User and Group Filter Support with LDAP Atn Provider in HiveServer2

- User and Group Filter Support with LDAP
  - Group Membership
    - hive.server2.authentication.ldap.groupDNPattern
    - hive.server2.authentication.ldap.groupFilter
    - hive.server2.authentication.ldap.groupMembershipKey
    - hive.server2.authentication.ldap.groupClassKey
  - User Search List
    - hive.server2.authentication.ldap.userDNPattern
    - hive.server2.authentication.ldap.userFilter
  - Custom Query String
    - hive.server2.authentication.ldap.customLDAPQuery
    - Support for Groups in Custom LDAP Query
  - Order of Precedence

User and Group Filter Support with LDAP

Starting in Hive 1.3.0, HIVE-7193 adds support in HiveServer2 for

- LDAP Group filters
- LDAP User filters
- Custom LDAP Query support.

Filters greatly enhance the functionality of the LDAP Authentication provider. They allow Hive to restrict the set of users allowed to connect to HiveServer2.

See Authentication/Security Configuration for general information about configuring authentication for HiveServer2. Also see Hive Configuration Properties – HiveServer2 for individual configuration parameters discussed below.

Group Membership

This enables HiveServer2 to enforce group membership for users. The authentication request will succeed if the user belongs to one or more of the groups listed in the Hive configuration file. If the user does not belong to at least one of the groups listed, the user authentication fails.

Four configuration parameters support group-membership based authentication:

- `hive.server2.authentication.ldap.groupDNPattern`
- `hive.server2.authentication.ldap.groupFilter`
- `hive.server2.authentication.ldap.groupMembershipKey` (version 2.1.0 via HIVE-13295)
- `hive.server2.authentication.ldap.groupClassKey` (version 2.1.0 via HIVE-13295)

hive.server2.authentication.ldap.groupDNPattern

This value represents a pattern for “distinguishedName” (DN) for groups in the directory. This value could be a single DN if the LDAP Group entities are co-located or could be a ***colon*** separated list of all DN patterns if the groups are scattered across different trees.

Each DN pattern can contain a “%s” in it that will be substituted with the group name (from the group filter) by the provider for group search queries.

Example 1 (single DN):

```
<property>
  <name>
    hive.server2.authentication.ldap.groupDNPattern
  </name>
  <value>CN=%s,OU=Groups,DC=apache,DC=org</value>
</property>
```

This indicates that all LDAP group entries are under the directory root *“OU=Groups,DC=apache,DC=org”*.

The LDAP Authentication Provider replaces the %s with the group name in the LDAP queries to locate the group entry. For example, if a group named “group1” is being queried for, it uses *“CN=group1,OU=Groups,DC=apache,DC=org”*.

Example 2 (two DNs):

```
<property>
  <name>
    hive.server2.authentication.ldap.groupDNPattern
  </name>
  <value>
      CN=%s,OU=Groups,DC=apache,DC=org:uid=%s,CN=Users,DC=apache,DC=org
  </value>
</property>
```

The above pattern advises the LDAPAtnProvider that LDAP group entities can exist in two separate trees in the directory and can have different attributes in their DNs. (Note the colon separator.)

hive.server2.authentication.ldap.groupFilter

This value represents the group name filter that is to be enforced by the LDAPAtnProvider. All individual groups are represented using a ***comma*** separated list. The user MUST belong to one or more of these groups for the authentication request to succeed.

Example:

```
<property>
  <name>hive.server2.authentication.ldap.groupFilter</name>
  <value>group1,group2</value>
</property>
```

hive.server2.authentication.ldap.groupMembershipKey

This value represents the LDAP attribute on group entries in LDAP that indicates its members. (Available starting in version 2.1.0.)

There could be multiple entries for this attribute, one for each of its members. By default, the LDAP authentication provider assumes "member" to search for users. To alter this default, set a value/key for property for the provider to accurately search for group members.

hive.server2.authentication.ldap.groupClassKey

This value represents the LDAP objectClass each of the groups implements in LDAP. By default, the LDAP Authentication provider uses "groupOfNames" in its search for groups. (Available starting in version 2.1.0.)

The properties above assist in correctly finding user-group associations in LDAP.

Example: If the LDAP Group "testGroup" has the following attributes, Hive's LDAP Authentication provider will not be able to find group members. Setting the 2 properties will help with this.

```
dn:uid=testGroup,ou=Groups,dc=domain,dc=com
objectClass: group
objectClass: top
memberUid: uid=testUser1,ou=Users,dc=domain,dc=com
memberUid: uid=testUser2,ou=Users,dc=domain,dc=com
cn: HiveUserGroup
```

```
<property>
  <name>hive.server2.authentication.ldap.groupMembershipKey</name>
  <value>memberUid</value>
</property>
<property>
  <name>hive.server2.authentication.ldap.groupClassKey</name>
  <value>group</value>
</property>
```

User Search List

This enables HiveServer2 to restrict access to a specified list of users. If the user being authenticated is not part of this userlist, access will be denied.

Two configuration parameters support this feature:

- `hive.server2.authentication.ldap.userDNPattern`
- `hive.server2.authentication.ldap.userFilter`

hive.server2.authentication.ldap.userDNPattern

This value represents a pattern for “distinguishedName” (DN) for users in the directory. This value could be a single DN if the LDAP User entities are co-located within a single root or could be a ***colon*** separated list of all DN patterns if the users are scattered across different trees/forests in the directory.

Each DN pattern can contain a “%s” in it that will be substituted with the username (from the user filter) by the provider for user search queries.

Example 1 (single DN):

```
<property>
  <name>
    hive.server2.authentication.ldap.userDNPattern
  </name>
  <value>
    CN=%s,CN=Users,DC=apache,DC=org
  </value>
</property>
```

In the example above, all users are co-located under a single root *“CN=Users,DC=apache,DC=org”*. To search for user “foo”, LDAPAtnProvider attempts to find the user with DN like *“CN=foo,CN=Users,DC=apache,DC=org”*.

Example 2 (two DNs):

```
<property>
  <name>
    hive.server2.authentication.ldap.userDNPattern
  </name>
  <value>
      CN=%s,OU=Users,DC=apache,DC=org:uid=%s,CN=UnixUsers,DC=apache,DC=org
  </value>
</property>
```

The above pattern advises the LDAPAtnProvider that LDAP user entities can exist in two separate trees in the directory and can have different attributes in their DNs. (Note the colon separator.)

hive.server2.authentication.ldap.userFilter

This is a ***comma*** separated list of usernames to grant access to. The Atn provider grants access if the user being authenticated is part of this list, and denies access otherwise.

Example:

```
<property>
  <name>
    hive.server2.authentication.ldap.userFilter
  </name>
  <value>
    hive-admin,hive,hivetest,hive-user
  </value>
</property>
```

Custom Query String

There are several LDAP implementations available for commercial use, with no standard set of attributes within each implementation. If either of the above filters does not meet the requirements for some unforeseen reasons, HiveServer2 can use a user-specified LDAP Query string to execute against the LDAP server. This configured query is expected to return a set of DNs that represent individual users (see below for support for groups). The returned result will then be used to adjudicate a GRANT/DENY decision to the authenticating user. To support this configuration, a new configuration property has been introduced.

hive.server2.authentication.ldap.customLDAPQuery

Example:

```
<property>
  <name>hive.server2.authentication.ldap.customLDAPQuery</name>
  <value><![CDATA[(&(objectClass=person)(|(memberOf=CN=Domain Admins,CN=Users,DC=apache,DC=org)
(memberOf=CN=Administrators,CN=Builtin,DC=apache,DC=org)))]]>
  </value>
</property>
```

The above query returns all users that are members of one of the groups (Domain Admins or Administrators). This offers a lot more flexibility that allows Hive users to customize the LDAP configuration for their implementation of LDAP.

Support for Groups in Custom LDAP Query

Version information

Available starting with Hive 2.1.1 (see HIVE-14513).

It is not always straightforward to be able to author queries that return users. For example, to allow "all users from group1 and group2" to be authenticated, the LDAP query has to return a union of all members of group1 and group2.

One common configuration is that groups contain a list of users:

```
"dn: uid=group1,ou=Groups,dc=example,dc=com",
"distinguishedName: uid=group1,ou=Groups,dc=example,dc=com",
"objectClass: top",
"objectClass: groupOfNames",
"objectClass: ExtensibleObject",
"cn: group1",
"ou: Groups",
"sn: group1",
"member: uid=user1,ou=People,dc=example,dc=com",
```

The query

```
(&(objectClass=groupOfNames)(|(cn=group1)(cn=group2)))
```

will return the entries

```
uid=group1,ou=Groups,dc=example,dc=com
uid=group2,ou=Groups,dc=example,dc=com
```

but there is no means to form a query that would return just the values of "member" attributes. (LDAP APIs allow filtering of the attributes on the result set.)

To allow for such queries to return user DNs for the members of the group instead of the group DN itself, as of Hive release 2.1.1 the LDAP authentication provider will (re)use the configuration property hive.server2.authentication.ldap.groupMembershipKey. This property represents the attribute name that represents the user DN on the Group entry. In the example from above, that attribute is "*member*".

This allows the Hive LDAP authentication provider to specify a query that returns groups and individual users as below (all users of `group1` + the user `user4` will be allowed to authenticate):

```
<property>
  <name>hive.server2.authentication.ldap.customLDAPQuery</name>
  <value><![CDATA[(|(&(objectClass=groupOfNames)(cn=group1))(&(objectClass=person)(sn=user4)))]]>
  </value>
</property>
```

Order of Precedence

The group membership parameters can be used in conjunction with the user lists to enforce a stricter access. The LDAP Atn provider adjudicates authentication decisions according to the following criteria:

1. If hive-site.xml contains *"hive.server2.authentication.ldap.customLDAPQuery"*, ***only*** the results of this query are used to adjudicate an authentication decision. All other property values (*.groupDNPattern, *.groupFilter, *.userDNPattern, *.userFilter) are ignored entirely.
2. If the *.groupFilter and *.userFilter parameters are both specified in the configuration file, access is granted ***if and only if*** the user being authenticated satisfies ***both*** conditions; access is denied otherwise. So the user has to be listed in the *.userFilter ***and*** the user MUST also belong to one of the groups listed in the *.groupFilter.
3. If only one of the filters ((*.userDNPattern + *.userFilter) || (*.groupDNPattern + *.groupFilter)) is specified, a decision is adjudicated based on whether the user being authenticated satisfies the specified filter.
4. If neither *.groupFilter nor *.userFilter is specified in the configuration file, the provider attempts to search for the user in the LDAP directory within the *baseDN* tree. Access is granted if user has been found, and denied otherwise. **IMPORTANT:** This implementation is a little more stringent compared to the prior implementation. In the prior implementation, if the baseDN was not provided, authentication would be granted if the provider is able to bind to LDAP with the user