

Schedules

Schedules allow to schedule crawls to run on a regular basis based on a Cron-like expression.

Schedules run on the main manager. When a schedule fires, it is an equivalent of the manager receiving a start request. Anything that is running already is ignored. Schedules may cause multiple seeds to start. Schedules are loaded when the manager becomes main manager (either at start-up or when a manager becomes main because of a failure) and updated by incoming REST requests. There is a chance that schedules will not "fire" if their scheduled execution time is during the election of a new main manager following a failure. This is deemed to be acceptable.

Scheduling is flexible:

- Allowing multiple seeds to be started by a single schedule
- Allowing schedules to chain actions (ie crawl seed 1, then crawl seed 2)

There are two ways of configuring Schedules:

1. Via UI [UI - Schedules](#)
2. Via REST [Schedules REST API](#)