

git_comments:

1. * Licensed to the Apache Software Foundation (ASF) under one or more * contributor license agreements. See the NOTICE file distributed with * this work for additional information regarding copyright ownership. * The ASF licenses this file to You under the Apache License, Version 2.0 * (the "License"); you may not use this file except in compliance with * the License. You may obtain a copy of the License at * * <http://www.apache.org/licenses/LICENSE-2.0> * * Unless required by applicable law or agreed to in writing, software * distributed under the License is distributed on an "AS IS" BASIS, * WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. * See the License for the specific language governing permissions and * limitations under the License.
2. * * hadoop custom param for sqoop job
3. * * sqoop job name - map-reduce job name
4. * * sqoop advanced param
5. * * sqoop job type: * CUSTOM - custom sqoop job * TEMPLATE - sqoop template job
6. * * customJob eq 1, use customShell
7. sqoop job type is template set task relation
8. set sqoop advanced custom param
9. set sqoop job name
10. set hadoop custom param
11. export hive to mysql
12. export hdfs to mysql using update mode
13. sqoop TEMPLATE job import mysql to HDFS with hadoo
14. * * test SqoopJobGenerator
15. import mysql to hive
16. * * get taskExecutionContext include mysql * @return TaskExecutionContext
17. test sqoop tash init method
18. sqoop CUSTOM job
19. 192.168.0.111:3306\","database":\',"test\',"jdbcUrl":\',"jdbc:mysql://192.168.0.111:3306/test\',"user":\',"kylo\',"password":\',"123456\',"";
192.168.0.111:3306\',"database":\',"test\',"jdbcUrl":\',"jdbc:mysql://192.168.0.111:3306/test\',"user":\',"kylo\',"password":\',"123456\',"";

git_commits:

1. **summary:** Merge pull request #2943 from Eights-Li/dev-sqoop-optimization
message: Merge pull request #2943 from Eights-Li/dev-sqoop-optimization Sqoop task optimization

github_issues:

1. **title:** [Feature] Sqoop component optimization
body: **Is your feature request related to a problem? Please describe.** dev branch sqoop task need to enhancement. optimization points: Sqoop's data access and data export do not support Hadoop-level custom parameters, that is, -D level parameters – MR task name – MR map and reduce memory and quantity, etc. • Split-by field is not supported. If -m is greater than 1, if the primary key of the relational database table is not self-increasing, Sqoop It may cause duplicate data imported into Hadoop. The general solution is to specify a split-by field. therefore, split-by needs support • Cannot customize parameters, such as import mysql, some tables can add –direct to speed up the import speed **Describe the solution you'd like** ideas: • The task name of Sqoop is universal, and it must be changed to the required parameter on the Sqoop page • Add Hadoop custom parameter input box for setting MR parameter memory, etc. • Add Sqoop task-level custom parameters, like –driect, –fetch-size and other parameters used in specific situations • Add option button to choose, custom script or use template script, refer to the design of DataX node
2. **title:** [Feature] Sqoop component optimization
body: **Is your feature request related to a problem? Please describe.** dev branch sqoop task need to enhancement. optimization points: Sqoop's data access and data export do not support Hadoop-level custom parameters, that is, -D level parameters – MR task name – MR map and reduce memory and quantity, etc. • Split-by field is not supported. If -m is greater than 1, if the primary key of the relational database table is not self-increasing, Sqoop It may cause duplicate data imported into Hadoop. The general solution is to specify a split-by field. therefore, split-by needs support • Cannot customize parameters, such as import mysql, some tables can add –direct to speed up the import speed **Describe the solution you'd like** ideas: • The task name of Sqoop is universal, and it must be changed to the required parameter on the Sqoop page • Add Hadoop custom parameter input box for setting MR parameter memory, etc. • Add Sqoop task-level custom parameters, like –driect, –fetch-size and other parameters used in specific situations • Add option button to choose, custom script or use template script, refer to the design of DataX node
label: code-design
3. **title:** [Feature] Sqoop component optimization
body: **Is your feature request related to a problem? Please describe.** dev branch sqoop task need to enhancement. optimization points: Sqoop's data access and data export do not support Hadoop-level custom parameters, that is, -D level parameters – MR task name – MR map and reduce memory and quantity, etc. • Split-by field is not supported. If -m is greater than 1, if the primary key of the relational database table is not self-increasing, Sqoop It may cause duplicate data imported into Hadoop. The general solution is to specify a split-by field. therefore, split-by needs support • Cannot customize parameters, such as import mysql, some tables can add –direct to speed up the import speed **Describe the solution you'd like** ideas: • The task name of Sqoop is universal, and it must be changed to the required parameter on the Sqoop page • Add Hadoop custom parameter input box for setting MR parameter memory, etc. • Add Sqoop task-level custom parameters, like –driect, –fetch-size and other parameters used in specific situations • Add option button to choose, custom script or use template script, refer to the design of DataX node
4. **title:** [Feature] Sqoop component optimization
body: **Is your feature request related to a problem? Please describe.** dev branch sqoop task need to enhancement. optimization points: Sqoop's data access and data export do not support Hadoop-level custom parameters, that is, -D level parameters – MR task name – MR map and reduce memory and quantity, etc. • Split-by field is not supported. If -m is greater than 1, if the primary key of the relational database table is not self-increasing, Sqoop It may cause duplicate data imported into Hadoop. The general solution is to specify a split-by field. therefore, split-by needs support • Cannot customize parameters, such as import mysql, some tables can add –direct to speed up the import speed **Describe the solution you'd like** ideas: • The task name of Sqoop is universal, and it must be changed to the required parameter on the Sqoop page • Add Hadoop custom parameter input box for setting MR parameter memory, etc. • Add Sqoop task-level custom parameters, like –driect, –fetch-size and other parameters used in specific situations • Add option button to choose, custom script or use template script, refer to the design of DataX node
5. **title:** [Feature] Sqoop component optimization
body: **Is your feature request related to a problem? Please describe.** dev branch sqoop task need to enhancement. optimization points: Sqoop's data access and data export do not support Hadoop-level custom parameters, that is, -D level parameters – MR task name – MR map and reduce memory and quantity, etc. • Split-by field is not supported. If -m is greater than 1, if the primary key of the relational database table is not self-increasing, Sqoop It may cause duplicate data imported into Hadoop. The general solution is to specify a split-by field. therefore, split-by needs support • Cannot customize parameters, such as import mysql, some tables can add –direct to speed up the import speed **Describe the solution you'd like** ideas: • The task name of Sqoop is universal, and it must be changed to the required parameter on the Sqoop page • Add Hadoop custom parameter input box for setting MR parameter memory, etc. • Add Sqoop task-level custom parameters, like –driect, –fetch-size and other parameters used in specific situations • Add option button to choose, custom script or use template script, refer to the design of DataX node

github_issues_comments:

1. The suggestion is very good. At present, the Sqoop node type does only support data import and export, and other Sqoop commands do not support it. But according to my understanding, the person in charge of the Sqoop class node may not want to open the function of custom parameters, because this is more difficult to verify. The above is just my understanding. I hope @zixi0825 can give us your opinions. Thank you.
2. **body:** The solution is good. When I designed, I thought that the graphical interface can meet the needs of simple and fast use. It is better to use custom scripts to support custom parameters. Custom scripts I think can be implemented with shell scripts, so they are not added to the sqoop task type.

- label:** code-design
3. merge into dev, close

github_pulls:

- 1. title:** Sqoop task optimization
body: ## What is the purpose of the pull request #2917 #2763 *sqoop task optimization* ## Brief change log - *Add CUSTOM and TAMELATE job for sqoop task* - *support hadoop params like mr job name, etc* - *suoopr sqoop advanced params like --direct & --fetch-size, etc* - *Add sqoop task front-end selection box cascade display* - *fix bug : java map fields and hive map fields generate error* ## Verify this pull request - */server/worker/task/sqoop/SqoopTaskTest.java*
- 2. title:** Sqoop task optimization
body: ## What is the purpose of the pull request #2917 #2763 *sqoop task optimization* ## Brief change log - *Add CUSTOM and TAMELATE job for sqoop task* - *support hadoop params like mr job name, etc* - *suoopr sqoop advanced params like --direct & --fetch-size, etc* - *Add sqoop task front-end selection box cascade display* - *fix bug : java map fields and hive map fields generate error* ## Verify this pull request - */server/worker/task/sqoop/SqoopTaskTest.java*
- 3. title:** Sqoop task optimization
body: ## What is the purpose of the pull request #2917 #2763 *sqoop task optimization* ## Brief change log - *Add CUSTOM and TAMELATE job for sqoop task* - *support hadoop params like mr job name, etc* - *suoopr sqoop advanced params like --direct & --fetch-size, etc* - *Add sqoop task front-end selection box cascade display* - *fix bug : java map fields and hive map fields generate error* ## Verify this pull request - */server/worker/task/sqoop/SqoopTaskTest.java*
- 4. title:** Sqoop task optimization
body: ## What is the purpose of the pull request #2917 #2763 *sqoop task optimization* ## Brief change log - *Add CUSTOM and TAMELATE job for sqoop task* - *support hadoop params like mr job name, etc* - *suoopr sqoop advanced params like --direct & --fetch-size, etc* - *Add sqoop task front-end selection box cascade display* - *fix bug : java map fields and hive map fields generate error* ## Verify this pull request - */server/worker/task/sqoop/SqoopTaskTest.java*
- 5. title:** Sqoop task optimization
body: ## What is the purpose of the pull request #2917 #2763 *sqoop task optimization* ## Brief change log - *Add CUSTOM and TAMELATE job for sqoop task* - *support hadoop params like mr job name, etc* - *suoopr sqoop advanced params like --direct & --fetch-size, etc* - *Add sqoop task front-end selection box cascade display* - *fix bug : java map fields and hive map fields generate error* ## Verify this pull request - */server/worker/task/sqoop/SqoopTaskTest.java*
- 6. title:** Sqoop task optimization
body: ## What is the purpose of the pull request #2917 #2763 *sqoop task optimization* ## Brief change log - *Add CUSTOM and TAMELATE job for sqoop task* - *support hadoop params like mr job name, etc* - *suoopr sqoop advanced params like --direct & --fetch-size, etc* - *Add sqoop task front-end selection box cascade display* - *fix bug : java map fields and hive map fields generate error* ## Verify this pull request - */server/worker/task/sqoop/SqoopTaskTest.java*
- 7. title:** Sqoop task optimization
body: ## What is the purpose of the pull request #2917 #2763 *sqoop task optimization* ## Brief change log - *Add CUSTOM and TAMELATE job for sqoop task* - *support hadoop params like mr job name, etc* - *suoopr sqoop advanced params like --direct & --fetch-size, etc* - *Add sqoop task front-end selection box cascade display* - *fix bug : java map fields and hive map fields generate error* ## Verify this pull request - */server/worker/task/sqoop/SqoopTaskTest.java*

github_pulls_comments:

1. # [Codecov](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943?src=pr&el=h1) Report > Merging [#2943](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943?src=pr&el=desc) into [dev](https://codecov.io/gh/apache/incubator-dolphinscheduler/commit/2749c7e5369459c1a410508cc4769db37f4a0043&el=desc) will **increase** coverage by `0.94%`. > The diff coverage is `60.86%`. [! [Impacted file tree graph](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/graphs/tree.svg? width=650&height=150&src=pr&token=by9iXRLi9)](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943?src=pr&el=tree) `` diff @@ Coverage Diff @@ ## dev #2943 +/- ## ===== + Coverage 37.20% 38.14% +0.94% - Complexity 2559 2621 +62 ===== Files 435 436 +1 Lines 20020 20079 +59 Branches 2426 2443 +17 ===== + Hits 7449 7660 +211 + Misses 11904 11711 -193 - Partials 667 708 +41 `` | [Impacted Files] (https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943?src=pr&el=tree) | Coverage Δ | Complexity Δ | |---|---|---|---| | [...ommon/task/sqoop/targets/TargetMysqlParameter.java](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1jb21tb24vc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvY29tbW9uL3Rhc2sv3Fvb3AvdGFyZ2V0cy9 | `100.00% <0> (0)` | `19.00 <0.00> (0)` | | [...ver/master/consumer/TaskPriorityQueueConsumer.java](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1zXJ2ZXIvc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvc2VydmdVYlL21hc3Rlc1jb25zdW1lciUUYXN | `51.82% <0.00%> (-7.00%)` | `16.00 <0.00> (0)` | | [.../sqoop/generator/sources/MysqlSourceGenerator.java](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1zXJ2ZXIvc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvc2VydmdVYlL3dvcmtlci90YXNlL3Nxb29wLz | `58.00% <42.85%> (+58.00%)` | `5.00 <0.00> (+5.00)` | | [...inscheduler/common/task/sqoop/SqoopParameters.java](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1jb21tb24vc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvY29tbW9uL3Rhc2sv3Fvb3AvU3Fvb3BQYX | `74.00% <58.06%> (-6.00%)` | `25.00 <12.00> (+10.00)` | :arrow_down: | | [...worker/task/sqoop/generator/SqoopJobGenerator.java] (https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1zXJ2ZXIvc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvc2VydmdVYlL3dvcmtlci90YXNlL3Nxb29wLz | `74.07% <72.72%> (+74.07%)` | `10.00 <2.00> (+10.00)` | | [...he/dolphinscheduler/common/enums/SqoopJobType.java] (https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1jb21tb24vc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvY29tbW9uL2VudW1zL1Nxb29wSm9iVHlwZ | `88.88% <88.88%> (0)` | `3.00 <3.00> (?)` | | [...r/worker/task/sqoop/generator/CommonGenerator.java](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1zXJ2ZXIvc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvc2VydmdVYlL3dvcmtlci90YXNlL3Nxb29wLz | `83.87% <89.47%> (+83.87%)` | `6.00 <0.00> (+6.00)` | | | [.../org/apache/dolphinscheduler/api/enums/Status.java](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1hcGkvc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvYXBpL2VudW1zL1N0YXR1cy5qYXZh) | `100.00% <100.00%> (0)` | `5.00 <0.00> (0)` | | | [.../sqoop/generator/targets/MysqlTargetGenerator.java](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff?src=pr&el=tree#diff-ZG9scGhpbNjJaGVkdWxlc1zXJ2ZXIvc3JlL21haW4vamF2YS9vcmcvYXBhY2hlL2RvbHB0aW5zY2hlZHVzZXIvc2VydmdVYlL3dvcmtlci90YXNlL3Nxb29wLz | `75.75% <100.00%> (+75.75%)` | `4.00 <0.00> (+4.00)` | | ... and [17 more](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943/diff? src=pr&el=tree-more) | | ----- [Continue to review full report at Codecov](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943? src=pr&el=continue). > **Legend** - [Click here to learn more](https://docs.codecov.io/docs/codecov-delta) > `Δ` = absolute <relative> (impact), `0` = not affected, `?` = missing data > Powered by [Codecov](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943?src=pr&el=footer). Last update [2749c7e...a420644](https://codecov.io/gh/apache/incubator-dolphinscheduler/pull/2943?src=pr&el=lastupdated). Read the [comment docs] (https://docs.codecov.io/docs/pull-request-comments).
2. @break60 please front-end. thx @zixi0825 hi, this pr for sqoop task @743294668 bro, Can this pr meet your requirements?
3. front-end +1
4. Kudos, SonarCloud Quality Gate passed! [] (https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=BUG) [] (https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=BUG) [0 Bugs] (https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=BUG) [] (https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=VULNERABILITY) [] (https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=VULNERABILITY) [0 Vulnerabilities] (https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=VULNERABILITY) (and [])

resources/v2/common/security_hotspot.png' alt='Security Hotspot' width='16' height='16' />](https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=SECURITY_HOTSPOT) [0 Security Hotspots](https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=SECURITY_HOTSPOT) to review) [](https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=CODE_SMELL) [](https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=CODE_SMELL) [1 Code Smell](https://sonarcloud.io/project/issues?id=apache-dolphinscheduler&pullRequest=2943&resolved=false&types=CODE_SMELL) [](https://sonarcloud.io/component_measures?id=apache-dolphinscheduler&pullRequest=2943&metric=new_coverage&view=list) [68.4% Coverage](https://sonarcloud.io/component_measures?id=apache-dolphinscheduler&pullRequest=2943&metric=new_coverage&view=list) [](https://sonarcloud.io/component_measures?id=apache-dolphinscheduler&pullRequest=2943&metric=new_duplicated_lines_density&view=list) [0.0% Duplication](https://sonarcloud.io/component_measures?id=apache-dolphinscheduler&pullRequest=2943&metric=new_duplicated_lines_density&view=list) [The version of Java (1.8.0_252) you have used to run this analysis is deprecated and we will stop accepting it from October 2020. Please update to at least Java 11. Read more [here](https://sonarcloud.io/documentation/upcoming/)

github_pulls_reviews:

- 1. If return null, I see the follow-up code, the task will eventually be judged as successful.
- 2. yes, if null should throw a RuntimeException

jira_issues:

jira_issues_comments: