

Volltextsuche in Mails mit Solr

Diese Anleitung wurde für Debian 8 (Jessie) erstellt.

Versionen:

- Dovecot 2.2.27 (aus den Debian Backports, da die im normalen Debian enthaltene Version ein Problem mit Solr 6.5 hat)
- Solr 6.5.1 (Standalone Installation)
- OpenJDK 8 (ebenfalls aus den Backports aufgrund der neueren Version)

Installation/Upgrade Dovecot

Debian Backports aktivieren:

/etc/apt/sources.list.d/backports.list

```
deb http://ftp.de.debian.org/debian jessie-backports main contrib non-free
```

```
aptitude -t jessie-backports install dovecot-core dovecot-solr (+ weitere Pakete, wenn nötig (z.
```

Installation OpenJDK 8

```
aptitude -t jessie-backports install openjdk-8-jdk-headless openjdk-8-jre-headless
```

Installation Solr 6.5.1

Das aktuelle Solr .tgz von apache.org herunterladen und entpacken. Mit dem Installerscript lässt sich der Solr sehr einfach einrichten:

```
cd /root
wget http://mirror.synyx.de/apache/lucene/solr/6.5.1/solr-6.5.1.tgz
tar xvzf solr-6.5.1.tgz
bin/install_solr_service.sh /root/solr-6.5.1.tgz
```

Der Solr-Admin sollte nun über den Browser erreichbar sein: <http://localhost:8983>

Auf dem Server nun die Datenverzeichnisse anlegen und die Config von Dovecot rüberkopieren:

```
sudo su - solr -c "/opt/solr/bin/solr create -c dovecot -n dovecot"
```

Das Datenverzeichnis liegt in /var/solr/data/dovecot. Bei Solr 6.X kann das Schema über die WebGUI konfiguriert werden. Dovecot kommt aber mit einem eigenen Schema, daher wird beim Deploy dieser Konfiguration diese Möglichkeit deaktiviert.

Nun noch schema.xml und solrconfig.xml nach `/var/solr/data/dovecot/conf` kopieren und die Datei `managed-schema` löschen. Aufgrund der Übersichtlichkeit habe ich diese Dateien im Anhang hinterlegt.

```
cp schema.xml solrconfig.xml /var/solr/data/dovecot/conf
rm /var/solr/data/dovecot/conf/managed-schema
```

Nach einem Neustart des Solr-Services kann Dovecot konfiguriert werden:

```
systemctl restart solr.service
```

Konfiguration von Dovecot

Ich verwende auf dem Dovecot-Server ISPConfig. ISPConfig verwendet nicht das `conf.d`-Modell mit einzelnen Konfigurationsdateien, sondern packt alles direkt in die `/etc/dovecot/dovecot.conf`. Bitte ändere deine Dateien dort wo notwendig.

`/etc/dovecot/dovecot.conf`

```
[...]
mail_plugins = quota fts fts_solr
[...]
plugin {
  [...]
  fts = solr
  fts_solr = break-imap-search url=http://localhost:8983/solr/dovecot/
  fts_autoindex = yes
  [...]
}
```

Selbstverständlich können Solr und Dovecot auch auf verschiedenen Servern gehostet werden. Dafür muss nur die IP-Adresse entsprechend angepasst werden. Der Solr-Port (8983) sollte nach außen selbstverständlich mit einer Firewall gesichert werden oder Solr auf eine interne IP gebunden werden. In meinem Setup liegt Solr auf einem internen Server ohne öffentliche IP-Adresse.

Nach einem Neustart per `systemctl restart dovecot.service` kann die Volltextsuche bereits verwendet werden. Dazu muss zuerst der Index erzeugt werden:

```
doveadm index -A -q '*'
```

Da die serverseitige Volltextsuche eine Erweiterung der Suche im IMAP-Protokoll ist, muss ein Client verwendet werden, der dies unterstützt. Mir sind aktuell bekannt: Claws Mail / Sylpheed, Roundcube, Thunderbird.

Revision #1

Created 30 April 2021 12:27:43 by magenbrot

Updated 30 April 2021 12:29:09 by magenbrot