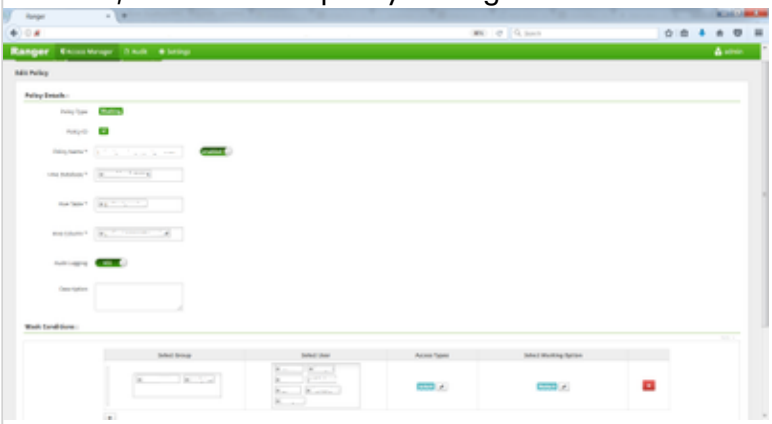
Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

Apache Hive | Apache Ranger | Hortonworks Data Platform (HDP)

navendugarg New Contributor
Created on 02-09-2017 02:24 AM - edited 08-18-2019 04:09 AM

I am running a HDP 2.5.3 cluster with HiveServer2 enabled on an edge node. Ranger is enforcing security. In addition, I've created a policy in ranger that redacts a certain column with pii on select.



Unfortunately, when I execute the following query in a hive client:
select * from <tablename> limit 100;
I get the following error:
[42000]: Error while compiling statement: FAILED: SemanticException org.apache.hadoop.hive.ql.parse.ParseException: line 1:62 rule Identifier failed predicate: {allowQuotedId()}? line 1:74 rule Identifier failed predicate: {allowQuotedId()}? line 1:94 rule Identifier failed predicate: {allowQuotedId()}? line 1:117 rule Identifier failed predicate: {allowQuotedId()}?
If I disable the above policy, the query runs successfully.

Reply

2,639 Views

0 Kudos

Tags (5)

Tags: hdp-2.5.0 hiveserver2 masking Ranger Security

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

ahallam Rising Star
Created 02-09-2017 08:40 AM

which user you are using to run the query ? only the allowed users can do the select * because it includes the the column in the policy.

Reply

844 Views

0 Kudos

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

navendugarg New Contributor
Created 02-10-2017 05:36 PM

I am logged-in as a user that is in the "select user". So yes, this user has permissions to run the query.

Reply

844 Views

0 Kudos

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

aleekha Contributor
Created 02-14-2017 11:37 PM

@Navendu Garg: I tried your scenario in my test environment and had no issues accessing column with masking option set to 'redact'. On giving select cmd, the values were shown as xxxxx instead of usual values. Also, based on https://issues.apache.org/jira/browse/HIVE-6013, I inserted values into my column with ` and ', also set column name with ` and ' but the behavior was fine.
What hive client are you using?
Also, can you let me know what kind of value was in your column? Did the column name or any of column values contain any kind of quotes?
I am assuming that you have an access level hive policy granting ur user atleast select privilege to the table and the columns you are accessing in ur select statement, and then a masking policy which masks one of the columns for the same user running the query.

Reply

844 Views

6 Kudos

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

navendugarg New Contributor

Created 02-15-2017 02:20 PM

I've tried it on SQL Workbench and Aginity Workbench, and I get errors on both clients.

Column is a varchar(9).

Here is what I discovered. A table with column names that have spaces (e.g. `account year`), this feature fails. When I added an underscore to the column (account_year), the masking started to work.

Reply

844 Views

1 Kudo

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

aleekha Contributor

Created 02-15-2017 07:47 PM

I tried the above scenario and @Navendu Garg's exact issue is described below:

There is a table say 'default.testtable1' with one of the column names containing a space say 'account name'.

[create table testtable1 (Name String,`Account number` String,Age int);

...log/webhcat/webhcat.log.2017-02-14:hive.support.quoted.identifiers=column]

A user say 'user1' has a Hive access policy granting all permissions to user1 on default.testtable1, all columns.

There is also a masking policy which masks column 'name' in same table 'default.testtable1' with mask type='redact' for user1.

select * from default.testtable1 throws:

```
Error: Error while compiling statement: FAILED: SemanticException org.apache.hadoop.hive.ql.parse.ParseExcept
line 1:56 rule Identifier failed predicate: {allowQuotedId()}?
line 1:129 rule Identifier failed predicate: {allowQuotedId()}?
line 1:144 rule Identifier failed predicate: {allowQuotedId()}? (state=42000,code=40000)
```

Audit shows only masking policy allowed the txn. There is no deny and no access policy mentioned in the audit. The flow didn't reach the access policies auth.
On disabling the masking policy, the select shows correct results.
NOTE: I tried it with mask type as 'Hash','Partial mask','Nullify' and same result.
@Madhan Neethiraj, Will you pls confirm if this is expected - If any of the column names in a table contain spaces and there is a mask policy on any of that table's columns, then select with fail with the above exception.

Reply

844 Views

7 Kudos

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

mneethiraj Contributor
Created 02-15-2017 10:19 PM

I think Hive folks should look into this one.
CC: @Eugene Koifman, @Wei Zheng

Reply

844 Views

0 Kudos

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

ekoifman Expert Contributor
Created 02-15-2017 10:42 PM

@Navendu Garg, could you share your table definition (show create table T)

Reply

844 Views

0 Kudos

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

aleekha Contributor
Created 02-22-2017 05:50 AM

Hi @Eugene Koifman: I was able to reproduce the exact scenario that Navendu mentioned and table create cmd I used was:
[create table testtable1 (Name String,`Account number` String,Age int);
...log/webhcat/webhcat.log.2017-02-14:hive.support.quoted.identifiers=column]

Reply

844 Views

5 Kudos

Re: Querying a Hive Table (via Hiveserver2) with Column Masking enabled via Ranger Hive Plugin returns with an error.

shaikidris_ali New Contributor
Created 10-11-2017 12:54 PM

I think its a bug in ranger, where quoted.indentifiers are not being honoured.
To reproduce:

- set hive.support.sql11.reserved.keywords=false
- create test-user who has restricted access to some rows in ranger (row filter) say, bangalore filter for service_city column.
- create a table with one of the column is reserved word like "date"

```
select date from temp_db.test limit 10; //or
select `date` from temp_db.test limit 10;
```

- Resolved query by Ranger:

```
select date from (SELECT *, BLOCK__OFFSET__INSIDE__FILE, INPUT__FILE__NAME, ROW__ID FROM `temp_db`.`test`  WH
```

- Exception:

```
2017-10-09 13:28:27,517 ERROR [HiveServer2-Handler-Pool: Thread-60373]: ql.Driver (SessionState.java:printEr
```

Workaround:
>hive.support.quoted.identifiers=none;
Fire query> select `date` from temp_db.test limit 10;
Note: select * works without any issues even though the table has columns with reserved words.

Reply

844 Views

0 Kudos

Powered by Khoros