程式, 需要milter輔助連結到SMTP服務軟體
    - SA-milter, procmail

# 過濾系統介紹

- SpamAssassin
  - 效率不佳
    - Perl語言運作效率較差, 佔用記憶體較多
    - 需連線到razor server取得是否為廣告信的資訊
    - 需詢問DNS確認正反查資訊
      - 建構本機DNS cache service
  - 能阻擋約60%~70%的廣告信
    - 緩慢下降中
    - 廣告信手法越來越新穎

# 過濾系統介紹

- Amavis + AntiSpam + AntiVirus
  - Amavis
  - 偽裝smtp service
  - 基本檢查功能
    - 附件性質過濾, 如執行檔等高危險檔案
    - 包含解壓縮功能, 附件為無加密壓縮檔會解壓縮再確認是否攔截
  - milter功能
    - 可連結廣告信掃描與病毒掃描程式
    - 例如連結 SpamAssassin 與 ClamAV

# 過濾系統介紹

- Dspam
  - 由Jonathan Zdziarski於2003年到2007年緩慢地研發而成
  - 2007年5月 賣給了Sensory Networks
  - 2008年12月 Sensory Networks成立了DSPAM-community 子公司
  - 2009年1月 將DSPAM釋放給DSPAM-community公司維護。

# 過濾系統介紹

- Dspam
  - 優點
    - 開放原始碼
    - 過濾內容方式
    - 過濾條件完全個人化
    - 程式效率高(以C語言撰寫)
    - 判別精準(經訓練後可達九成九的判別率)
  - 缺點
    - 設定不易!!
    - 沒有訓練 = 沒有過濾效果

# Amavis+SpamAssassin+ClamAV

- Amavis
  - 阻擋不良的附件
  - 解壓縮附件
  - 連結antispam 與 antivirus
- SpamAssassin
  - 基本過濾, 阻擋70%廣告信
- ClamAV
  - 過濾病毒與釣魚信的惡意script

# Smtp傳輸

- Hello
  - 打招呼, 表明自己的來源
- Mail from: <Email_address>
  - 寄件者
- Rcpt to: <Email_address>
  - 收件者
- Data
  - 最後一行一個半形句點作為結束
  - 包含header, subject, content等等資訊

# Postfix的檢查點

- Hello
  - 對於helo指令所送的ip/dn做檢查
  - 非標準
- Client
  - 連線來源端, 可做access管控, 反查檢查, FQDN檢查等等
- Sender
  - 可拒絕特定的Email來源, 此為信件內顯示的寄件者, 不代表真正的送信端, 可假造之
- Recipient
  - 收信者, 目的Email, 可做分流或管控等, 若為最終MTA, 不應收不該收的hostname
- Header
  - 信件內看不到的前面部分, 會帶有信件傳遞路徑紀錄與各種檢查結果標示
- Body
  - 信件內容, 除內文外, 包括標題與附件.

# Postfix+amavis

- master.cf
  - smtp-amavis unix - - n - **10** lmtp
    - -o soft_bounce=yes
  - 0.0.0.0:10025 inet n - n - - smtpd
    - -o mynetworks=127.0.0.0/8
    - -o content_filter=
    - -o local_recipient_maps=
    - -o relay_recipient_maps=
    - -o smtpd_restriction_classes=
    - -o smtpd_client_restrictions=
    - -o smtpd_helo_restrictions=
    - -o smtpd_sender_restrictions=
    - -o smtpd_recipient_restrictions=permit_mynetworks,reject
    - -o strict_rfc821_envelopes=yes
    - -o smtpd_error_sleep_time=0
    - -o receive_override_options=no_unknown_recipient_checks, no_header_body_checks

# Postfix+amavis

- main.cf
  - content_filter = smtp-amavis:[127.0.0.1]:10024
    - 不管什麼信件都丟給amavis檢查
      - Smtp傳送到localhost的port 10024
      - Port 10024由amavis開啟
  - smtpd_recipient_restrictions = check_recipient_access pcre:<pathfile>
    - 在recipient檢查點決定是否送給amavis檢查

# Postfix+amavis

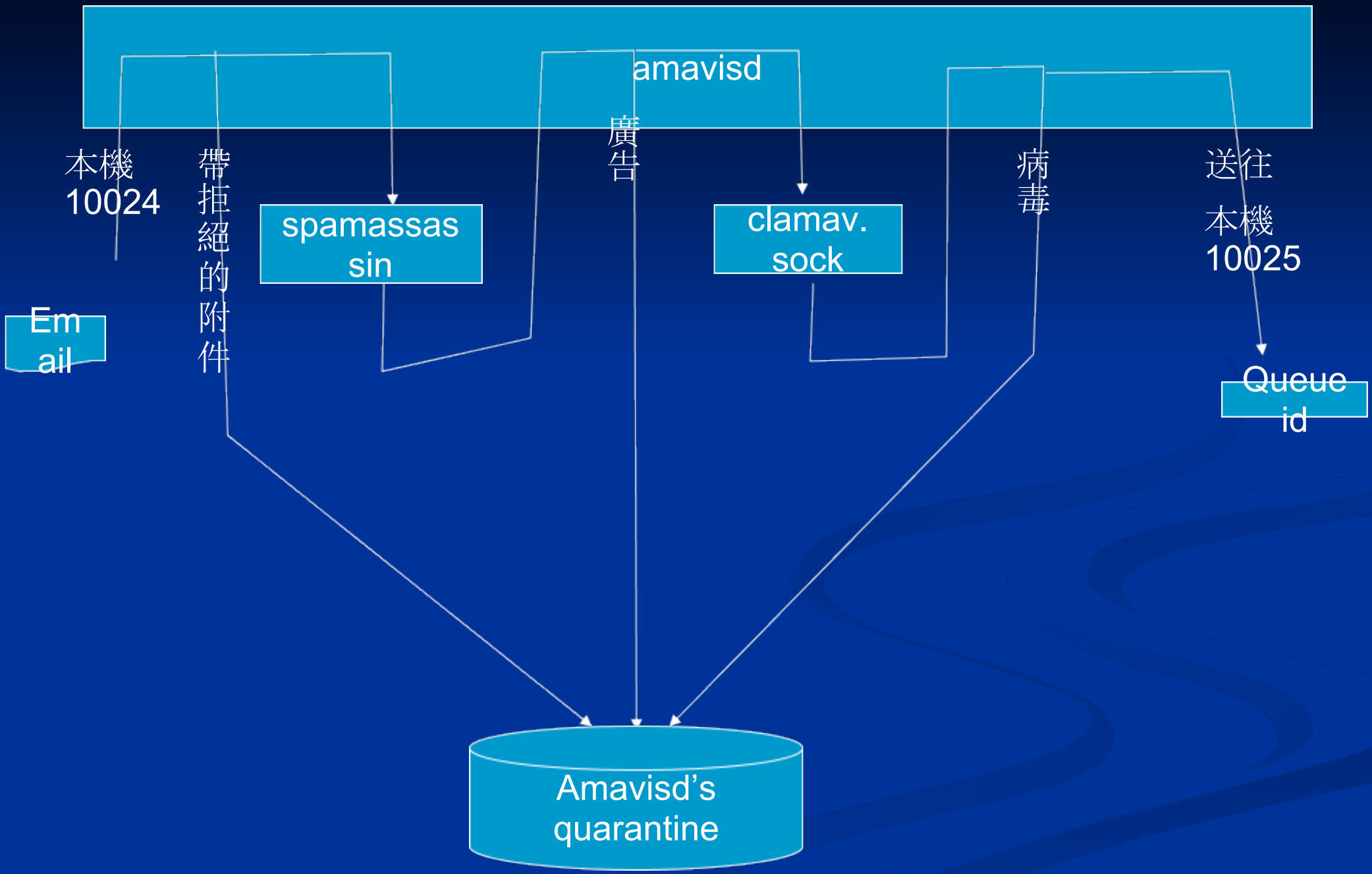
- pcre file sample
  - /spam@cc.nctu.edu.tw/ OK
  - /notspam@cc.nctu.edu.tw/ OK
  - /@cc.nctu.edu.tw/ FILTER smtp-amavis:[127.0.0.1]:10024
  - /./ REJECT

# Amavis+SA

- amavisd.conf
  - $max_servers = 10;
  - $sa_tag_level_deflt = 2.0;
    - # add spam info headers if at, or above that level
  - $sa_tag2_level_deflt = 6.2;
    - # add 'spam detected' headers at that level
  - $sa_kill_level_deflt = 7;
    - # triggers spam evasive actions (e.g. blocks mail)
  - $sa_dsn_cutoff_level = 40;
    - # spam level beyond which a DSN is not sent
  - $sa_crediblefrom_dsn_cutoff_level = 38;
    - # likewise, but for a likely valid From
  - $final_spam_destiny = D_DISCARD;
    - 原本為 D_BOUNCE;

# Amavis+ClamAV

- 可跟多種知名掃毒軟體組合
  - ClamAV, Kaspersky, Avira AntiVir, Symantec, F-Secure, avast!, NOD32, Panda
  - 請參閱amavisd.conf中, 對應的設定區
- ClamAV部分注意事項
  - vscan group需加入clamav
  - socket file需配合clamav.conf中的設定修改
    - /var/run/clamav/clamd.sock 或 /var/run/clamav/clamd

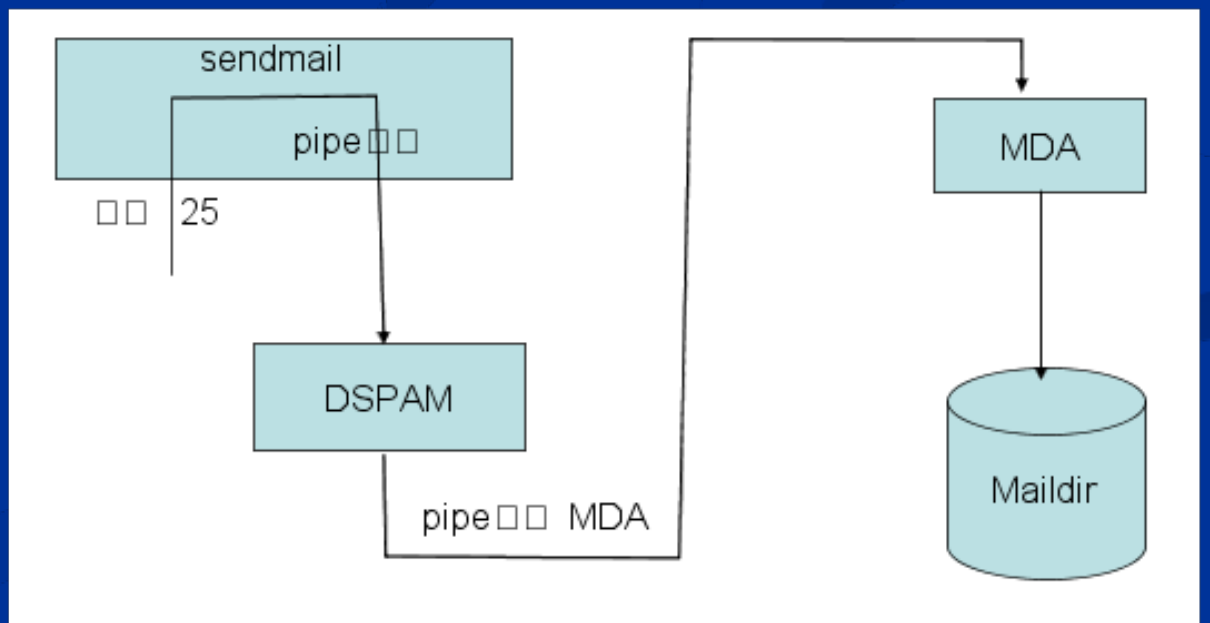

amavisd
本機
10024
帶拒絕的附件
spamassas
sin
廣告
clamav.
sock
病毒
送往
本機
10025
Em
ail
Queue
id
Amavisd's
quarantine

# Dspam的聯結方式

- 與sendmail連結
  - mc file (hostname.mc)
    - define(`LOCAL_MAILER_PATH', `/usr/local/bin/dspam')
    - define(`LOCAL_MAILER_ARGS', `dspam "--deliver=innocent" --user $u -d %u ')
  - dspam.conf
    - TrustedDeliveryAgent "/usr/local/bin/maildrop“
      - 或 “/usr/libexec/mail.local“ 若用mbox格式
  - maildrop.rc
    - DEFAULT="$HOME/Maildir“
  - 使用maildrop作為MDA的話, 請記得編譯時要加入
    - TRUSTED_USERS=“root mail www”

# Dspam的聯結方式

- 與sendmail連結
  - aliases設定
    - spam: "|/usr/local/bin/dspam --user root --class=spam --source=error"
    - notspam: "|/usr/local/bin/dspam --user root --class=innocent --source=error"



