

🔍 Aggregategroups Action API module allows deleting translatable page metadata for any group without trace (CVE-2021-36129)

🔒 Closed, ResolvedPublicSECURITY

Actions

Assigned To

abi_

Authored By

Nikerabbit

2021-05-15 08:22:24 (UTC+0)

Tags

👤 Security-Team (Our Part Is Done)

👤 Security

👤 Data-Persistence (work done)

👤 Data-Persistence-Backup (Done)


👤 MediaWiki-extensions-Translate (group management)


👤 Patch-For-Review

👤 SecTeam-Processed (Completed)

🔍 MW-1.37-notes (1.37.0-wmf.14; 2021-07-12)

Referenced Files

 F34479703: 01-T282932.patch
2021-06-03 20:34:15 (UTC+0)

 F34457802: T282905.patch
2021-05-17 15:14:09 (UTC+0)

Subscribers

abi_

Aklapper

Base

jcrespo

LSobanski

Marostegui

Nikerabbit

sbassett

Description

While investigating **T282905: Translate syntax version update and translation-aware transclusion lost** I audited all code manipulating `translate_metadata` table. There is a small utility class `TranslateMetadata`. All writes to this table go through that class (there are a couple places where the table is read directly): https://codesearch.wmcloud.org/search?q=translate_metadata&i=nope&files=&excludeFiles=&repos=

There is `TranslateMetadata::deleteGroup` that is only called from `ApiAggregateGroups`. That API module does not validate the parameter for `aggregategroup` when `action=remove`. Only restriction is that this module requires the `translate-manage` right. We have little above 200 people with that right on MetaWiki: <https://meta.wikimedia.org/w/index.php?title=Special:ListUsers&offset=&limit=500&group=translationadmin>

I was able to confirm that deletion of any group's metadata is possible using this module. The code has innocent looking `// @todo Logging`, which is really unfortunate in this case, because I can't for sure say that this is or is not the cause for the parent issue.





It seems the most likely cause, however, by ruling out other possibilities:

- (1) The value was never saved in the first place. This is implausible because we have the log entries, other metadata in other tables and nothing in Logstash. Why would writes to this table silently fail, but not the other writes? In addition `achive.org` seems to confirm that the metadata was saved and was effective at some point.
- (2) The value was deleted by calling `TranslateMetadata::set` with `false` as the value. When marking the page for translation, type declarations would prevent false for appearing for all the different `tmd_key` values.
- (3) Page is moved or deleted. This would certainly leave log entries and other visible traces.

I think it would be good to silently patch `do=remove` for `action=aggregategroups` to do nothing but add logging that would identify the attacker (if any).

It would be really helpful if we could find DELETE statements for `translate_metadata` from binlogs/query logs (if we have any). This would help to find the extent of this issue and hopefully confirm the cause too.

Details


Author Affiliation WMF Product	
Project	Subject
 mediawiki/extensions/Translate	SECURITY: Enhance validation and logging for AggregateGroups API deletions
 mediawiki/extensions/Translate	SECURITY: Enhance validation and logging for AggregateGroups API deletions
 mediawiki/extensions/Translate	SECURITY: Enhance validation and logging for AggregateGroups API deletions
 mediawiki/extensions/Translate	SECURITY: Enhance validation and logging for AggregateGroups API deletions
Customize query in gerrit	

Related Objects

Mentions	
Mentioned In	
T279733: Write and send supplementary release announcement for extensions and skins with security patches (1.31.15/1.35.3/1.36.1)	


Mentioned Here

~~#279733~~ Write and send supplementary release announcement for extensions and skins with security patches (1-31-15/1-35-3/1-36-1)
~~#202205~~ Translate syntax version update and translation-aware transclusion lost


 **Nikerabbit** created this task. 2021-05-15 08:22:24 (UTC+0)

  Restricted Application added a subscriber: **Aklapper**. · View Herald Transcript 2021-05-15 08:22:25 (UTC+0)

 **Marostegui** added a project: **Data-Persistence (work done)**. 2021-05-15 09:14:07 (UTC+0)

 **Marostegui** added a subscriber: **LSobanski**.

If you have some concrete dates and a wiki, we can check binlogs and backups to see the status of the data for that given wiki.


 **Nikerabbit** added a comment. 2021-05-15 09:45:11 (UTC+0)

With the Wikimedia CEE Spring 2021 page the deletion of the rows (where translate_metadata.tmd_group='page-Wikimedia CEE Spring 2021') would need to have happened between 2021-05-01T23:30:30Z and today on metawiki. Is that too long of a period?

 **Marostegui** added a subscriber: **jcrespo**. 2021-05-15 10:38:29 (UTC+0)

We could probably scan the backups from that day till the last one and see on which one the page exists and on which one the page is no longer there. Once we have that narrowed, we could go ahead and check the binlogs around those days to find out exactly when it was deleted.

Otherwise it will be hard to scan all those generated binlogs.

 **@jcrespo** could you help out here?

 **Marostegui** added a project: **Data-Persistence-Backup**. 2021-05-15 10:38:44 (UTC+0)


 **Reedy** added a project: **MediaWiki-extensions-Translate**. 2021-05-16 15:23:43 (UTC+0)

 **jcrespo** added a comment. 2021-05-17 09:36:35 (UTC+0)

Hey,  **@Nikerabbit** could you please confirm the request, as I think I understand the overall issue, but want to have 100% clear the request.

- There were (presumably, that is the question) some DELETE statements ran against the s7 metadata table metawiki.translate_metadata and you want to know any possible details about them. In particular -but there could be more- one was presumably run between 2021-05-01T23:30:30Z and 2021-05-16T00:00:00Z. If I run `select count(*) FROM translate_metadata where translate_metadata.tmd_group='page-Wikimedia CEE Spring 2021'` I get 0 results, but at some point between these 2 dates it should return some results. You want to know when that happened, if it did, and the context (e.g. api call run/any other information about related queries in the transaction).

 **Nikerabbit** added a comment. 2021-05-17 10:43:34 (UTC+0)

 **@jcrespo** Yep that sounds correct. If that query returns results at some time between those dates, we can rule out my possible cause number 1. If you find details of a DELETE query, I might be able to rule out other possible cases. This table should have very few DELETE queries overall, it's mostly updated with REPLACE (insert if missing) queries.

 **jcrespo** added a comment. Edited · 2021-05-17 11:04:22 (UTC+0)


```
live results:
mysql.py -h db1116:3317 metawiki -e "SELECT * FROM translate_metadata ORDER BY tmd_group, tmd_key" | grep 'page-Wikimedia CEE Spring 2021' ~> 0 results
backup 11 May:
zgrep 'page-Wikimedia CEE Spring 2021' dump.s7.2021-05-11--00-00-02/metawiki.translate_metadata.sql.gz ~> 0 results
backup 4 May:
zgrep 'page-Wikimedia CEE Spring 2021' dump.s7.2021-05-04--02-24-36/metawiki.translate_metadata.sql.gz ~> 3 results
("page-Wikimedia CEE Spring 2021", "maxid", "31"),
("page-Wikimedia CEE Spring 2021", "transclusion", "1"),
("page-Wikimedia CEE Spring 2021", "version", "2"),
```

So a delete, replace or update happened between (approximately, the backup dates are not accurate to the second), between 2021-05-11 00:00:02 and 2021-05-04 02:24:36. In binlog coordinates that is between db1136-bin.002276:170694204 and db1136-bin.002293:1027651177.


The following events were obtained from grepping the binlogs:

```
mysqlbinlog db1136-bin.002286 | grep -C 10 'page-Wikimedia CEE Spring 2021'
COMMIT/*!*/;
# at 893088481
#210508 6:46:37 server id 171978861 end_log_pos 893088519 GTID 171978861-171978861-202827326 trans
/*!100001 SET @@session.gtid_seq_no=202827326/*!*/;
BEGIN
/*!*/;
# at 893088519
#210508 6:46:37 server id 171978861 end_log_pos 893088747 Query thread_id=938197070 exec_time=0 error_code=0
use 'metawiki'/*!*/;
SET TIMESTAMP=1620456397/*!*/;
REPLACE /* ApiGroupReview::changeState */ INTO `translate_groupreviews` (tgr_group,tgr_lang,tgr_state) VALUES ('page-Wikimedia CEE Spring 2021','tr','progress')
/*!*/;
# at 893088747
#210508 6:46:37 server id 171978861 end_log_pos 893088775 Intvar
SET INSERT_ID=41213765/*!*/;
# at 893088775
```


Which points to deletes happening on May 8, 2021 7:49:11 AM (actual commit took effect 2 seconds later) from the function TranslateMetadata::set


 **jcrespo** added a comment. 2021-05-17 11:11:41 (UTC+0)

In ~~#202932#7092272~~,  **@Nikerabbit** wrote:


 **@jcrespo** Yep that sounds correct. If that query returns results at some time between those dates, we can rule out my possible cause number 1. If you find details of a DELETE query, I might be able to rule out other possible cases. This table should have very few DELETE queries overall, it's mostly updated with REPLACE (insert if missing) queries.

