

## GDP - Feature #46

### Need way of rebuilding log indices ("fsck for logs")

08/22/2016 09:06 AM - Eric Allman

<b>Status:</b> Closed	<b>Start date:</b> 08/22/2016
<b>Priority:</b> High	<b>Due date:</b>
<b>Assignee:</b> Eric Allman	<b>% Done:</b> 0%
<b>Category:</b> log-check	<b>Estimated time:</b> 0.00 hour
<b>Target version:</b>	
<b>Description</b> Since the tidx (timestamp index) is relatively new, many logs don't have the index necessary to make "query by timestamp" work. Also, some of the timestamp indices have somehow been corrupted. An application to rebuild the indices from the raw log data would be useful. It's a bit tricky if we want it to be able to run while gdplogd is running: somehow some mechanism to inform gdplogd to re-open logs will be required. Alternatively, it might be necessary to take gdplogd down, rebuild the indices, and then restart it (this could possibly happen as part of gdplogd startup).	
<b>Related issues:</b> Related to GDP - Feature #48: Scanning for corrupt indices at gdplogd startup <b>Closed</b> <b>08/23/2016</b>	

#### History

##### #1 - 08/24/2016 10:44 AM - Eric Allman

See Issue [#47](#) for a brief discussion of the issues.

##### #2 - 08/24/2016 02:46 PM - Eric Allman

- Priority changed from Normal to High

We have some known corrupt logs right now, so this really needs to get fixed.

##### #3 - 08/30/2016 11:11 AM - Nitesh Mor

- Related to Feature #48: Scanning for corrupt indices at gdplogd startup added

##### #4 - 09/02/2016 04:39 PM - Eric Allman

- Status changed from New to In Progress

A version of gdp-log-check has been checked in, but it isn't well tested yet. I recommend against trying it in production until it is better tested.

Notably, although it does lock basic data structures, it has no way of signalling to a running gdplogd that it has made changes. For this reason, it should *never* be run while gdplogd is running.

##### #5 - 09/09/2016 02:43 PM - Eric Allman

- Category changed from gdplogd to log-check

- Status changed from In Progress to Feedback

There's a version of gdp-log-check in the apps directory that at least does the basics. I'm putting this into Feedback state in the hopes of getting feedback (duh).

In the process I found that duplicates and gaps in logs didn't work and in fact left the log indices in pretty curdled state. This may affect Kaz as well. I'm adding him (and Nitesh and Ken) as watchers on this issue so they know what's going on.

Briefly, "gdp-log-check logname" will check a log for index consistency, and "gdp-log-check -r logname" will rebuild the indices. See the man page (gdp-log-check.8) for more information.

**#6 - 11/11/2016 04:25 PM - Eric Allman**

- *Status changed from Feedback to Closed*

I'm closing this issue since I think a preliminary program is available that works acceptably. Enhancements should be separate issues.