# Dspam的聯結方式

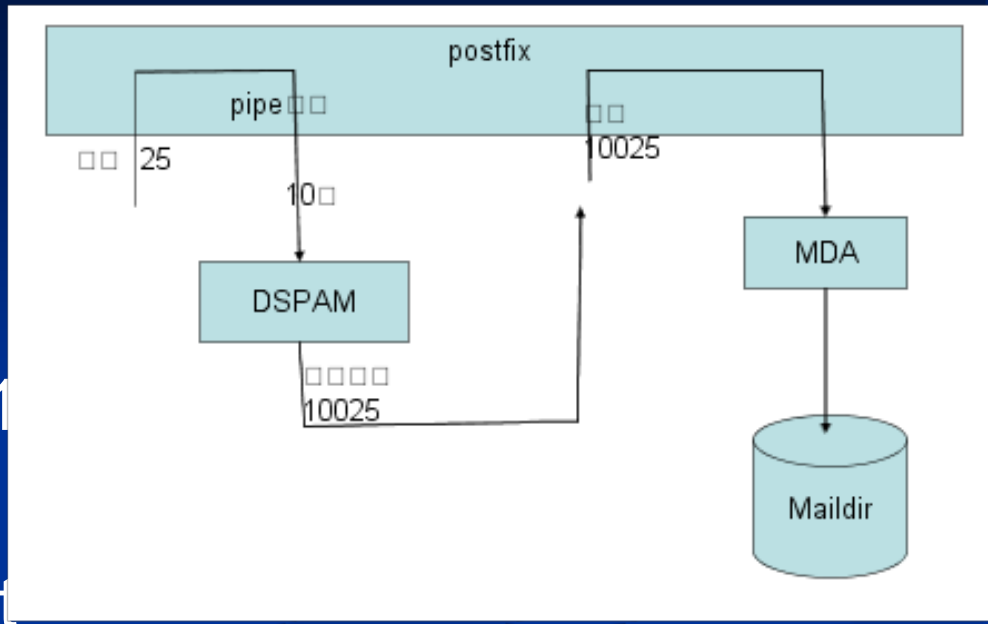
- 與postfix連結
  - master.cf
    - dspam unix - n n - **10** pipe
      - -o **soft_bounce=yes**
      - flags=Ru eol=\n user=dspam argv=/usr/local/bin/dspam --deliver=innocent --user $user -d $user
    - spam unix - n n - - pipe
      - flags=Ru user=dspam argv=/usr/local/bin/dspam --class=spam --source=error --user dspam
    - notspam unix - n n - - pipe
      - flags=Ru user=dspam argv=/usr/local/bin/dspam --class=innocent --source=error --user dspam
    - 0.0.0.0:10025 inet n - n - - smtpd
      - -o mynetworks=127.0.0.0/8
      - -o content_filter=
      - -o local_recipient_maps=
      - -o relay_recipient_maps=
      - -o smtpd_restriction_classes=
      - -o smtpd_client_restrictions=
      - -o smtpd_helo_restrictions=
      - -o smtpd_sender_restrictions=
      - -o smtpd_recipient_restrictions=permit_mynetworks,reject
      - -o strict_rfc821_envelopes=yes
      - -o smtpd_error_sleep_time=0

# Dspam的聯結方式

- 與postfix連結
  - main.cf
    - 方式一: content_filter=dspam:
    - 方式二: smtpd_recipient_restrictions = check_recipient_access pcre:<pcre file>
  - <pcre file>
    - /spam@cc.nctu.edu.tw/ OK
    - /notspam@cc.nctu.edu.tw/ OK
    - /@cc.nctu.edu.tw/ FILTER dspam:
    - /./ REJECT

# Dspam的聯結方式

- 與postfix連結
  - dspam.conf
    - DeliveryHost 127.0.0.1
    - DeliveryPort 10025
    - DeliveryIdent localhost
    - DeliveryProto SMTP

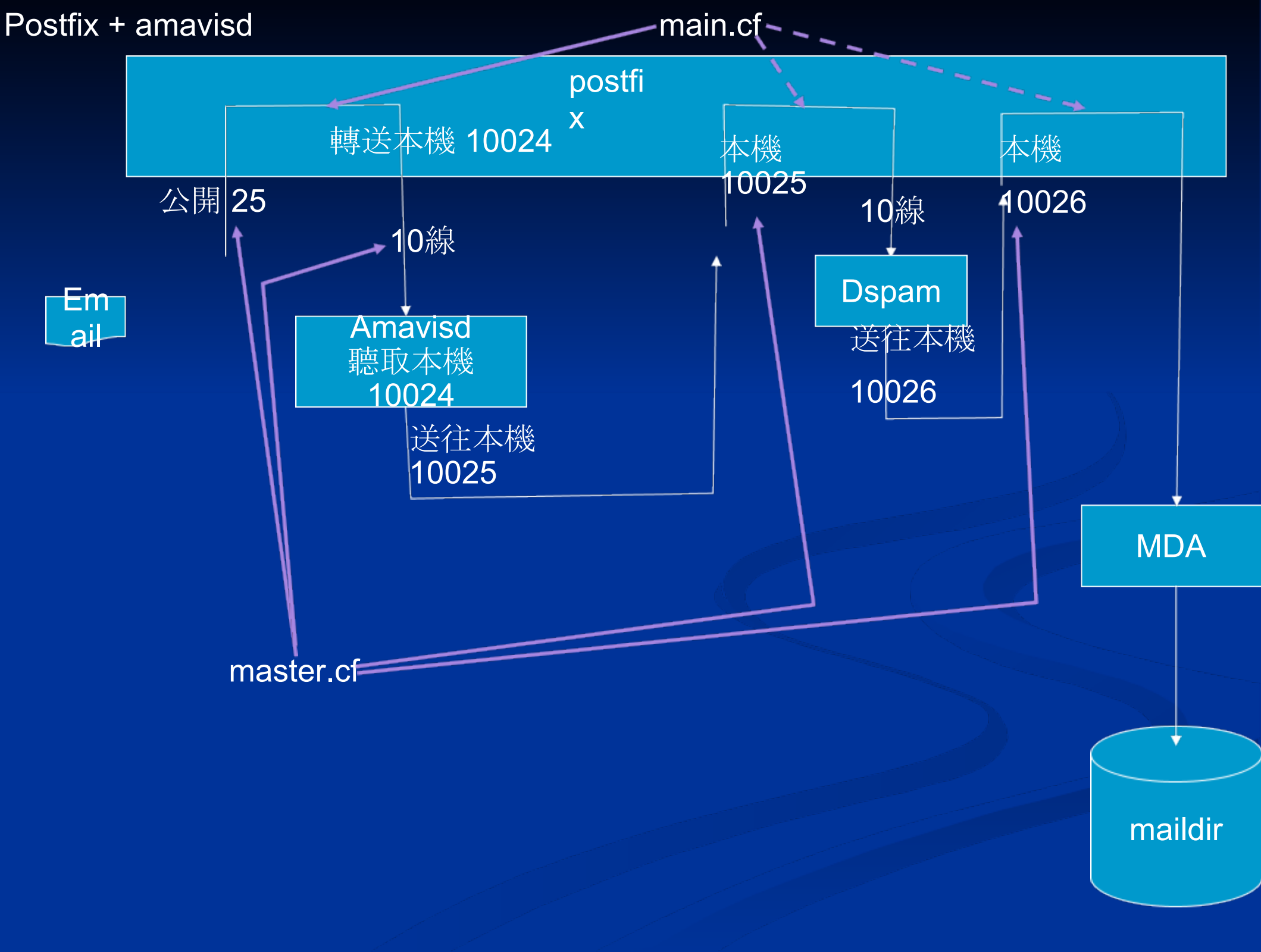


# postfix+amavis+dspam大串聯

- 首先先建立好postfix+amavis的連結
- master.cf
  - 增加一個dspam送回來的port
    - 0.0.0.0:**10026** inet n - n - - smtpd
      - -o mynetworks=127.0.0.0/8
      - -o content_filter=
      - . . . . .
  - 修改amavis送回來的port設定
    - 0.0.0.0:10025 inet n - n - - smtpd
      - -o mynetworks=127.0.0.0/8
      - **-o content_filter=dspam:**
      - . . . . .
  - 增加dspam的設定如前面的例子

# postfix+amavis+dspam大串聯

- dspam.conf
  - DeliveryHost 127.0.0.1
  - DeliveryPort **10026**
  - . . . . .


Postfix + amavisd
main.cf
postfix
轉送本機 10024
本機 10025
本機 10026
公開 25
10線
10線
Email
Amavisd 聽取本機 10024
送往本機 10025
Dspam
送往本機 10026
MDA
master.cf
maildir

# dspam的WebUI

- dspam內附一個WebUI管理介面

# dspam的WebUI

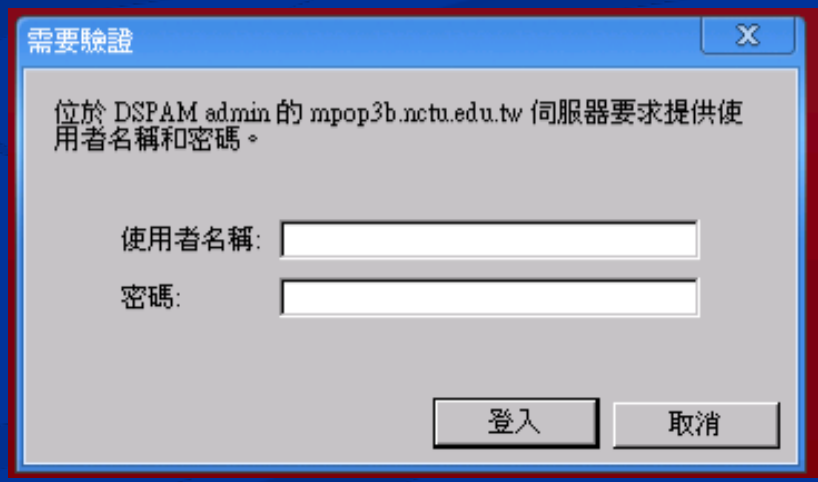
- 圖解功能說明
- http://mail.nctu.edu.tw/dspam.html

