

Traženje igle u stogu sijena 2: mairix



Prije nekog vremena opisali smo kako putem alata [grepmail](#) [1] pronaći mail prema određenim kriterijima u vašim mailbox datotekama. Iako se samom alatu nema što prigovoriti, ovaj način pretraživanja kod velikih mailbox datoteka može biti spor, te ne može rekurzivno pretraživati mailbox datoteke u drugim direktorijima. Sve ovo rješava **mairix**, alat koje će sve vaše mailbox foldere rekurzivno indeksirati, te će pretraživanje višegigabajtne arhive biti gotovo u samo nekoliko sekundi.

Mairix može indeksirati mbox, Maildir i MH formate zapisa, tako da je pokrivena i situacija na tipičnom CARNet poslužitelju. Instalacija je jednostavna:

```
# apt-get install mairix
The following NEW packages will be installed: mairix
Unpacking mairix (from .../mairix_0.17-2_i386.deb) ...
Setting up mairix (0.17-2) ...
#
```

Sljedeći korak je, kao i uvijek, malo konfiguracije. U paketu dolazi datoteka `/usr/share/doc/mairix/dotmairixrc.eg` iz koje možete vidjeti kako bi trebala izgledati tipična konfiguracijska datoteka. No, za početak možete upotrijebiti ovu konfiguraciju:

```
base=$HOME/mail
database=$HOME/.mairix_database
mfolder=Search
mformat=mbox
mbox=...
```

Objasnit ćemo svaku stavku. Kao što se može pretpostaviti iz samog njenog naziva, opcija **base** određuje gdje se nalaze vaše mailbox datoteke. Ovo uključuje i sve eventualne direktorije ispod toga. Opcija **database** određuje lokaciju indeksne datoteke, i možete je ostaviti kako je predloženo.

Opcija **mfolder** određuje gdje će se snimiti mailovi koji odgovaraju kriteriju pretraživanja. Ukoliko je navedeno samo ime datoteke, njena staza se dodaje stazi zadanoj u opciji **base**, u ovom primjeru to će biti `$HOME/mail/Search`. Ukoliko navedemo apsolutnu stazu `"/tmp/Search"`, rezultati pretraživanja će se tamo i snimati, bez obzira na vrijednost u opciji **base** (iako, ne preporučujemo snimanje u `/tmp!`).

Mformat određuje o kojem se formatu zapisa radi, a u našem slučaju je to klasični mbox. Ukoliko rabite Maildir, tako i upišite.

Najvažnija opcija je **mbox**. Ukoliko navedete samo tri točke "...", to označava da želite rekurzivno indeksiranje svega u `$HOME/mail`. Ukoliko želite indeksiranje samo određenih datoteka u `$HOME/mail`, primjerice samo mailboxova koji pocinju sa 2008-*,

```
mbox=2008*
```

U slučaju da u \$HOME/mail imate mailboxove raspodjeljene po godina, 2006, 2007, 2008, i želite ih indeksirati, a ostalo ne, stavite:

```
mbox=200...
```

Dakle, trostruka točka određuju rekursiju, odnosno da ukoliko neki direktorij počinje s 200, i njega će se indeksirati. Ukoliko želite indeksirati samo određene foldere, navedite ih, odijeljene dvotočkama:

```
mbox=mbox1:mbox2:mbox3
```

Za ostale mogućnosti pogledajte u dokumentaciju, mada je indeksiranje svih mboxova u potpunosti realna i korisna opcija. Vrijedi napomenuti da mairix ne radi locking mboxova, što znači da ukoliko želite indeksirati INBOX, možda neće biti indeksirane sve poruke. No, indeksiranje INBOX-a i nije tako korisna opcija.

Ukoliko postoje samo neki mailboxovi koje ne želite indeksirati, postoji i opcija **omit**. Njena uporaba je malo drugačija, primjerice:

```
mbox=...  
omit=spam*
```

će indeksirati sve, osim mboxova koji počinju sa spam (primjerice spam, spam2008, spam2009...). Ukoliko je spam direktorij, pomoći će:

```
mbox=...  
omit=spam**
```

što će indeksirati sve, osim mailboxova i direktorija koji počinju sa spam, primjerice i spam2009 i spam/bilosto.

