

Item 290

git_comments:

git_commits:

1. **summary:** Merge pull request #273 from smartive/NUTCH-2493
message: Merge pull request #273 from smartive/NUTCH-2493 NUTCH-2493: Add new parameter to crawler script to configure sitemap processing frequency

github_issues:

github_issues_comments:

github_pulls:

1. **title:** NUTCH-2493: Add new parameter to crawler script to configure sitemap processing frequency
body: While using the crawler script with the sitemap processing feature introduced in #271 I encountered some performance issues when working with large sitemaps. Therefore one should be able to specify if sitemap processing based on HostDB should take place and if so how frequently it should be done.
2. **title:** NUTCH-2493: Add new parameter to crawler script to configure sitemap processing frequency
body: While using the crawler script with the sitemap processing feature introduced in #271 I encountered some performance issues when working with large sitemaps. Therefore one should be able to specify if sitemap processing based on HostDB should take place and if so how frequently it should be done.
3. **title:** NUTCH-2493: Add new parameter to crawler script to configure sitemap processing frequency
body: While using the crawler script with the sitemap processing feature introduced in #271 I encountered some performance issues when working with large sitemaps. Therefore one should be able to specify if sitemap processing based on HostDB should take place and if so how frequently it should be done.
label: code-design
4. **title:** NUTCH-2493: Add new parameter to crawler script to configure sitemap processing frequency
body: While using the crawler script with the sitemap processing feature introduced in #271 I encountered some performance issues when working with large sitemaps. Therefore one should be able to specify if sitemap processing based on HostDB should take place and if so how frequently it should be done.

github_pulls_comments:

1. I am +1 here, tested on local test crawl.
2. +1, patch looks good

github_pulls_reviews:

jira_issues:

1. **summary:** Add configuration parameter for sitemap processing to crawler script
description: While using the crawler script with the sitemap processing feature introduced in NUTCH-2491 I encountered some performance issues when working with large sitemaps. Therefore one should be able to specify if sitemap processing based on HostDB should take place and if so how frequently it should be done.
2. **summary:** Add configuration parameter for sitemap processing to crawler script
description: While using the crawler script with the sitemap processing feature introduced in NUTCH-2491 I encountered some performance issues when working with large sitemaps. Therefore one should be able to specify if sitemap processing based on HostDB should take place and if so how frequently it should be done.
3. **summary:** Add configuration parameter for sitemap processing to crawler script
description: While using the crawler script with the sitemap processing feature introduced in NUTCH-2491 I encountered some performance issues when working with large sitemaps. Therefore one should be

