



Kostenloses eBook

LERNEN

solr

Free unaffiliated eBook created from
Stack Overflow contributors.

#solr

Inhaltsverzeichnis

Über.....	1
Kapitel 1: Erste Schritte mit solr.....	2
Bemerkungen.....	2
Examples.....	2
Einführung.....	2
SOLR installieren.....	4
Einführung.....	4
Voraussetzungen.....	4
SOLR herunterladen.....	6
Installieren.....	7
Testen Sie Ihre Installation.....	8
Beispiel für die Solr-Suche.....	14
Kapitel 2: Apache Solr.....	16
Examples.....	16
Solr Installation.....	16
Solr-Schnellstart (Solar starten und stoppen).....	16
Kapitel 3: Fehlerbehebung Solr.....	17
Bemerkungen.....	17
Examples.....	17
Ist mein Inhalt tatsächlich indiziert?.....	17
Kapitel 4: INDEXIERUNG DER TABELLE HIVE2 IN SOLR MIT SOLR DIH.....	19
Einführung.....	19
Bemerkungen.....	19
Examples.....	19
Schritte.....	19
Kapitel 5: Lucene-Abfragesyntax.....	21
Examples.....	21
Nähe suchen.....	21
Einfache Suche.....	21
Boolesche Suche.....	21

Phrasensuche.....	21
Suchbegriffe verstärken.....	21
Platzhaltersuche.....	21
Bereichssuche.....	22
Join über Kerne.....	22
Kapitel 6: So erstellen Sie einen benutzerdefinierten Feldtyp.....	23
Bemerkungen.....	23
Examples.....	23
Erstellen Sie einen benutzerdefinierten Solr-Feldtyp aus der eigenen benutzerdefinierten J.....	23
Erstellen Sie einen benutzerdefinierten Feldtyp aus verfügbaren Feldtypen.....	24
Credits.....	26



You can share this PDF with anyone you feel could benefit from it, downloaded the latest version from: [solr](#)

It is an unofficial and free solr ebook created for educational purposes. All the content is extracted from [Stack Overflow Documentation](#), which is written by many hardworking individuals at Stack Overflow. It is neither affiliated with Stack Overflow nor official solr.

The content is released under Creative Commons BY-SA, and the list of contributors to each chapter are provided in the credits section at the end of this book. Images may be copyright of their respective owners unless otherwise specified. All trademarks and registered trademarks are the property of their respective company owners.

Use the content presented in this book at your own risk; it is not guaranteed to be correct nor accurate, please send your feedback and corrections to info@zzzprojects.com

Kapitel 1: Erste Schritte mit solr

Bemerkungen

In diesem Abschnitt erhalten Sie einen Überblick darüber, was solr ist und warum ein Entwickler es verwenden möchte.

Es sollte auch alle großen Themen in solr erwähnen und auf die verwandten Themen verweisen. Da die Dokumentation für solr neu ist, müssen Sie möglicherweise erste Versionen dieser verwandten Themen erstellen.

Examples

Einführung

Solr ist ein eigenständiger Server für die Unternehmenssuche mit einer REST-ähnlichen API. Sie legen darin Dokumente ("Indexierung" genannt) über JSON, XML, CSV oder Binary über HTTP an. Sie fragen es über HTTP GET ab und erhalten JSON-, XML-, CSV- oder binäre Ergebnisse. Solr verwendet die Lucene-Suchbibliothek und erweitert diese.

Hier sind einige der wichtigsten Funktionen von solr:

- **Erweiterte Volltextsuchfunktionen**

Dank Lucene™ bietet Solr leistungsstarke Matching-Funktionen, darunter Phrasen, Platzhalter, Verknüpfungen, Gruppierungen und vieles mehr für jeden Datentyp

- **Optimiert für hohen Verkehrsaufkommen**

Solr hat sich weltweit in extrem großen Maßstäben bewährt

- **Auf Standards basierende offene Schnittstellen - XML, JSON und HTTP**

Solr verwendet die Tools, die Sie verwenden, um das Erstellen von Anwendungen zu einem Snap zu machen

- **Umfassende Administrationsschnittstellen**

Solr wird mit einer integrierten, reaktionsfähigen Benutzeroberfläche für die Verwaltung geliefert, um die Steuerung Ihrer Solr-Instanzen zu vereinfachen

- **Einfache Überwachung**

Benötigen Sie mehr Einblick in Ihre Instanzen? Solr veröffentlicht eine Vielzahl von Metrikdaten über JMX

- **In hohem Maße skalierbar und fehlertolerant**

Solr ist auf dem kampferprobten Apache Zookeeper aufgebaut und erleichtert das Skalieren nach oben und unten. Solr backt beim Replizieren, Verteilen, Neuausgleichen und Fehlertoleranz aus der Box.

- **Flexibel und anpassungsfähig mit einfacher Konfiguration**

Solr's ist so konzipiert, dass es sich Ihren Bedürfnissen anpasst und gleichzeitig die Konfiguration vereinfacht

- **Nahe Echtzeit-Indizierung**

Möchten Sie Ihre Updates jetzt sehen? Solr nutzt die Near Real-Time Indexing-Funktionen von Lucene, um sicherzustellen, dass Sie Ihren Inhalt sehen, wenn Sie ihn sehen möchten

- **Erweiterbare Plugin-Architektur**

Solr veröffentlicht viele genau definierte Erweiterungspunkte, die das Einfügen von Plugins für Index- und Abfragezeiten erleichtern. Da es sich um von Apache lizenzierte Open Source-Software handelt, können Sie jeden Code beliebig ändern!

