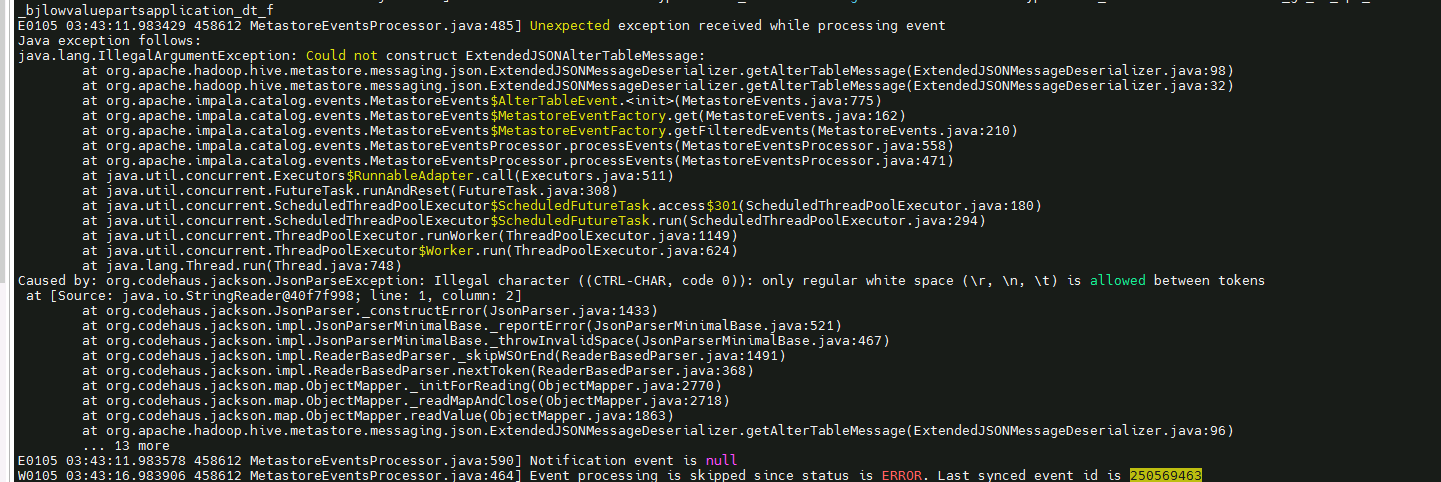
Impala 3.2.0

Hive 2.1.1



E0105 03:43:11.983429 458612 MetastoreEventsProcessor.java:485] Unexpected exception received while processing event

Java exception follows:

java.lang.IllegalArgumentException: Could not construct ExtendedJSONAlterTableMessage:

at org.apache.hadoop.hive.metastore.messaging.json.ExtendedJSONMessageDeserializer.getAlterTableMessage(ExtendedJSONMessageDeserializer.java:98)

at org.apache.hadoop.hive.metastore.messaging.json.ExtendedJSONMessageDeserializer.getAlterTableMessage(ExtendedJSONMessageDeserializer.java:32)

at org.apache.impala.catalog.events.MetastoreEvents$AlterTableEvent.<init>(MetastoreEvents.java:775)

at org.apache.impala.catalog.events.MetastoreEvents$MetastoreEventFactory.get(MetastoreEvents.java:162)

at org.apache.impala.catalog.events.MetastoreEvents$MetastoreEventFactory.getFilteredEvents(MetastoreEvents.java:210)

at org.apache.impala.catalog.events.MetastoreEventsProcessor.processEvents(MetastoreEventsProcessor.java:558)

at org.apache.impala.catalog.events.MetastoreEventsProcessor.processEvents(MetastoreEventsProcessor.java:471)

at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)

at java.util.concurrent.FutureTask.runAndReset(FutureTask.java:308)

at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.access$301(ScheduledThreadPoolExecutor.java:180)

at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.run(ScheduledThreadPoolExecutor.java:294)

at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)

at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)

at java.lang.Thread.run(Thread.java:748)

Caused by: org.codehaus.jackson.JsonParseException: Illegal character ((CTRL-CHAR, code 0)): only regular white space (\r, \n, \t) is allowed between tokens

at [Source: java.io.StringReader@40f7f998; line: 1, column: 2]

at org.codehaus.jackson.JsonParser.\_constructError(JsonParser.java:1433)

at org.codehaus.jackson.impl.JsonParserMinimalBase.\_reportError(JsonParserMinimalBase.java:521)

at org.codehaus.jackson.impl.JsonParserMinimalBase.\_throwInvalidSpace(JsonParserMinimalBase.java:467)

at org.codehaus.jackson.impl.ReaderBasedParser.\_skipWSOrEnd(ReaderBasedParser.java:1491)

at org.codehaus.jackson.impl.ReaderBasedParser.nextToken(ReaderBasedParser.java:368)

at org.codehaus.jackson.map.ObjectMapper.\_initForReading(ObjectMapper.java:2770)

at org.codehaus.jackson.map.ObjectMapper.\_readMapAndClose(ObjectMapper.java:2718)

at org.codehaus.jackson.map.ObjectMapper.readValue(ObjectMapper.java:1863)

at org.apache.hadoop.hive.metastore.messaging.json.ExtendedJSONMessageDeserializer.getAlterTableMessage(ExtendedJSONMessageDeserializer.java:96)

... 13 more

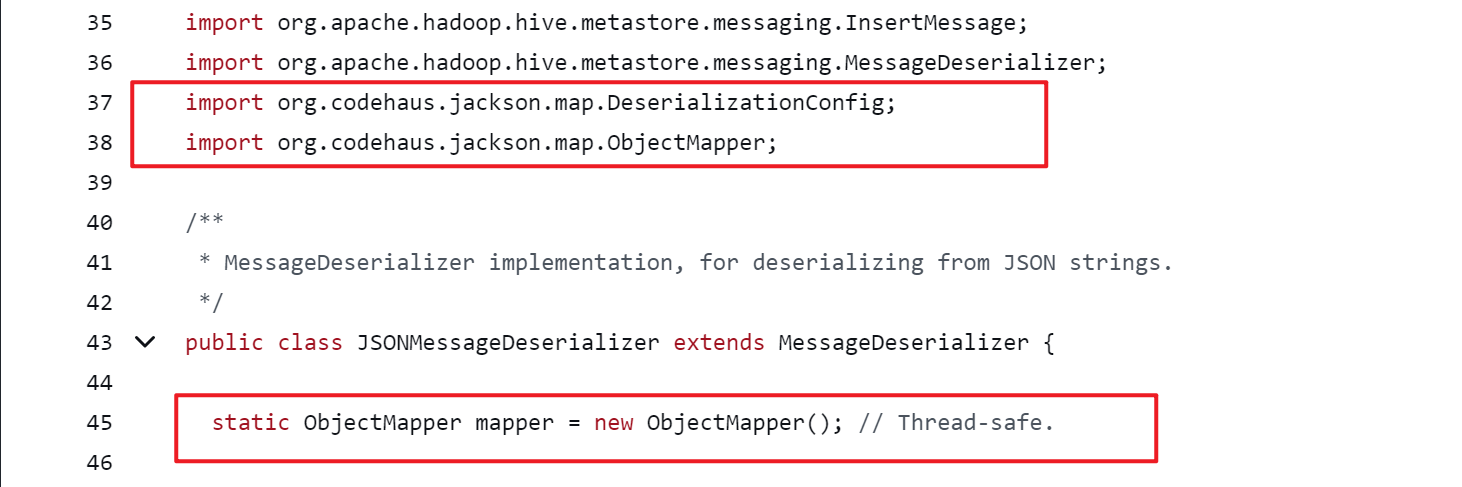
E0105 03:43:11.983578 458612 MetastoreEventsProcessor.java:590] Notification event is null

W0105 03:43:16.983906 458612 MetastoreEventsProcessor.java:464] Event processing is skipped since status is ERROR. Last synced event id is 250569463

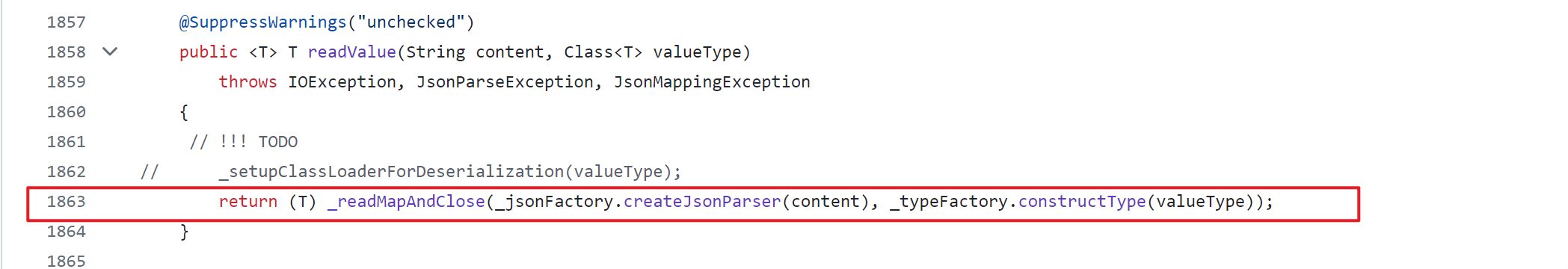
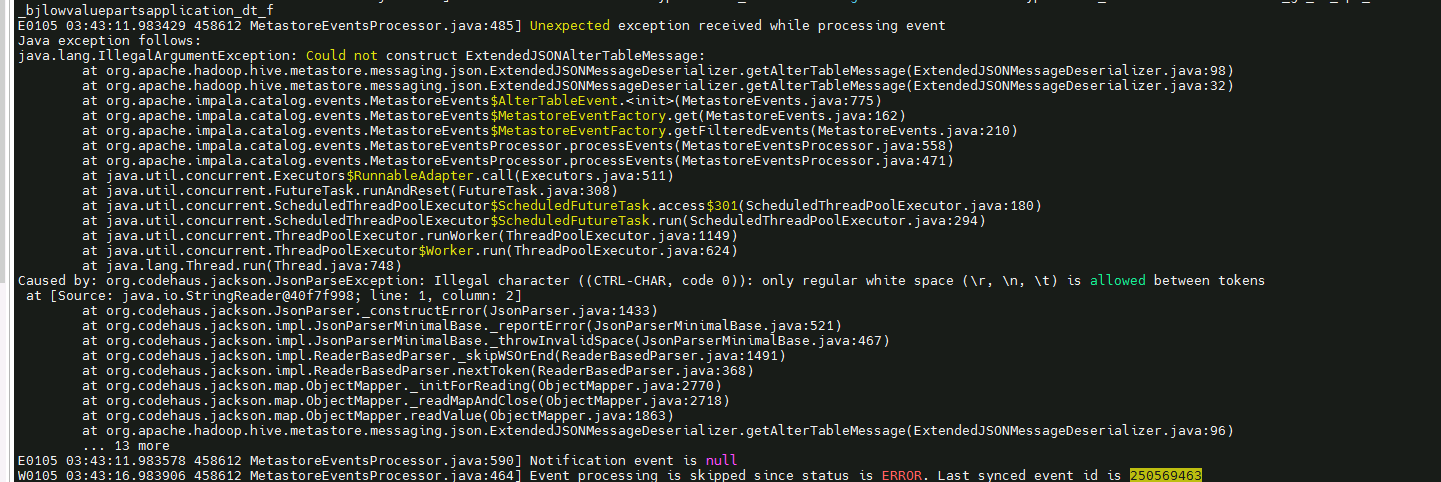
根据报错显示catalog同步的元数据中有非法字符，但元存储中没有找到非法字符的存在。

因此根据报错打印的类调用信息查询源码：

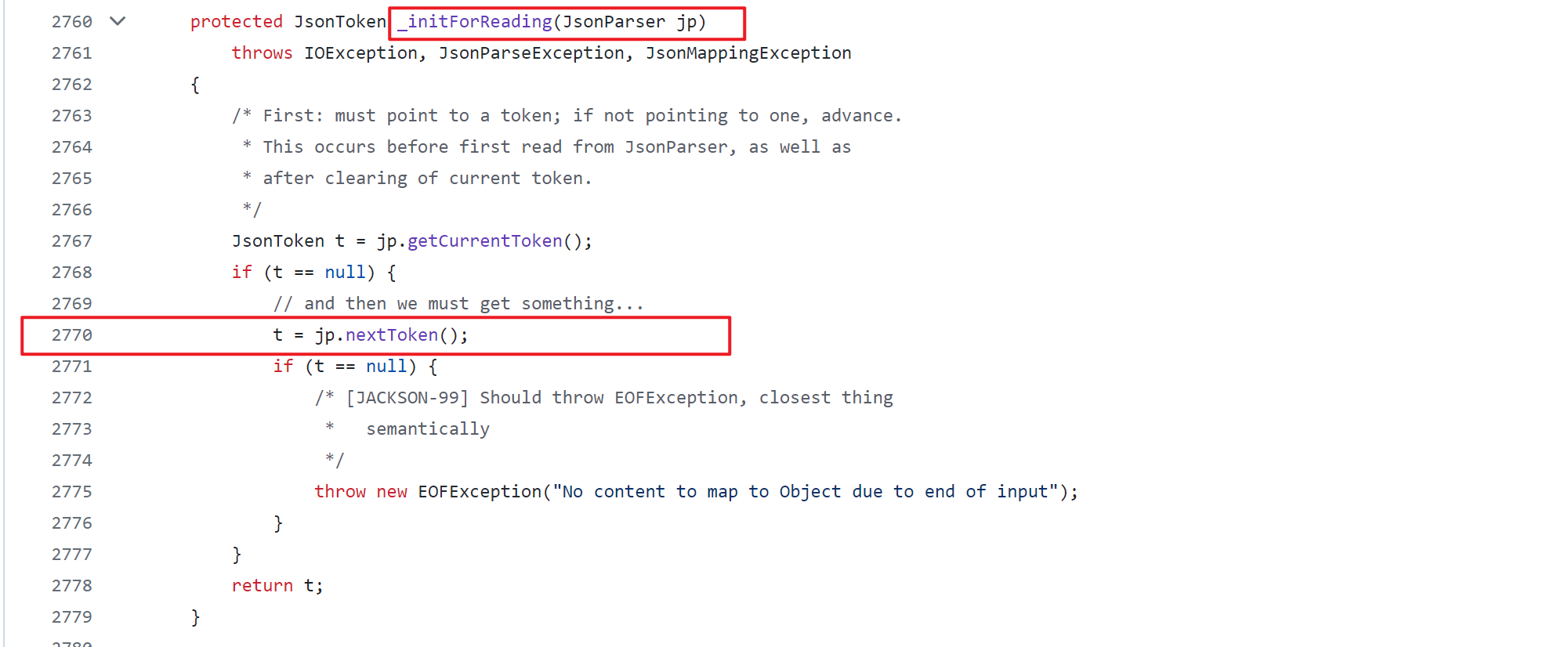
最开始的报错是hive 源码中ExtendedJSONMessageDeserializer.java类中的getAlterTableMessage方法打印的Could not construct ExtendedJSONAlterTableMessage报错， ExtendedJSONMessageDeserializer.java这个类中引入了codehaus jackson这个解析json的包，然后创建了一个 codehaus jackson的mapper对象，在getAlterTableMessage方法中使用mapper对象中的readValue方法读取元存储通知事件中的json数据。



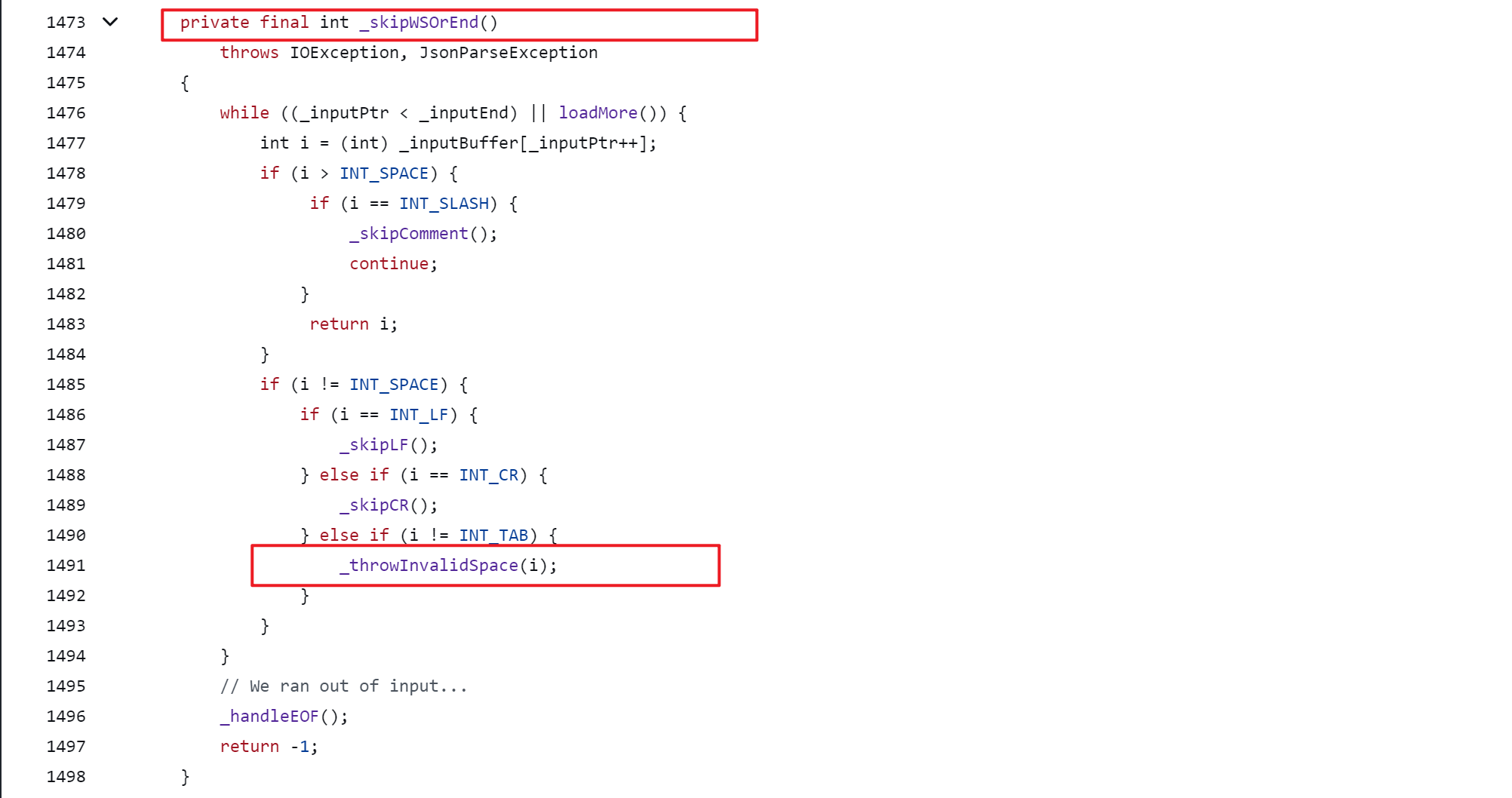
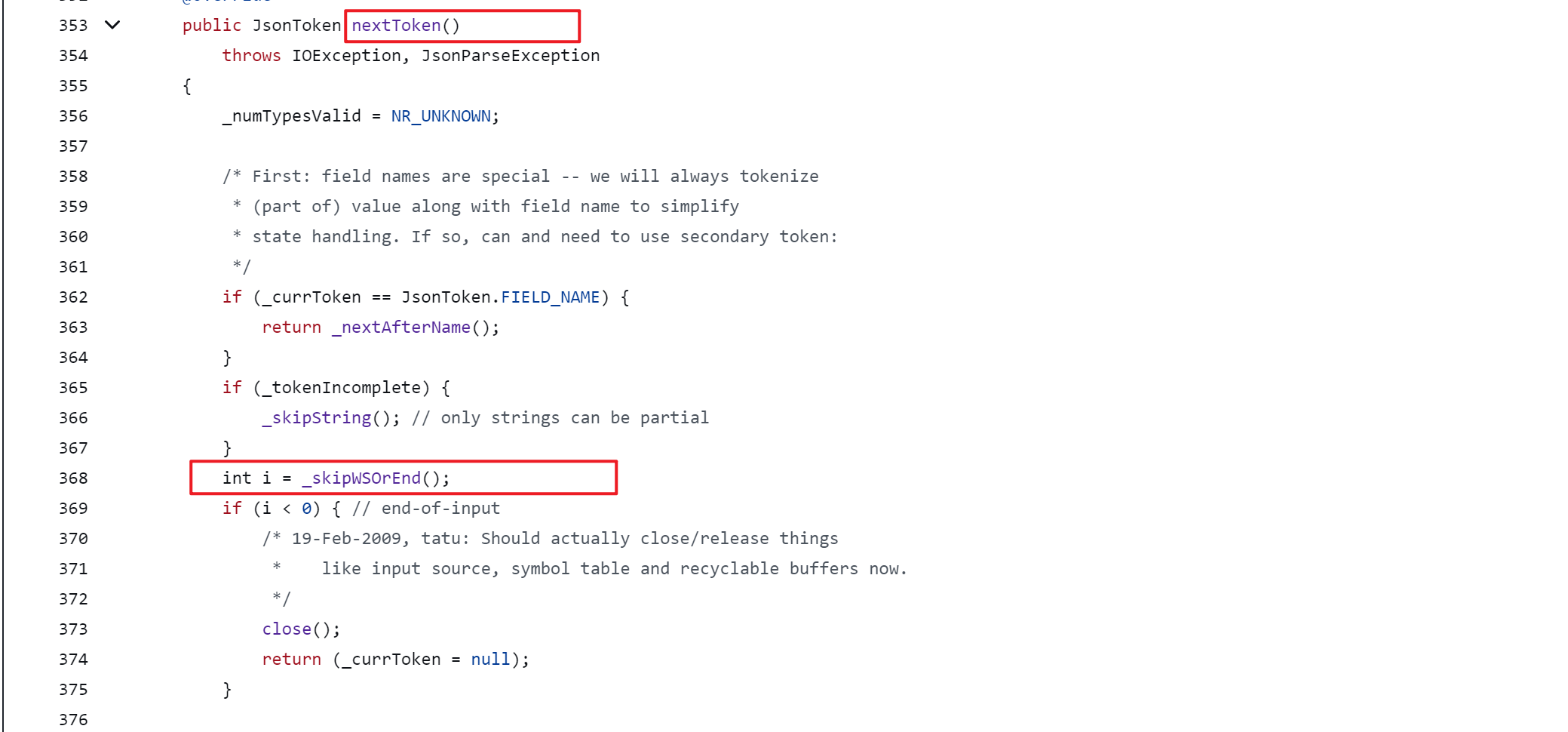
然后是codehaus jackson中的ObjectMapper.java类中的readValue方法抛出异常



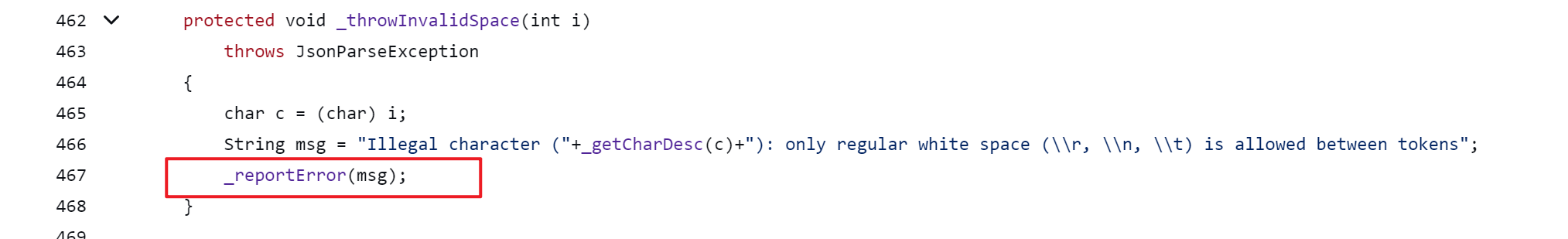
继续追溯调用路径，readValue方法中的异常来自于\_readMapAndClose方法，\_readMapAndClose方法又调用了\_initForReading方法，\_initForReading方法又调用了nextToken方法。

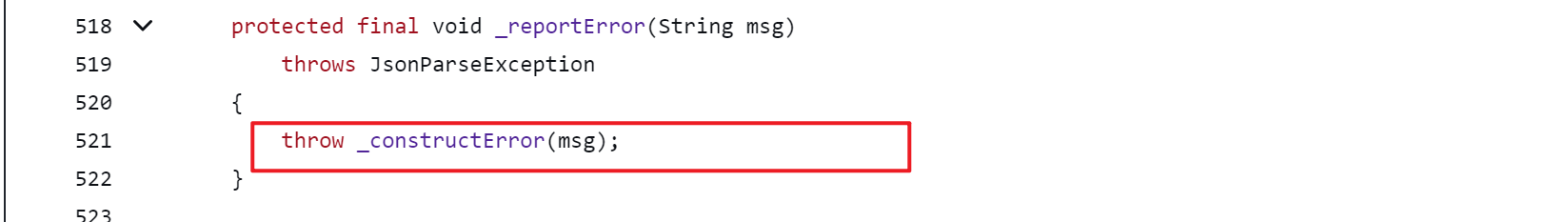


nextToken方法在impl包中的ReaderBasedParser.java类中，nextToken的异常来自于 \_skipWSOrEnd方法，\_skipWSOrEnd方法又调用了\_throwInvalidSpace方法。

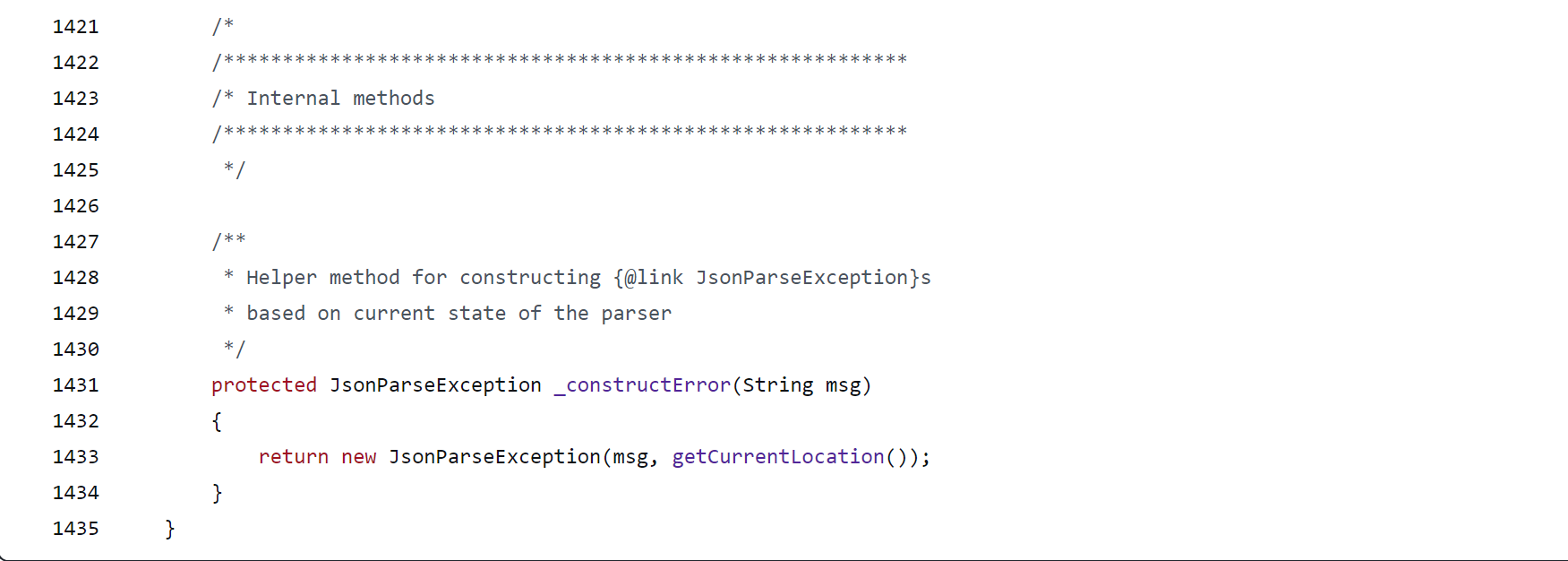


\_throwInvalidSpace方法在JsonParserMinimalBase.java类中，\_throwInvalidSpace方法的异常信息来自于\_reportError方法，\_reportError方法的异常信息是\_constructError方法抛出的。

