

Grouper data cleanup

Wiki Home	Grouper Release Announcements	Grouper Guides	Grouper Deployment Guide	Community Contributions	Internal Developer Resources
---------------------------	---	--------------------------------	--	---	--

This page presents suggestions for ongoing Grouper administration tasks, including:

[Pruning logs and registry](#)
[Monitoring Grouper Functions](#)
[Setting up Notifications](#)

Pruning Logs and Registry

[change log](#)
[daemon logs](#)
[user audit logs](#)
[point in time logs](#)
[registry](#)

Pruning the Change Log

When updates are made in Grouper (group add, membership add, etc), updates are also made to the temp change log (grouper_change_log_entry_temp) as part of the same transaction. The Grouper Daemon later (by default every minute), moves these changes to grouper_change_log_entry with sequence numbers and also adds flattened notifications and permission notifications.

By default, change log entries are retained for 14 days in grouper_change_log_entry. This is configurable in grouper-loader.properties. Note that this option doesn't exist before Grouper 2.0.

```
# number of days to retain db rows in grouper_change_log_entry.  -1 is forever.  default is 14
loader.retain.db.change_log_entry.days=14
```

Pruning the Daemon logs

When the [Grouper Daemon](#) runs jobs, logs are kept in the grouper_loader_log table with information about the job, such as when it started, ended, and status. Note that these logs don't just get created for the jobs that load data into Grouper, but also for the other jobs that are run by the Grouper Daemon, such as the job to populate grouper_change_log_entry which runs every minute.

By default, these logs are retained for 7 days. This is also configurable in grouper-loader.properties.

```
# number of days to retain db logs in table grouperloader_log.  -1 is forever.  default is 7
loader.retain.db.logs.days=7
```

Pruning the User Audit Logs

[User Audit](#) logs are created when updates are made in Grouper. These logs are stored in the grouper_audit_entry table. For Grouper 2.0, I think the only option for pruning these logs are by deleting them directly from the database.

Here's an example to delete rows based on time. Note that created_on is milliseconds from epoch.

```
delete from grouper_audit_entry where created_on < '1318599550575'
```

Pruning the Point in Time Logs

Grouper has several tables that have names starting with "grouper_pit_" that are used to store point in time data. These tables are populated by the Grouper Daemon. If an object has been deleted from Grouper, the object remains in the point in time tables with an end date. Using GSH, you can delete objects in the point in time tables that have end dates. See [Point in Time Auditing](#).

Pruning the Registry

You may need to delete old groups, e.g. if you don't need to keep course groups with unlimited history. At Penn we only need courses for the current semester, next semester, and previous semester. This might be as easy as an obliterate GSH command (note you can generate a report first), or a SQL query that generates a GSH script. See the [GSH documentation](#) for details

See Also

[Grouper Diagnostics](#)