Einige solr coole Funktionen:

- **Schema, wenn Sie wollen, Schemalos, wenn Sie nicht möchten**

Verwenden Sie den datengesteuerten schematischen Modus von Solr, wenn Sie mit den ersten Schritten beginnen, und sperren Sie ihn dann, wenn die Produktion beginnt.

- **Leistungsfähige Erweiterungen**

Solr wird mit optionalen Plugins für die Indizierung umfangreicher Inhalte (z. B. PDFs, Word), Spracherkennung, Clusterbildung von Suchergebnissen und mehr geliefert

- **Facettierte Suche und Filterung**

Schneide und schneide deine Daten so, wie du sie für richtig hältst, und zwar mit einer Vielzahl von Facettenalgorithmen

- **Geodatenuche**

Mit der integrierten Unterstützung von Solr für die räumliche Suche ist das Aktivieren der standortbasierten Suche einfach

- **Abfragevorschläge, Rechtschreibung und mehr**

Solr wird mit erweiterten Funktionen für die automatische Vervollständigung (Typheadersuche), Rechtschreibprüfung und mehr geliefert

- **Rich Document Parsing**

Solr wird mit dem integrierten Apache Tika ausgeliefert, was die Indizierung umfangreicher Inhalte wie Adobe PDF, Microsoft Word und mehr vereinfacht.

SOLR installieren

Einführung

Das folgende Verfahren wurde an einer Testinstanz in AWS mit Redhat und Solr 6.1.0. Möglicherweise müssen Sie den Prozess entsprechend an Ihr Betriebssystem und Ihre Umgebung anpassen.

Voraussetzungen

1. Stellen Sie sicher, dass Sie RedHat oder ein ähnliches (Fedora-basiertes) Betriebssystem verwenden.

```
cat /etc/redhat-release  
zeigt Ihre OS-Version an.
```

```
[ec2-user@ip-172-31-28-149 ~]$ cat /etc/redhat-release  
Red Hat Enterprise Linux Server release 7.2 (Maipo)
```

2. Prüfen Sie, ob Java 1.6 oder höher installiert ist

```
which java
```

```
[ec2-user@ip-172-31-28-149 ~]$ which Java  
/usr/bin/which: no Java in (/usr/local/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/home/ec2-user/.local/bin:/home/ec2-user/bin)
```

3. Installieren Sie gegebenenfalls Java

```
sudo yum list available java*
```

```
[ec2-user@ip-172-31-28-149 ~]$ sudo yum list available java*  
Loaded plugins: amazon-id, rhui-lb, search-disabled-repos  
Available Packages  
java-1.6.0-openjdk.x86_64                1:1.6.0.40-1.13.12.5.e17_  
java-1.6.0-openjdk-devel.x86_64        1:1.6.0.40-1.13.12.5.e17_  
java-1.7.0-openjdk.x86_64                1:1.7.0.111-2.6.7.2.e17_  
java-1.7.0-openjdk-devel.x86_64        1:1.7.0.111-2.6.7.2.e17_  
java-1.7.0-openjdk-headless.x86_64     1:1.7.0.111-2.6.7.2.e17_  
java-1.8.0-openjdk.x86_64                1:1.8.0.111-1.b15.e17_2
```

```
sudo yum install java-1.8.0-openjdk.x86_64
```

4. Prüfen Sie, ob es richtig installiert ist

```
which java das Java-Zuhause anzeigt  
java -version zeigt die Java-Version an
```

```
[ec2-user@ip-172-31-28-149 ~]$ which java
/usr/bin/java
[ec2-user@ip-172-31-28-149 ~]$ java -version
openjdk version "1.8.0_111"
OpenJDK Runtime Environment (build 1.8.0_111-b15)
OpenJDK 64-Bit Server VM (build 25.111-b15, mixed mode)
```

5. Erstellen Sie einen SOLR-Benutzer.

```
sudo adduser solr
```

6. Fügen Sie ein Passwort für den Benutzer hinzu.

```
sudo passwd solr
```

```
[ec2-user@ip-172-31-28-149 ~]$ sudo adduser solr
[ec2-user@ip-172-31-28-149 ~]$ sudo passwd solr
Changing password for user solr.
New password:
Retype new password:
passwd: all authentication tokens updated successfully.
```

7. Aktivieren Sie sudo für den SOLR-Benutzer, und führen Sie visudo aus

```
sudo visudo
```

8. Suchen Sie die folgenden Zeilen:

```
## Allows people in group wheel to run all commands
# %wheel ALL=(ALL) ALL
```

9. Wenn `%wheel` auskommentiert ist, kommentieren Sie die zweite Zeile ab, indem Sie das Zeichen `#` entfernen.

```
%wheel ALL=(ALL) ALL
```

```
## Allows people in group wheel to run all commands
%wheel ALL=(ALL) ALL
```

10. Wenn Sie eine Änderung vorgenommen haben, verwenden Sie `:wq` Andernfalls verwenden Sie `:q` zum Beenden.

11. Fügen Sie den Solr-Benutzer der Radgruppe hinzu.

```
sudo usermod -aG wheel solr
```

12. `solr` Sie zum `solr` Benutzer und prüfen Sie, ob Sie über Root-Berechtigungen verfügen:

```
su solr -
sudo whoami
```



```
[ec2-user@ip-172-31-28-149 ~]$ sudo usermod -aG wheel solr
[ec2-user@ip-172-31-28-149 ~]$ su solr
Password:
[solr@ip-172-31-28-149 ec2-user]$ groups
solr wheel
[solr@ip-172-31-28-149 ec2-user]$ sudo whoami

We trust you have received the usual lecture from the local
Administrator. It usually boils down to these three things:

    #1) Respect the privacy of others.
    #2) Think before you type.
    #3) With great power comes great responsibility.

[sudo] password for solr:
root
```

