Item 294
git_comments:

git_commits:

1. **summary:** [GOBBLIN-1303] Catch error when failing to clean staging data
   message: [GOBBLIN-1303] Catch error when failing to clean staging data Closes #3141 from jack-moseley/cleanup-failure
   label: code-design

github_issues:

github_issues_comments:

github_pulls:

1. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data
   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"
2. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data
   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"
3. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data
   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative

mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"

4. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data
   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"

5. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data
   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"

6. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data
   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"
   label: code-design

7. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data
   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit

message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"

8. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data

   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"

9. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data

   body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"

10. **title:** [GOBBLIN-1303] Catch error when failing to clean staging data

    body: Dear Gobblin maintainers, Please accept this PR. I understand that it will not be reviewed until I have checked off all the steps below! ### JIRA - [x] My PR addresses the following [Gobblin JIRA] (https://issues.apache.org/jira/browse/GOBBLIN/) issues and references them in the PR title. For example, "[GOBBLIN-XXX] My Gobblin PR" - https://issues.apache.org/jira/browse/GOBBLIN-1303 ### Description - [x] Here are some details about my PR, including screenshots (if applicable): Currently if there is a failure in cleanupStagingData, it is not caught, so the rest of launchJob (which includes sending GobblinTrackingEvents) will not be run ### Tests - [x] My PR adds the following unit tests __OR__ does not need testing for this extremely good reason: trivial ### Commits - [x] My commits all reference JIRA issues in their subject lines, and I have squashed multiple commits if they address the same issue. In addition, my commits follow the guidelines from "[How to write a good git commit message](http://chris.beams.io/posts/git-commit/)": 1. Subject is separated from body by a blank line 2. Subject is limited to 50 characters 3. Subject does not end with a period 4. Subject uses the imperative mood ("add", not "adding") 5. Body wraps at 72 characters 6. Body explains "what" and "why", not "how"

github_pulls_comments:

1. +1 LGTM
2. @sv2000 please merge

github_pulls_reviews:

1. But this change will swallow the exception and any errors in `cleanupStagingData` will not fail the job and hence will go undetected (even if there is log message). Should we rather move the code we want to

run (sending GTEs) to the finally block at line 600?
2. Sure that's a good point, will change that
3. **body:** Actually not sure if that's the correct place for it either, because if there is an error when notifying listeners, then the cleanup code to close the metrics reporting wouldn't run. Maybe it should be in the finally block at line 542
   label: code-design
4. Actually isn't this kind of the normal behaviour for a failed job? Any exception in work unit creation is caught at line 530 and not rethrown, does the job not failed for those cases either?
5. Got it. Yes 542 is the right place.
6. Okay, so the current behavior is do not fail the job (just mark the gobblin job as FAILED) when workunit creation fails, or anything in the main try block fails. However, according to the current behavior any exception in `cleanupStagingData` and `notifyListeners` are thrown (i.e. making job fail). So, if we want to preserve the same behavior, we should move the code form L565-L599 to finally block at 542, right? Why the current behavior treats `cleanupStagingData` failures differently, and whether we should change this behavior is a different question that I am not fully sure about.

jira_issues:

1. **summary:** Catch error when failing to clean staging data
   description:
2. **summary:** Catch error when failing to clean staging data
   description:

jira_issues_comments: