git_comments:

1. return String.format("%,24d", numEntries);
2. * * Utility that generates single line tablet info. The output of this could be fed to sort, awk, * grep, etc inorder to answer questions like which tablets have the most files.
3. * * Wrapper for tablename and id. Comparisons, equals and hash code use tablename (id is ignored)
4. return String.format("%,24d", size);
5. * * Process the command line for table names using table option, table name pattern, or default to * current table. * * @param cl * command line * @param shellState * shell state * @return set of table names. * @throws NamespaceNotFoundException * if the namespace option is specified and namespace does not exist
6. * 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.
7. If we didn't get any tables, and we have a table selected, add the current table
8. second tablet info
9. * 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.
10. first table info
11. third tablet info
12. kill a tserver and see if its gone from the list
13. * 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.
14. * * Tests features of the Ample TabletMetadata class that can't be tested in TabletMetadataTest
15. * * Check for tserver ZooLock at the ZK location. Return Optional containing TServerInstance if a * valid Zoolock exists.
16. * * Get the tservers that are live from ZK. Live servers will have a valid ZooLock. This method was * pulled from org.apache.accumulo.server.master.LiveTServerSet

git_commits:

1. **summary:** Create listtablets shell command. Closes #1317 (#1821)
   message: Create listtablets shell command. Closes #1317 (#1821) * New command for debugging tablets called listtablets * Added getLiveTServers() to TabletMetadata for generating a list of tservers that currently have a lock in ZK, similar to master. * The list of live tservers is passed to TabletMetadata in order to get the current state of a tablet * Command will print one line for every tablet in a table * Created TabletMetadataIT for testing getLiveTServers() Co-authored-by: EdColeman <dev1@etcoleman.com>

github_issues:

1. **title:** Provide single-line output of tablet info.
   body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.
2. **title:** Provide single-line output of tablet info.
   body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.
   label: code-design
3. **title:** Provide single-line output of tablet info.
   body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.
4. **title:** Provide single-line output of tablet info.
   body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.
5. **title:** Provide single-line output of tablet info.
   body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.
6. **title:** Provide single-line output of tablet info.
   body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.
   label: code-design
7. **title:** Provide single-line output of tablet info.

body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.

8. **title:** Provide single-line output of tablet info.
   body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.

9. **title:** Provide single-line output of tablet info.
   body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.

10. **title:** Provide single-line output of tablet info.
    body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.

11. **title:** Provide single-line output of tablet info.
    body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.

12. **title:** Provide single-line output of tablet info.
    body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.

13. **title:** Provide single-line output of tablet info.
    body: Copied from ACCUMULO-2873 - It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc in order to answer questions like which tablets have the most files.The output could look something like the following ``` $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful ``` All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory.

github_issues_comments:

1. I've taken a first run at this - current available at: https://github.com/EdColeman/accumulo/tree/listTabletsCmd exploring what may be possible at to start a discussion.

2. **body:** This still needs addition work. - What is the proper way to determine status? One way would be to check if the tablet has a loc: assigned. Is there something else that is more appropriate? - The original ticket mentioned that some information could be pulled from zookeeper - with the current implementation, all information is coming from the metadata table. What information is missing / would be helpful? - Formatting of size and number of entries to allow human readable option.
   label: code-design

3. > What is the proper way to determine status? Could use [TabletLocationState.getState()][2]. One way to get TabletLocationState objects is via [TabletStateStore] [1]. The set of live tservers can be obtained using [LiveTServerSet][3]. > The original ticket mentioned that some information could be pulled from zookeeper - with the current implementation, all information is coming from the metadata table. What information is missing / would be helpful? [LiveTServerSet][3] pulls the set of active tablet servers from zookeeper. [1]:
   https://github.com/apache/accumulo/blob/f39c2fef1c1c911e86985ce6352c5d1bc9967c5b/server/base/src/main/java/org/apache/accumulo/server/master/state/TabletS [2]:
   https://github.com/apache/accumulo/blob/f39c2fef1c1c911e86985ce6352c5d1bc9967c5b/server/base/src/main/java/org/apache/accumulo/server/master/state/TabletL [3]:
   https://github.com/apache/accumulo/blob/f39c2fef1c1c911e86985ce6352c5d1bc9967c5b/server/base/src/main/java/org/apache/accumulo/server/master/LiveTServer

4. @keith-turner @ctubbsii I have been working on this as it seems like a very useful tool. One of the issues I am running into is that the only way to get the current state of tablets is to use objects in the master and this introduces a dependency in the shell on accumulo-server-base. There are other parts of the code that depend on accumulo-server-base, like the GC [here]
   (https://github.com/apache/accumulo/blob/72e8bc41335ae542087cc70d80d2b7b7e8652575/server/gc/src/main/java/org/apache/accumulo/gc/GarbageCollectWriteA The information is coming from ZK and metadata so I was wondering if having something in Ample would be appropriate. I am not sure what is the best solution.

5. Ampl might be okay for this. It might also be okay to have a public API for the shell to use.

6. **body:** @milleruntime I suspect the code could be moved to accumulo core OR reimplemented in core (and existing code migrated to Ample). One possible way to add it to ample would be to modify TabletMetadata class to add a method like the following. ```java TabletState getTabletState(Set<Location> liveTservers) { //use tablet metadata + liveTservers to determine the state of this tablet. // Location is an inner class TabletMetadata } ``` Then maybe Ample could have a getLiveTservers() method.
   label: code-design

7. For my previous comment, the Location type will not really work for this case. Need a type that just encapsulates `(server,session)` for the set of live tservers. TabletMetadata.Location has those plus LocationType, which makes no sense for a live tserver.

8. > For my previous comment, the Location type will not really work for this case. Need a type that just encapsulates `(server,session)` for the set of live tservers. TabletMetadata.Location has those plus LocationType, which makes no sense for a live tserver. What do you think about a new type in core.metadata (similar to what I did for TabletFile) that could be used everywhere? And would we want to keep the ZK session Id? @EdColeman and I were talking about the usefulness of storing the internal session ID. It seems like its used as a unique identifier of the ephemeral ZooLock for that tserver. This is a separate issue but I don't think we would want to have the session in this new type.

9. And I suppose it would need to replace this type here:
   https://github.com/apache/accumulo/blob/d1b06d317983d460039861d4ab462d5a65ca3c11/core/src/main/java/org/apache/accumulo/core/metadata/schema/Ample.ja

10. >. What do you think about a new type in core.metadata (similar to what I did for TabletFile) that could be used everywhere? And would we want to keep the ZK session Id? I like that. That Location type in TabletMetadata could possibly extend it. Not sure about ZK session id would have to look at it. Using the same type to represent the same info throughout the code seems like a big win. Seems like it can follow the pattern of add it and then later adapt code to use it.

11. > > . What do you think about a new type in core.metadata (similar to what I did for TabletFile) that could be used everywhere? And would we want to keep the ZK session Id? > > I like that. That Location type in TabletMetadata could possibly extend it. Not sure about ZK session id would have to look at it. Using the same type to represent the same info throughout the code seems like a big win. Seems like it can follow the pattern of add it and then later adapt code to use it. Instead of a new type, since the class is much simpler than TabletFile, I am playing around with just moving TServerInstance from master to core and using it. https://github.com/apache/accumulo/blob/main/server/base/src/main/java/org/apache/accumulo/server/master/state/TServerInstance.java

github_pulls:

github_pulls_comments:

github_pulls_reviews:

jira_issues:

1. **summary:** Create utility that generates single line tablet information
   description: It would be very useful to have a utility that generates single line tablet info. The output of this could be fed to sort, awk, grep, etc inorder to answer questions like which tablets have the most files. The output could look something like the following {noformat} $accumulo admin listTablets --table bigTable3 #files #walogs #entries #size #status #location #tableid #endrow 6 2 40,001 50M ASSIGNED 10.1.9.9 4:9997[abc] 3 admin 3 1 50,002 40M ASSIGNED 10.1.9.9 5:9997[abc] 3 helpful {noformat} All of the information can be obtained by scanning the metadata table and looking into zookeeper. Could possibly contact tablet servers to get info about entires in memory. The order of the columns in the example above is arbitrary, except for end row. Maybe end row column should come last because it can be of arbitrary length. Also the end row could contain any character, could look into using a CSV library. It would be nice to design the utility so that columns can be added in future versions w/o impacting current scripts that use the utility.

jira_issues_comments:

1. if we used Avro for this output, it would easily handle the binary/text issue. We could also use the existing avro-tools utilities to get a textual representation or to do projections of some subset of data. Also it'd be super easy then to work with it programmatically. It'd also leverage Avro's extensive schema evolution support, so changes would be easy to do.
2. Resolving as "Won't Fix", because it there does not appear to be any interest in working on this. If this is still an issue, please open a new issue or pull request at https://github.com/apache/accumulo/issues
3. Duplicated by the issue on GitHub: https://github.com/apache/accumulo/issues/1317