Please report asap if you will require some kind of data recovery- the more time it passes, the harder it gets. Recovering data from last week: very easy, from last month: easy, from last 3 months: possible, from over 3 months: probably impossible

 **Nikerabbit** added a comment. 2021-05-17 12:18:54 (UTC+0)

 **@jcrespo** Thanks a lot for this! It seems we can rule out exploitation of this API as the cause. Unfortunately it points to a bug in our code which we have not yet identified. I imagine it will be hard to selectively restore from backups, and given this metadata is not super essential, it is probably not worth it, but I'll let you know asap if I change my mind.

▼

 **T282905.patch** 1 KB
Download

 **Nikerabbit** moved this task from **Backlog** to **group management** on the **MediaWiki-extensions-Translate** board. 2021-05-24 10:37:34 (UTC+0)

It's low impact, since it requires `translate-manage` right to be exploited. It is not currently being exploited to my knowledge.

 sbassett moved this task from **Watching** to **Security Patch To Deploy** on the **Security-Team** board.

Patch looks good to me. There are a few code style issues which can be fixed later.

I'll go ahead and deploy this now and keep an eye on logstash for any unexpected errors. If someone with more knowledge of the extension, appropriate rights, etc. could further test the patch in production, and confirm its efficacy, that would be great. Also - updated patch with new subject, bug, author and correct naming convention for production deployment:

 **01-T282932.patch** 1 KB
Download

The above patch has been [deployed](#) to wmf.7. Logstash errors seem fine, in that there do not currently appear to be any obvious, related errors from the patch. Again, if someone with better knowledge of the extension and appropriate rights could perform some additional UAT, that'd be appreciated.

🔒 sbassett changed the edit policy from "Custom Policy" to "All Users".


<https://qerrit.wikimedia.org/r/702760>

<https://gerrit.wikimedia.org/r/702718>

<https://gerrit.wikimedia.org/r/702719>




Change 702720 had a related patch set uploaded (by SBassett; author: Abijeet Patro):
[mediawiki/extensions/Translate@REL1_31] SECURITY: Enhance validation and logging for AggregateGroups API deletions
<https://gerrit.wikimedia.org/r/702720>


 gerritbot added a comment. 2021-07-02 15:46:49 (UTC+0)

Change 702760 **merged** by jenkins-bot:
[mediawiki/extensions/Translate@master] SECURITY: Enhance validation and logging for AggregateGroups API deletions
<https://gerrit.wikimedia.org/r/702760>


 ReleaseTaggerBot added a project: ~~MW-1.37-notes (1.37.0-wmf.14, 2021-07-12)~~ 2021-07-02 16:00:17 (UTC+0)

 gerritbot added a comment. 2021-07-02 16:09:15 (UTC+0)


Change 702719 **merged** by jenkins-bot:
[mediawiki/extensions/Translate@REL1_35] SECURITY: Enhance validation and logging for AggregateGroups API deletions
<https://gerrit.wikimedia.org/r/702719>


 gerritbot added a comment. 2021-07-02 16:10:05 (UTC+0)


Change 702720 **merged** by jenkins-bot:
[mediawiki/extensions/Translate@REL1_31] SECURITY: Enhance validation and logging for AggregateGroups API deletions
<https://gerrit.wikimedia.org/r/702720>

 gerritbot added a comment. 2021-07-02 16:11:50 (UTC+0)

Change 702718 **merged** by jenkins-bot:
[mediawiki/extensions/Translate@REL1_36] SECURITY: Enhance validation and logging for AggregateGroups API deletions
<https://gerrit.wikimedia.org/r/702718>

 sbassett renamed this task from *Aggregategroups Action API module allows deleting translatable page metadata for any group without trace* to *Aggregategroups Action API module allows deleting translatable page metadata for any group without trace (CVE-2021-36129)*. 2021-07-02 20:02:54 (UTC+0)

 sbassett closed this task as *Resolved*. 2021-07-02 20:13:49 (UTC+0)

 sbassett moved this task from *Watching to Our Part Is Done* on the *Security-Team* board.