SOLR herunterladen

13. Finden Sie Ihren lokalen Spiegel unter: <http://www.apache.org/dyn/closer.lua/lucene/solr/>

14. Ändern Sie Ihren Verzeichnisstandort zur Auswahl:

```
cd /opt/
```

15. Laden Sie eine Kopie des Pakets vom Spiegel herunter:

```
sudo curl -O http://www.trieuvan.com/apache/lucene/solr/6.1.0/solr-6.1.0.tgz
```

```
[solr@ip-172-31-28-149 opt]$ sudo curl -O http://www.trieuvan.com/apache/lucene/solr/6.2.0/solr-6.2.0.tgz
```

% Total Time	% Received Current	% Xferd	Average Speed Dload	Upload	Time Total	Time Spent
0	0	0	0	0	0	0
3	142M	3	5343k	0	0	5501k
9	142M	9	13.5M	0	0	7041k
13	142M	13	19.3M	0	0	6656k
16	142M	16	23.4M	0	0	6028k
19	142M	19	27.4M	0	0	5656k

16. Entpacken Sie das Paket:

```
sudo tar zxvf solr-6.1.0.tgz
```

17. Kopieren Sie das Installationskript in Ihren Ordner:

```
sudo cp /opt/solr-6.1.0/bin/install_solr_service.sh .
```

18. Entfernen Sie die nicht benötigten Dateien:

```
sudo rm -rf solr-6.1.0
```

Installieren

1. Führen Sie das Installationskript aus:

```
sudo ./install_solr_service.sh solr-6.1.0.tgz
```

```
Started Solr server on port 8983 (pid=9607). Happy searching!  
  
Found 1 Solr nodes:  
  
Solr process 9607 running on port 8983  
{  
  "solr_home":"/var/solr/data",  
  "version":"6.2.0 764d0f19151dbff6f5fcd9fc4b2682cf934590c5 - mi  
ke - 2016-08-20 05:41:37",  
  "startTime":"2016-10-20T04:02:37.028Z",  
  "uptime":"0 days, 0 hours, 0 minutes, 15 seconds",  
  "memory":"85.4 MB (%17.4) of 490.7 MB"}  
  
Service solr installed.
```

2. Führen Sie den Autostart des SOLR-Dienstes durch, wenn der Server neu gestartet wird.

```
sudo chkconfig --add solr
```

```
chkconfig | grep solr
```

```
[solr@ip-172-31-28-149 ec2-user]$ sudo chkconfig --add solr  
[sudo] password for solr:  
[solr@ip-172-31-28-149 ec2-user]$ chkconfig |grep solr  
  
Note: This output shows SysV services only and does not include native  
systemd services. SysV configuration data might be overridden by native  
systemd configuration.  
  
If you want to list systemd services use 'systemctl list-unit-files'.  
To see services enabled on particular target use 'systemctl list-dependencies [target]'.  
  
solr          0:off  1:off  2:on   3:on   4:on   5:off  
n            6:off
```

3. Service-Besitzer ändern

```
sudo chown -R solr:solr /var/solr/
```

Testen Sie Ihre Installation

1. Erstellen Sie einen Kern von der Befehlszeile aus:

```
sudo su - solr -c "/opt/solr/bin/solr create -c NewCore1 -n data_driven_schema_configs"
```

```
[solr@ip-172-31-28-149 ec2-user]$ sudo su - solr -c "/opt/solr/bin/solr create -c NewCore1 -n data_driven_schema_configs"
```

```
Copying configuration to new core instance directory:  
/var/solr/data/NewCore1
```

```
Creating new core 'NewCore1' using command:
```

```
http://localhost:8983/solr/admin/cores?action=CREATE&name=NewCore1&instanceDir=NewCore1
```

```
{  
  "responseHeader": {  
    "status": 0,  
    "QTime": 2856},  
  "core": "NewCore1"}  
}
```

2. Öffnen Sie den Admin in einem Browser:

```
http://<solr_server>:8983/solr/#/
```

Solr Admin

ec2-35-160-125-216.us-west-2.compute.amazonaws.com:8983/solr/#/

Solr

- Dashboard
- Logging
- Core Admin
- Java Properties
- Thread Dump
- Core Selector

Instance

Start 7 minutes ago

Versions

solr-spec	6.2.0
solr-impl	6.2.0 764d0f19151dbff6f5fcd9fc4b2682cf934590c5 - mike - 2016-
lucene-spec	6.2.0
lucene-impl	6.2.0 764d0f19151dbff6f5fcd9fc4b2682cf934590c5 - mike - 2016-

JVM

- Runtime: Oracle Corporation OpenJDK 64-Bit Server VM 1.8.0_111 25.111-b1
- Processors: 1
- Args:
 - DSTOP.KEY=solrrocks
 - DSTOP.PORT=7983
 - Djetty.home=/opt/solr/server
 - Djetty.port=8983
 - Dlog4j.configuration=file:/var/solr/log4j.properties

3. Öffnen Sie die Liste der Kerne im Menü, um den NewCore1-Kern anzuzeigen.

Statistics

- Last Modified: -
- Num Docs: 0
- Max Doc: 0
- Heap Memory 0
- Usage: Deleted Docs: 0
- Version: 2
- Segment Count: 0
- Optimized: ✓
- Current: ✓

Replication (Master)

	Version	Gen	Size
Master (Searching)	0	1	71 bytes
Master (Replicable)	-	-	-

+ Admin Extra

4. Testen Sie, ob Kerne klebrig sind:

```
sudo service solr restart
```

5. Aktualisieren Sie den Admin in einem Browser:

`http://<solr_server>:8983/solr/#/` Sie sicher, dass die Admin-Seite neu geladen wird und der Kern nach dem Neustart erneut `http://<solr_server>:8983/solr/#/` .

6. Zeigen Sie den Serverstatus in der Befehlszeile an:

```
sudo service solr status
```

7. Bereiten Sie eine neue Kernkonfiguration für die Kernerstellung in Web Admin vor, indem Sie Ihren Verzeichnispfad in Daten ändern:

```
cd var/solr/data/
```

```

[solr@ip-172-31-28-149 bin]$ pwd
/opt/solr/bin
[solr@ip-172-31-28-149 bin]$ cd ../../..
[solr@ip-172-31-28-149 /]$ pwd
/
[solr@ip-172-31-28-149 /]$ cd var/solr/data/
[solr@ip-172-31-28-149 data]$ ll
total 4
drwxrwxr-x. 4 solr solr  50 Oct 20 22:54 NewCore1
-rw-r-----. 1 solr solr 2117 Oct 20 00:02 solr.xml
[solr@ip-172-31-28-149 data]$ ll NewCore1/
total 8
drwxrwxr-x. 3 solr solr 4096 Oct 20 00:02 conf
-rw-rw-r--. 1 solr solr  78 Oct 20 22:54 core.properties
drwxrwxr-x. 5 solr solr  53 Oct 20 22:54 data

```

8. Hier werden die neuen Kerne gespeichert:

```
ll
```

9. Der Ordner "conf" des neu erstellten Kerns kann als Vorlage verwendet werden:

```
ll NewCore1/
```

10. Erstellen Sie einen Ordner für einen anderen Kern, den Sie in der Webverwaltung erstellen:

```
mkdir CoreFromWebAdmin
```

11. Kopieren Sie das conf-Verzeichnis in den neuen Speicherort:

```
sudo cp -R NewCore1/conf/ CoreFromWebAdmin
```

12. Wechseln Sie in Ihrem Browser zur Web Admin-Oberfläche

13. Klicken Sie auf Core hinzufügen

14. Fügen Sie CoreFromWebAdmin als Namen und Ordner für den neuen Kern hinzu.

The screenshot shows the Solr Admin web interface. On the left is a navigation menu with options: Dashboard, Logging, Core Admin (selected), Java Properties, and Thread Dump. Below the menu is a 'Core Selector' dropdown. The main area displays a modal dialog for adding a new core. The dialog contains the following fields:

- name: CoreFromWebAdmin
- instanceDir: CoreFromWebAdmin
- dataDir: data
- config: solrconfig.xml
- schema: schema.xml

Below the fields is an information icon and a message: "instanceDir and dataDir need to exist before you can create the core". At the bottom of the dialog are "Add Core" and "Cancel" buttons. To the right of the dialog, a terminal window shows the following output:

```

[solr@ip-172-31-...
[solr@ip-172-31-...
total 4
drwxrwxr-x. 4 so
-rw-r-----. 1 so
[solr@ip-172-31-...
total 8
drwxrwxr-x. 3 so
-rw-rw-r--. 1 so
drwxrwxr-x. 5 so
[solr@ip-172-31-...
[solr@ip-172-31-...
eFromWebAdmin/
[sudo] password
[solr@ip-172-31-...
total 4
drwxr-xr-x. 3 ro
[solr@ip-172-31-...
[solr@ip-172-31-...
[solr@ip-172-31-...

```

Below the dialog, a table shows the status of the core:

deletedDocs:	0
optimized:	✓
current:	✓
directory:	org.rockwell.lock

15. Öffne den neuen Kern.

16. Klicken Sie auf Dokumente, um Dokumente hinzuzufügen.

17. Wählen Sie das XML-Format aus und fügen Sie den folgenden Code ein:

```

<add><doc>
  <field name="id">F9V7464-APL-KIT</field>
  <field name="name">Belkin Mobile Power Cord for iPod w/ Dock</field>
  <field name="manu">Belkin</field>
  <!-- Join -->
  <field name="manu_id_s">belkin</field>
  <field name="cat">electronics</field>
  <field name="cat">connector</field>
  <field name="features">car power adapter, white</field>
  <field name="weight">4.0</field>
  <field name="price">19.95</field>
  <field name="popularity">1</field>
  <field name="inStock">>false</field>
  <!-- Buffalo store -->
  <field name="store">45.18014,-93.87741</field>
  <field name="manufacturedate_dt">2005-08-01T16:30:25Z</field>

