



Toolforge webservices are in the final stages of [migrating to the toolforge.org domain](#).  
Please help us clean up older documentation referring to [tools.wmflabs.org](#)!

# Incident documentation/20171017-webm-audio-mediatype

[< Incident documentation](#)

## Summary

audio/webm files were incorrectly labeled as `img_media_type="VIDEO"` & `img_major_mime="video"`. After having fixed the cause (<https://phabricator.wikimedia.org/T156135>), there were a few old entries in the database that had to be fixed (`image`: 30, `oldimage`: 4, `filearchive`: 94). Because of the small amount of rows to be updated, `mlitn` misjudged the impact.

## Timeline

`mlitn` logs in to tin & connects to mysql commonswiki.

~12:58: `mlitn` runs

```
UPDATE image SET
img_media_type="AUDIO", img_major_mime="audio" WHERE img_media_type="VIDEO"
AND img_major_mime="video" AND img_minor_mime="webm" AND img_metadata LIKE
's:s:9:"mime_type";s:10:"audio/webm";%';
Query OK, 30 rows affected (0.35 sec)
Rows matched: 30  Changed: 30  Warnings: 0
```

~13:17: `mlitn` runs

```
UPDATE oldimage SET
oi_media_type="AUDIO", oi_major_mime="audio" WHERE oi_media_type="VIDEO" AND
oi_major_mime="video" AND oi_minor_mime="webm" AND oi_metadata LIKE
's:s:9:"mime_type";s:10:"audio/webm";%';
Query OK, 4 rows affected (7.45 sec)
Rows matched: 4  Changed: 4  Warnings: 0
```

~13:20: `mlitn` runs

```
UPDATE filearchive SET
fa_media_type="AUDIO", fa_major_mime="audio" WHERE fa_media_type="VIDEO" AND
fa_major_mime="video" AND fa_minor_mime="webm" AND fa_metadata LIKE
's:s:9:"mime_type";s:10:"audio/webm";%';
Query OK, 94 rows affected (8.46 sec)
Rows matched: 94  Changed: 94  Warnings: 0
```

~13:25: `zeljkof` & `jynus` notice lag on s4 (<https://grafana.wikimedia.org/dashboard/db/mysql?orgId=1&var-dc=eqiad%20prometheus%2Fops&var-server=db1053&var-port=9104&from=1508243685418&to=1508248037877>)

~13:25: `mlitn` forgot to `!log` what he was doing: there was confusion, and execution of another (unrelated) script was halted as a result

As a general comment, the reason why it has so much impact was because, despite the number of rows being very low, the actual number of rows read was very high (probably most or a significant part of the table, which not only was read, but also blocked, preventing other writes to go on, and blocking replication on the non-masters. There are scripts such as `pt-archiver` that assure writes are done by primary key on very small batches, ensuring minimal impact (e.g. for deleting >30M rows on commonswiki <https://phabricator.wikimedia.org/T177772#3670119>).

Main page  
Recent changes  
Server admin log (Prod)  
Server admin log (RelEng)  
Deployments  
SRE/Operations Help  
Incident status

Cloud VPS & Toolforge  
Cloud VPS documentation  
Toolforge documentation  
Request Cloud VPS project  
Server admin log (Cloud VPS)

Tools  
What links here  
Related changes  
Special pages  
Permanent link  
Page information  
Cite this page

Print/export  
Create a book  
Download as PDF  
Printable version

## Conclusions

- Scripting should properly have been done instead of running from CLI
- DBAs should have been pinged, who are in a better position to judge impact & advise
- `!log` should have been used to warn what was about to be made
- Before performing database edits- backups should be performed to allow easy reversions
- Should dangerous maintenance scripts executions be limited/managed differently/documented better?

Category: [Incident documentation](#)

This page was last edited on 13 November 2017, at 16:08.

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. See [Terms of Use](#) for details.

[Privacy policy](#) [About](#)

[Disclaimers](#) [Code of Conduct](#) [Developers](#) [Statistics](#) [Cookie statement](#) [Mobile view](#)

[Wikitech](#)