# dspam的WebUI

- 其WebUI沒有認證檢查功能, 留給管理者自行設計
- 設定環境變數 REMOTE_USER 為帳號之後再呼叫 dspam.cgi 即可
- 可自行設計一個登入畫面, 於驗證過後設定此環境變數, 再導向dspam.cgi
- 透過apache+pwauth 做本機認證最簡單

# dspam的WebUI

- 安裝pwauth
- 安裝 mod_authnz_external
- httpd.conf中, 增加以下設定範例
  - AddExternalAuth pwauth /usr/local/bin/pwauth
  - SetExternalAuthMethod pwauth pipe
  - <Directory "/usr/local/www/apache22/cgi-bin/dspam"
    >
  - AuthName "DSPAM admin"
  - AuthType Basic
  - AuthBasicProvider external
  - AuthExternal pwauth
  - Require valid-user
  - </Directory>



# Dspam的過濾原理與訓練方式

- 沒有任何預先設定的rules
- 對於拆解的字串稱之為factor(s)
- 對於拆解的字串做的hash稱之為token(s)
- 資料庫不存在的token即視為預設值
  - --deliver=innocent 預設為正常信的值
  - --deliver=spam 預設為廣告信的值

Email

取出Factors, 產生 tokens

From*u@mail.nctu.edu.tw>, 0.00414,
Date*00+0800, 0.01000,
Date*Tue, 0.14553,
is+a, 0.15627,
This+is, 0.15702,
MIME, 0.16252,
Received*Tue, 0.16330,
Received*Tue, 0.16330,
message, 0.16923,
is, 0.17303,
This, 0.17411,
a, 0.18613,
Received*(dspam.cc.nctu.edu.tw+[140.113.2.71]), 0.79506,
Received*dspam.cc.nctu.edu.tw+(dspam.cc.nctu.edu.tw, 0.79506,
message+in, 0.21247,
part+message, 0.21281,
a+multi, 0.21281,
multi+part, 0.21377,
part, 0.21438,
multi, 0.21480,
format, 0.21542,
in+MIME, 0.21679,
MIME+format, 0.21713,
Content-Type*charset="big5", 0.22757,
From*mail.nctu.edu.tw>, 0.23795,
Received*(localhost, 0.24205,
Received*(localhost, 0.24205

找DB, 取出參考值

MySQL Dspam_token_data

參考值加總

X-DSPAM-Confidence: 0.7595

做出判決

X-DSPAM-Probability: 0.0000

使用者設定的寬鬆度

MySQL Dspam_preferences

Em
ail

根據使用者設定學習行為

DSPAM should train:

- On every new message scanned by the filter
- (•) Only when the filter makes a mistake
- Only with new data or if the filter makes a mistake

MySQL
Dspam_preferences

將factors以前面抓出的值減0.01寫入DB

不處理

僅對沒找到值的factors以-0.01寫入DB

MySQL
Dspam_token_data

抓取信件前64k內容(包含header), 寫入DB
作為未來更正學習用

MySQL
Dspam_signature_data

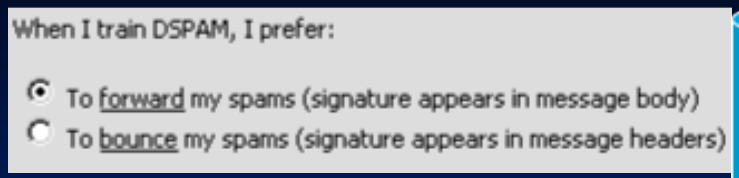
寫入統計資訊

| | SPAM messages | Good messages |
|---|---|---|
| Since last reset | 432 missed | 314 missed |
| | 8840 caught | 34658 delivered |
| | 95.341% caught | 0.898% missed |
| Total processed by filter | 665 missed | 329 missed |
| | 9235 caught | 42999 delivered |
| From corpus | 955 fed | 3 fed |

MySQL
Dspam_stats

Em
ail

根據設定寫入header與body

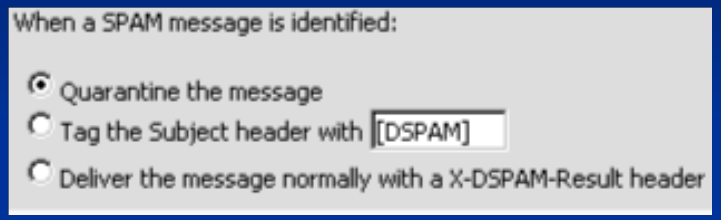



MySQL
Dspam_preference
s

X-DSPAM-Signature: 2842,49f6c001539039982617038

格式為 UID, dspam mail id

根據設定決定垃圾信處理方式




置入攔截區的mbox file

quarantin
e

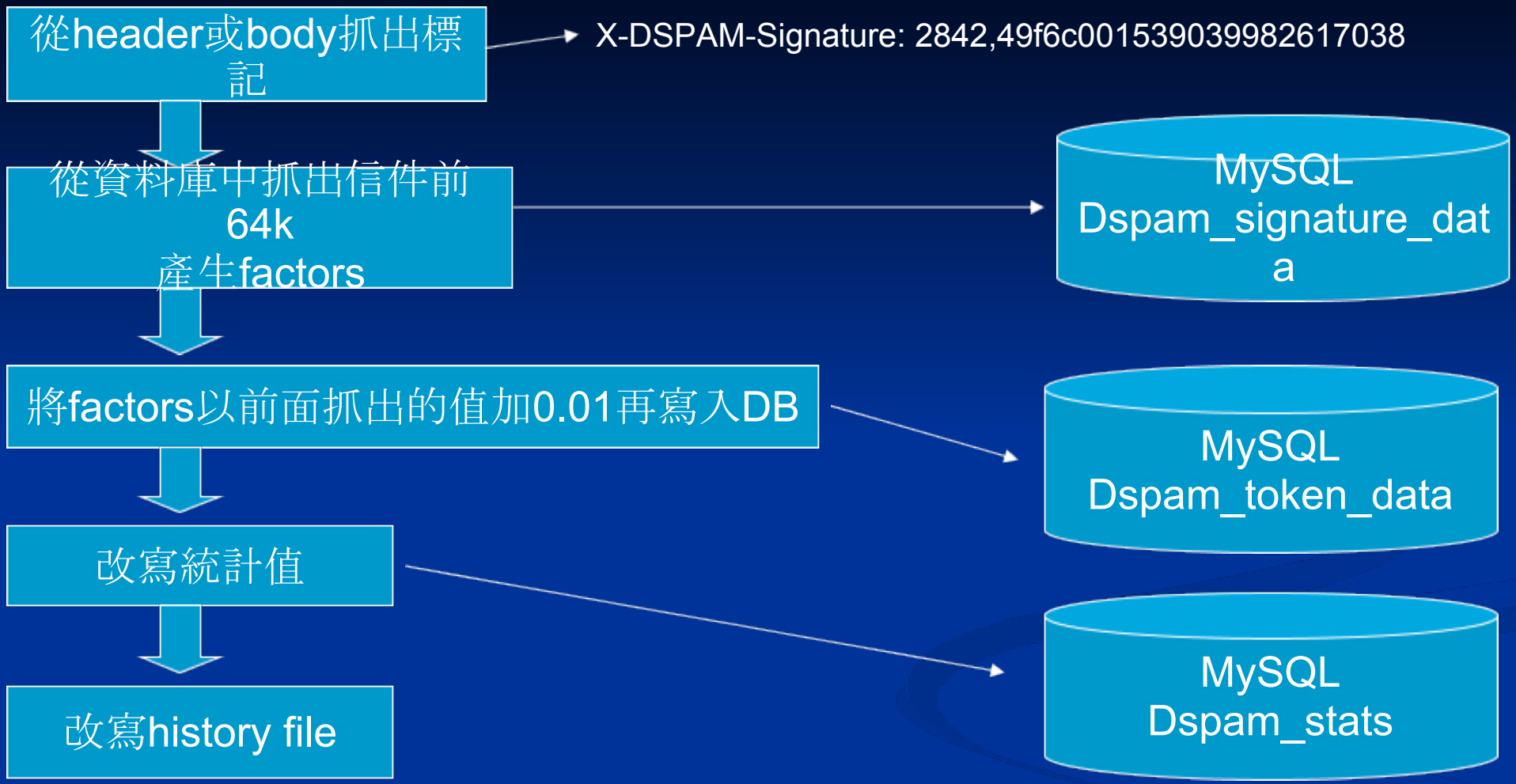
標題前加入指定字串

Header中加入標記

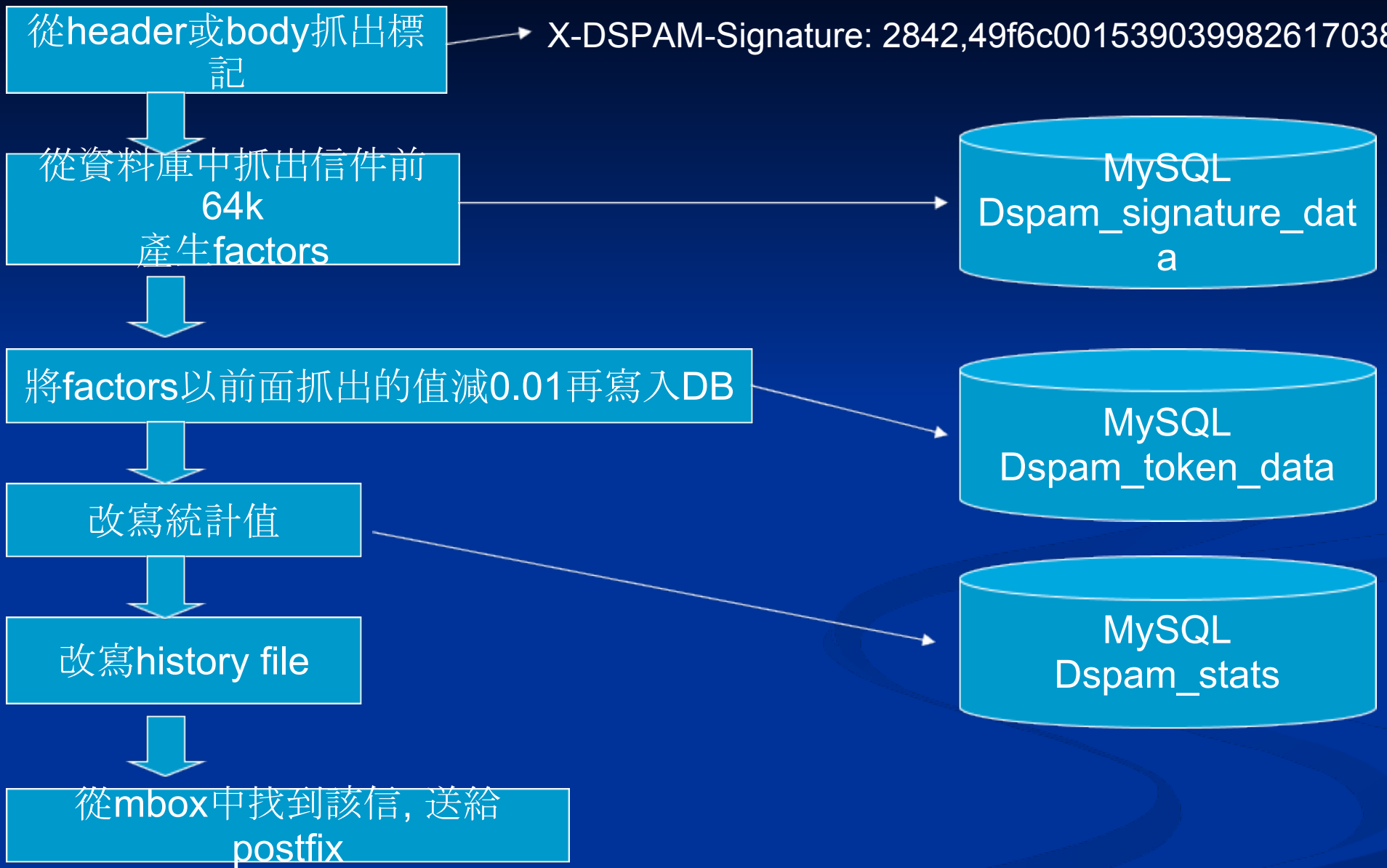
送回postfix

寫入history file

更正為廣告信


從header或body抓出標記
X-DSPAM-Signature: 2842,49f6c0015390399826170 38
從資料庫中抓出信件前64k產生factors
MySQL Dspam_signature_data
將factors以前面抓出的值加0.01再寫入DB
MySQL Dspam_token_data
改寫統計值
MySQL Dspam_stats
改寫history file

更正為正常信


從header或body抓出標記
X-DSPAM-Signature: 2842,49f6c00153903998261703 8
從資料庫中抓出信件前64k產生factors
MySQL Dspam_signature_data
將factors以前面抓出的值減0.01再寫入DB
MySQL Dspam_token_data
改寫統計值
MySQL Dspam_stats
改寫history file
從mbox中找到該信, 送給postfix