```



```

</doc>

<doc>
  <field name="id">IW-032</field>
  <field name="name">iPod &amp; iPod Mini USB 2.0 Cable</field>
  <field name="manu">Belkin</field>
  <!-- Join -->
  <field name="manu_id_s">belkin</field>
  <field name="cat">electronics</field>
  <field name="cat">connector</field>
  <field name="features">car power adapter for iPod, white</field>
  <field name="weight">2.0</field>
  <field name="price">11.50</field>
  <field name="popularity">1</field>
  <field name="inStock">>false</field>
  <!-- San Francisco store -->
  <field name="store">37.7752,-122.4232</field>
  <field name="manufacturedate_dt">2006-02-14T23:55:59Z</field>
</doc>

<doc>
  <field name="id">F887464-APL-KIT</field>
  <field name="name">Belkin Mobile Power Cord for iPod w/ Dock</field>
  <field name="manu">Belkin</field>
  <!-- Join -->
  <field name="manu_id_s">belkin</field>
  <field name="cat">electronics</field>
  <field name="cat">connector</field>
  <field name="features">car power adapter, black</field>
  <field name="weight">4.0</field>
  <field name="price">19.95</field>
  <field name="popularity">1</field>
  <field name="inStock">>true</field>
  <!-- Buffalo store -->
  <field name="store">45.18014,-93.87741</field>
</doc>

<doc>
  <field name="id">FAV7464-APL-KIT</field>
  <field name="name">Belkin Mobile Power Cord for iPod w/ Dock</field>
  <field name="manu">Belkin</field>
  <!-- Join -->
  <field name="manu_id_s">belkin</field>
  <field name="cat">electronics</field>
  <field name="cat">connector</field>
  <field name="features">car power adapter, blue</field>
  <field name="weight">4.0</field>
  <field name="price">15.95</field>
  <field name="popularity">2</field>
  <field name="inStock">>true</field>
  <!-- Buffalo store -->
  <field name="store">45.18014,-93.87741</field>
  <field name="manufacturedate_dt">2015-09-21T16:30:25Z</field>
</doc></add>

```

Wenn Ihre Antwort erfolgreich ist, haben Sie SOLR erfolgreich installiert und Ihre Installation überprüft.



- Dashboard
- Logging
- Core Admin
- Java Properties
- Thread Dump
- NewCore1
 - Overview
 - Analysis
 - Dataimport
 - Documents
 - Files
 - Ping
 - Plugins / Stats
 - Query
 - Replication
 - Schema
 - Segments info

Request-Handler (qt)

Document Type

Document(s)

```
<add>
<doc>
  <field name="id">F9V7464-APL-KIT</field>
  <field name="name">Belkin Mobile Power Cord for iPod w/ Dock</field>
  <field name="manu">Belkin</field>
  <!-- Join -->
  <field name="manu_id_s">belkin</field>
  <field name="cat">electronics</field>
  <field name="cat">connector</field>
  <field name="features">car power adapter, white</field>
</doc>
</add>
```

Commit Within

Overwrite

Status: success
Response:
{
 "response":
 "status":
 "queryString":
 }
}

Beispiel für die Solr-Suche



- Dashboard
- Logging
- Core Admin
- Java Properties
- Thread Dump
- demo
- Overview
- Analysis
- Dataimport
- Documents
- Files
- Ping
- Plugins / Stats
- Query
- Replication
- Schema
- Segments info

Request-Handler (qt)

/select

— common —

q

:

fq

publisher_s="Bantam"

sort

pubyear_i desc

start, rows

0 3

fl

title_t, pubyear_i

df

Raw Query Parameters

key1=val1&key2=val2

wt

json

indent

debugQuery

dismax

```
http://localhost:8983/solr/demo/select?fl=title_t, puby
{
  "responseHeader":{
    "status":0,
    "QTime":1,
    "params":{
      "q":":*",
      "indent":"on",
      "fl":"title_t, pubyear_i",
      "fq":"publisher_s=\"Bantam\"",
      "sort":"pubyear_i desc",
      "rows":"3",
      "wt":"json",
      "_":"1498631813496"}},
  "response":{"numFound":5,"start":0,"docs":[
    {
      "pubyear_i":1999,
      "title_t":["A Clash of Kings"]},
    {
      "pubyear_i":1996,
      "title_t":["A Game of Thrones"]},
    {
      "pubyear_i":1992,
      "title_t":["Snow Crash"]}]}
}}
```

Erste Schritte mit solr online lesen: <https://riptutorial.com/de/solr/topic/1015/erste-schritte-mit-solr>

Kapitel 2: Apache Solr

Examples

Solr Installation

Sie können Solr auf jedem System installieren, auf dem eine geeignete Java Runtime Environment (JRE) verfügbar ist (siehe unten). Derzeit sind dies Linux, OS X und Microsoft Windows. Sie benötigen die Java Runtime Environment (JRE) Version 1.8 oder höher. Überprüfen Sie in einer Befehlszeile Ihre Java-Version folgendermaßen:

```
$ java -version
```

Solr ist auf der Solr-Website unter <http://lucene.apache.org/solr/> verfügbar. Extrahieren Sie das Solr-Distributionsarchiv in Ihr lokales Verzeichnis

Solr-Schnellstart (Solar starten und stoppen)

Sie können Solr starten, indem Sie `bin / solr` im Solr-Verzeichnis ausführen. Dadurch wird Solr im Hintergrund gestartet und Port 8983 überwacht.

```
$ bin / solr start
```

Um den Port zu ändern, auf den Solr zugreift, können Sie beim Start den Parameter `-p` verwenden, z. B. .:

```
$ bin / solr start -p 8984
```

Da es sich bei Solr um einen Server handelt, ist es üblicher, ihn im Hintergrund auszuführen, insbesondere unter Unix / Linux. Um Solr im Vordergrund zu starten, führen Sie einfach Folgendes aus:

```
$ bin / solr start -f
```

Wenn Sie Solr im Vordergrund ausführen (mit `-f`), können Sie es mit `Strg-c` stoppen. Wenn Sie jedoch im Hintergrund ausführen, sollten Sie den Befehl `stop` verwenden, z.

```
$ bin / solr stop -p 8983
```

Für den Befehl `stop` müssen Sie den Port angeben, auf den Solr wartet, oder Sie können den Parameter `-all` verwenden, um alle laufenden Solr-Instanzen zu stoppen.