Nakon konfiguracije, slijedi indeksiranje. Jednostavno pokrenite mairix:

```
$ mairix
```

Ova će operacija potrajati neko vrijeme, i možda će ispisati pokoju grešku, no nemate neki izbor osim zanemariti ih. Kako bi indeks bio aktualan, ovu operaciju stavite u cron, npr:

```
0 7 * * * cd $HOME && mairix
```

Najzad, uporaba. Ona je različita nego ona kod grepmaila, ali nije ništa teža za zapamtiti. Primjerice,

```
$ mairix pojam
```

će generirati mailbox kojeg ste definirali u mfolder opciji, a koji će sadržavati mailove koji u zaglavlju ili tijelu poruke sadržavaju taj pojam. Kako to može dati previše rezultata, ograničit ćemo pretraživanje:

```
$ mairix f:pero@pero.hr
```

će dati sve mailove, isfiltrirane po kriteriju pošiljatelja (f=From). Slično tome:

\$ mairix t:netko@domena.hr

će pronaći sve mailove koji su poslani na adresu "netko@domena.hr". Vodeći se tom logikom, "c" označava Cc: polje, dok "a" označava "Address" dakle sva adresna polja: From:, To:, Cc:. Nadalje, "s" označava "Subject:" polje:

\$ mairix s:Jucer

Analogno tome, "b" označava "body", dakle tijelo poruke. Moguće je i kombinirati:

\$ mairix fb:Pero

ili

\$ mairix f:Pero b:Marko # (u ovom slučaju može i drugi pojam)

će pronaći sve mailove gdje se pojam "Pero" može pronaći u tijelu poruke ili u From: zaglavlju. Moguće su i negacije:

\$ mairix f:~pero@pero.hr

će pronaći sve mailove koji ne potječu od korisnika "pero@pero.hr". Zanimljiva je opcija koja omogućava približno pretraživanje:

\$ mairix s:pojam=1

koje će pronaći sve retke Subject iz zaglavlja koji imaju najviše jednu grešku od navedenog pojma, dakle uvjete zadovoljava i "pojam" i "dojam". Ukoliko vam je potreban samo dio pojma (*substring*), upotrijebit ćete:

\$ mairix s:ivati=

koji će pronaći sve pojmove u Subjectu koji sadržavaju "ivati", npr. "otkrivati", "prikrivati" i slično. Vremenske odrednice možete upotrijebiti na način:

\$ mairix d:1m-
\$ mairix d:3w-6w
\$ mairix d:-6y
\$ mairix d:20091201- s:pojam

će pronaći sve mailove unazad jedan mjesec do danas, sve mailove stare između 3 i 6 tjedana, te sve mailove starije od 6 godina (respektivno). U zadnjem slučaju, mairix će pronaći sve mailove od 1. prosinca 2009. do danas, i koji u Subjectu imaju "pojam". Ukoliko vas zanima veličina mailova, upotrijebit ćete sličnu sintaksu:

\$ mairix z:1k-
\$ mairix z:1k-5k
\$ mairix z:-5m

što će pronaći sve mailove veće od jednog kilobajta, sve mailove između jednog i pet kilobajta, te sve mailove manje od 5 megabajta.

Čitanje mailboxa koji se stvori pretraživanjem je lako, u ovisnosti o mail klijentu kojeg rabite.

```
$ alpine -f $HOME/mail/Search  
$ mutt -f $HOME/mail/Search
```

Dokumentacija također dolazi u paketu, i nalazi se u HTML formatu u datoteci /usr/share/doc/mairix/mairix.html. Službeno web-sjedište projekta možete naći na adresi: <http://www.rpcurnow.force9.co.uk/mairix/> [2].

UPDATED: 2010-01-05

- [Logirajte](#) [3] se za dodavanje komentara

sri, 2009-02-11 22:40 - Željko Boroš**Vote:** 0

No votes yet

Source URL: <https://sysportal.carnet.hr/node/515>

Links

[1] <https://sysportal.carnet.hr/node/480>

[2] <http://www.rpcurnow.force9.co.uk/mairix/>

[3] <https://sysportal.carnet.hr/sysportallogin>