Wenn Sie nicht sicher sind, ob Solr lokal ausgeführt wird, können Sie den Statusbefehl verwenden:

```
$ bin / solr status
```

Apache Solr online lesen: <https://riptutorial.com/de/solr/topic/9635/apache-solr>

Kapitel 3: Fehlerbehebung Solr

Bemerkungen

Bei der Suche nach dem Ziel einer copyField-Anweisung ist der verwendete Analyser-Stack derjenige des copyField-Ziels und nicht die Quelle. Dies ist besonders wichtig bei der Standardkonfiguration, die nach generischen **Textfeldern** sucht.

Examples

Ist mein Inhalt tatsächlich indiziert?

Oft versuchen die Leute, einige Inhalte zu indizieren und dann zu finden. Wenn das erwartete Ergebnis nicht angezeigt wird, versuchen Sie, den gesamten durchgehenden Prozess zu behandeln. Der bessere Weg ist zu sehen, ob der Inhalt tatsächlich in den erwarteten Feldern indiziert ist. Auf diese Weise wird das Problem in zwei Teile aufgeteilt: Indizieren und Suchen.

Der einfachste Weg, um zu überprüfen, was indiziert wird, ist im Bildschirm Schema der Administrator-Benutzeroberfläche (1). Wählen Sie einfach das entsprechende Feld (2) und laden Sie die Term-Informationen (3). Das sind alle indizierten Begriffe in diesem Feld. Die Liste (4) kann ziemlich lang sein, aber Sie können die Anzahl der angezeigten Elemente und / oder den Index nur für das Debugging in ein separates Testfeld ändern.

- Dashboard
- Logging
- Core Admin
- Java Properties
- Thread Dump
- techproducts
- Overview
- Analysis
- Dataimport
- Documents
- Files
- Ping
- Plugins / Stats
- Query
- Replication
- Schema**
- Segments info

2

Field
cat
Copied to
text **x**
Type
string

Unique Key Field
id

Field: cat

Field-Type: org.apache.solr.schema.StrField
Docs: 20

Flags:	Indexed	Tokenized	Stored
Properties	✓		✓
Schema	✓		✓
Index	✓	✓	✓

3

Autoload

10 / 16 Top-Terms

12	electronics
4	currency
3	memory
2	search connector software hard drive graphics card
1	electronics and s electronics and c

Wenn Sie nichts oder keinen erwarteten Inhalt sehen, bedeutet dies normalerweise, dass die Indizierung fehlgeschlagen ist. Wenn Sie den Inhalt hier sehen, aber nicht in der Abfrage, müssen Sie die Problembehandlung durchführen.

Fehlerbehebung Solr online lesen: <https://riptutorial.com/de/solr/topic/1969/fehlerbehebung-solr>

Kapitel 4: INDEXIERUNG DER TABELLE HIVE2 IN SOLR MIT SOLR DIH

Einführung

Diese Dokumentation bietet eine Möglichkeit zum Herstellen einer Verbindung mit hive mithilfe des SOLR Data Import Handler und zum Indizieren der Daten in SOLR. Dies ist eine interessante Dokumentation, da ich sie nicht über das Internet finden konnte.

Der Handler wickelt im Grunde mehr als 80 Millionen Datensätze ab, was eine starke Infrastruktur mit guten CPUs und Arbeitsspeicher erfordert.

Bemerkungen

Gerne können Sie sich mit uns in Verbindung setzen und wir werden versuchen, so viel wie möglich zu helfen.

Examples

Schritte

Wir haben zuerst die Hive2-Gläser bekommen und es durch Java erledigt, um die Konnektivität zu überprüfen. Dann stellten wir fest, dass die Gläser verwendet werden:

1. hadoop-common-2.7.0-mapr-1703.jar
2. hive-common-2.1.1-mapr-1703-r1.jar
3. hive-jdbc-2.1.1-mapr-1703-r1-standalone.jar

Wenn Sie die SOLR Cloud verwenden, müssen diese Jars auf die VM übertragen werden, auf der SOLR installiert ist, und anschließend in solrconfig.xml folgendermaßen referenziert:

Importieren Sie Part in solrconfig.xml

```
<lib dir = "/ users / path_to_folder_with_jar" regex = ". *. jar" />
```

Dann ist dies der wichtigste Teil: Ihre Hive-Verbindungszeichenfolge:

Verbindungsteil

```
<dataConfig> <dataSource name = "ABC" driver = "org.apache.hive.jdbc.HiveDriver" url = "jdbc:hive2: //...connectionString" user = "Benutzername" password = "password" />
```

```
<document name = "collection_name">
```

```
<entity name = "collection_lookup" query = "unique_key als ID aus Tabellename auswählen">
```

</ entity>

</ document>

</ dataConfig>

Schieben Sie die Konfiguration durch den Zookeeper

```
server / scripts / cloud-scripts / zkcli.sh -zkhost host1: 2181, host2: 2181 -cmd upconfig -confname  
configName -confdir server / solr / configsets / folder /
```

Gehen Sie zu [http:// host: 8983 / solr / # / collection_name / dataimport // dataimport](http://host:8983/solr/#/collection_name/dataimport//dataimport), überprüfen Sie dann das Debugging und prüfen Sie zuerst mit 10 oder 20 Datensätzen.

Sie sehen die Daten fließen. PROST !! Ich kann helfen, wenn Sie weiter diskutieren wollen, aber ich gehe davon aus, dass dies tun sollte. Es arbeitet für mich.

INDEXIERUNG DER TABELLE HIVE2 IN SOLR MIT SOLR DIH online lesen:

<https://riptutorial.com/de/solr/topic/10697/indexierung-der-tabelle-hive2-in-solr-mit-solr-dih>

Kapitel 5: Lucene-Abfragesyntax

Examples

Nähe suchen

```
name:"john doe"~1
```

Sucht nach mehreren Begriffen innerhalb einer bestimmten Begriffsentfernung (~ 1), dh es wird Text gefunden, der **john anonymous doe enthält**, aber nicht **john second name doe**

Einfache Suche

```
name:john
```

Sucht nach einem einzelnen Begriff (Joe) in einem einzelnen Feld (Name)

Boolesche Suche

```
+firstname:john +surname:doe
```

Stimmt mit Dokumenten überein, bei denen Vorname John und Nachname Doe ist. + prefix gibt an, dass der Suchbegriff *vorkommen muss* (AND).

```
+firstname:john -surname:doe
```

Stimmt mit Dokumenten überein, bei denen der Vorname John ist und der Nachname kein Doe ist. - prefix zeigt an, dass der Suchbegriff *nicht vorkommen darf* (NOT).

```
+firstname:john surname:(doe bloggs)
```

Stimmt mit Dokumenten überein, bei denen der Vorname john und der Nachname doe oder bloggs ist. Kein Präfix bedeutet, dass der Nachname vorkommen *sollte* (ODER)

Phrasensuche

```
name:"john doe"
```

Sucht nach mehreren Begriffen in einer bestimmten Reihenfolge.

Suchbegriffe verstärken

```
name:(john doe^5)
```

Mit dem ^ -Kennzeichen kann ein Suchbegriff verstärkt werden, um das Relevanzniveau zu erhöhen. Dies bedeutet, dass Dokumente, die *Doe* enthalten, relevanter sind als diejenigen, die *John* enthalten

Platzhaltersuche

name:john*

Mit dem Indikator * können Sie nach dem Suchbegriff *john* eine Wildcard-Suche durchführen, die 0 oder mehr Zeichen enthält. Es werden Dokumente zurückgegeben, die john, johnson, john's, johnny usw. enthalten.

name:do?

Das ? Mit dem Indikator können Sie eine Platzhaltersuche mit einem einzelnen Zeichen im Suchbegriff durchführen und Dokumente mit Doe, Dog, Dot usw. zurückgeben.

Bereichssuche

age:[50 TO 60]

Übereinstimmungen mit Dokumenten, deren Alter zwischen 50 und 60 Jahren einschließlich 50 und 60 liegt

age:{50 TO 60}

Stimmt mit Dokumenten überein, bei denen das Alter zwischen 50 und 60 Jahren (außer 50 und 60) liegt

age:[* TO 60]

Stimmt mit Dokumenten überein, bei denen das Alter 60 oder weniger beträgt

age:[50 TO *]

Stimmt mit Dokumenten überein, bei denen das Alter mindestens 50 beträgt

age:{50 to 60]

Sie können geschweifte und eckige Klammern mischen. Übereinstimmungen mit Dokumenten, bei denen das Alter größer als 50 Jahre ist, aber weniger als oder gleich 60 ist

Join über Kerne

```
{!join from=personid to=id fromIndex=AddressCore}address:Address1
```

Wenn Sie also zwei Kerne haben, die so aussehen:

PersonCore- ID, Name

AddressCore- ID, Adresse, Personen-ID

Dadurch werden alle PersonCore-Dokumente an einer bestimmten Adresse gefunden

Lucene-Abfragesyntax online lesen: <https://riptutorial.com/de/solr/topic/2242/lucene-abfragesyntax>

Kapitel 6: So erstellen Sie einen benutzerdefinierten Feldtyp

Bemerkungen

Hinweise zu benutzerdefinierten Java-Klassenfeldern:

Dies ist ein kleiner Abschnitt eines umfangreichen Artikels, der [in Solr anhand von externen Feldern](#) sortiert wird, um Solr-Dokumente nach benutzerdefiniertem Feldvergleich zu sortieren.

Hinweise zu benutzerdefinierten Feldern, die aus vorhandenen Solr-Feldern erstellt wurden:

Apache hat eine ausführliche Dokumentation zu diesem Thema erstellt - [Grundlegendes zu Analysatoren, Tokenisierern und Filtern](#) .

Examples

Erstellen Sie einen benutzerdefinierten Solr-Feldtyp aus der eigenen benutzerdefinierten Java-Klasse

Schemaänderungen:

Sie müssen in Ihrer Solr-Schemadatei einen neuen Feldtyp definieren. Anschließend können Sie Felder dieses Typs erstellen. Beispielschema-Snippet:

```
<!-- Source: solr/example/.../conf/schema.xml -->
<?xml version="1.0" encoding="UTF-8" ?>
<schema name="adam" version="1.3">
  <types>
    ...
    <fieldType name="rank_t" class="org.apache.solr.schema.ext.RankFieldType"/>
  </types>
  <fields>
    ...
    <field name="rank" type="rank_t" indexed="true" stored="true"/>
  </fields>
  ...
</schema>
```

Java-Klasse für benutzerdefinierten Feldtyp:

```
// Source: src/java/org/apache/solr/schema/ext/RankFieldType.java
package org.apache.solr.schema.ext;

import java.io.IOException;

import org.apache.lucene.document.Fieldable;
```

```

import org.apache.lucene.search.SortField;
import org.apache.solr.response.TextResponseWriter;
import org.apache.solr.schema.FieldType;
import org.apache.solr.schema.SchemaField;
import org.apache.solr.search.ext.RankFieldComparatorSource;

public class RankFieldType extends FieldType {

    @Override
    public SortField getSortField(SchemaField field, boolean top) {
        return new SortField(field.getName(), new RankFieldComparatorSource(), top);
    }

    @Override
    // copied verbatim from GeoHashField method
    public void write(TextResponseWriter writer, String name, Fieldable f) throws IOException
    {
        writer.writeStr(name, f.stringValue(), false);
    }
}

```

Erstellen Sie einen benutzerdefinierten Feldtyp aus verfügbaren Feldtypen

Lassen Sie uns einige theoretische Kenntnisse erwerben, bevor Sie zum Beispiel übergehen. Hier werden drei wichtige Begriffe verwendet: [Analysatoren](#) , [Tokenizer](#) und [Filter](#) . Um ein solches benutzerdefiniertes Feld zu erstellen, müssen Sie ein Analysegerät mit einem Tokenizer und einem oder mehreren Filtern erstellen. Wie [hier](#) erwähnt, können Sie nur einen Tokenizer pro Analysator verwenden. Es gibt jedoch Möglichkeiten, diese Einschränkung zu überwinden.

```

<fieldType name="alphaOnlySort" class="solr.TextField" sortMissingLast="true"
omitNorms="true">
  <analyzer>
    <tokenizer class="solr.KeywordTokenizerFactory"/>
    <filter class="solr.LowerCaseFilterFactory"/>
    <filter class="solr.TrimFilterFactory"/>
    <filter class="solr.PatternReplaceFilterFactory" replace="all" replacement=""
pattern="([a-z])"/>
  </analyzer>
</fieldType>

```

Ein anderes Beispiel:

```

<fieldType name="lowercase_text" class="solr.TextField" positionIncrementGap="150">
  <analyzer>
    <tokenizer class="solr.KeywordTokenizerFactory" />
    <filter class="solr.LowerCaseFilterFactory" />
  </analyzer>
</fieldType>

```

Noch ein Beispiel mit Beschreibung:

```

<fieldType name="text_stem" class="solr.TextField">
  <analyzer>
    <tokenizer class="solr.StandardTokenizerFactory"/>
    <filter class="solr.StandardFilterFactory"/>
  </analyzer>
</fieldType>

```

```
<filter class="solr.LowerCaseFilterFactory"/>
<filter class="solr.EnglishPorterFilterFactory"/>
</analyzer>
</fieldType>
```

Dieses Beispiel beginnt mit dem Standard-Tokenizer von Solr, der den Feldtext in Token aufteilt. Diese Token durchlaufen dann den Standardfilter von Solr, der Punkte aus Akronymen entfernt und einige andere übliche Operationen durchführt. Alle Token werden dann auf Kleinbuchstaben gesetzt, was die Übereinstimmung von Groß- und Kleinschreibung zum Abfragen erleichtert. Der letzte Filter im obigen Beispiel ist ein Stemmerfilter, der den Porter-Stemming-Algorithmus verwendet. Ein Stemmer ist im Wesentlichen ein Satz von Abbildungsregeln, der die verschiedenen Formen eines Wortes auf das Basis- oder Stammwort abbildet, von dem sie abgeleitet sind. In Englisch sind beispielsweise die Wörter "Umarmungen", "Umarmen" und "Umarmt" alle Formen des Stammworts "Umarmung". Der Stemmer wird alle diese Begriffe durch "hug" ersetzen, was indiziert wird. Dies bedeutet, dass eine Abfrage nach "hug" dem Begriff "hugged" entspricht, aber nicht "riesig".

Beispiel für die Verwendung eines solchen benutzerdefinierten Feldes:

```
<field name="keywords" type="text_stem" indexed="true" stored="true" />
```

Liste der verfügbaren Tokenizer-Typen: [Liste der Tokenizer-Typen](#)

Liste der verfügbaren Filtertypen: [Liste der Filtertypen](#)

So erstellen Sie einen benutzerdefinierten Feldtyp online lesen:

<https://riptutorial.com/de/solr/topic/7409/so-erstellen-sie-einen-benutzerdefinierten-feldtyp>

Credits

S. No	Kapitel	Contributors
1	Erste Schritte mit solr	Community , esiprogrammer , Lefty G Balogh , mfatihk , SimplyInk
2	Apache Solr	Pralhad Awasthi
3	Fehlerbehebung Solr	Alexandre Rafalovitch
4	INDEXIERUNG DER TABELLE HIVE2 IN SOLR MIT SOLR DIH	Khalid Imam
5	Lucene-Abfragesyntax	BunkerMentality , Fuu , MatsLindh
6	So erstellen Sie einen benutzerdefinierten Feldtyp	RamenChef , saurav