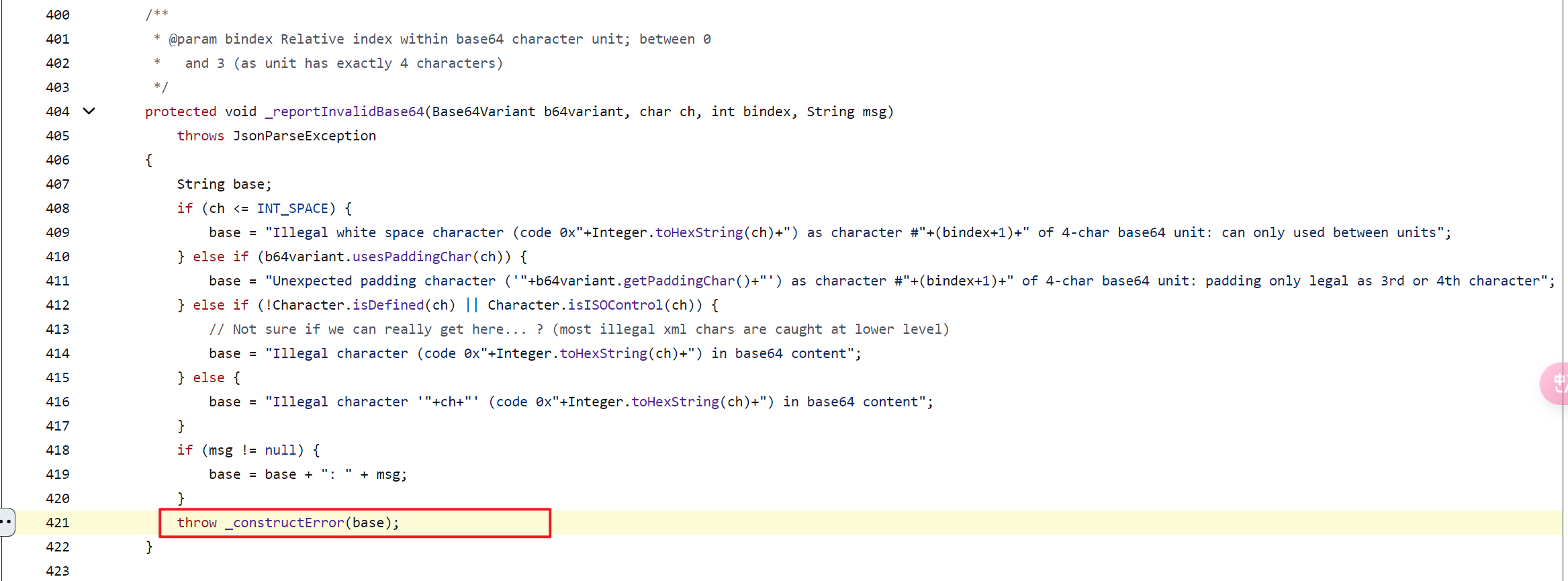




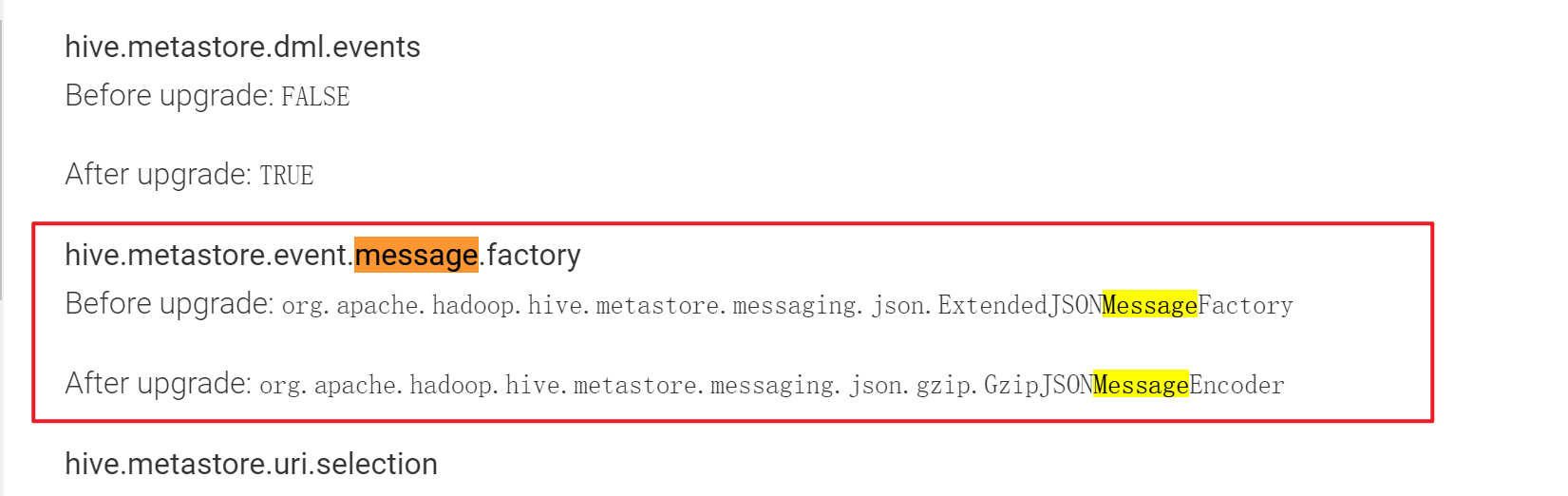
\_constructError方法在JsonParser.java类中。



\_constructError方法返回了一个异常消息，调用这个类并传入msg的是 JsonParserMinimalBase.java类中的\_reportInvalidBase64方法，这里也是最终的错误日志组装的地方。



查看cloudera文档，在CDP文档中发现从CDH升级到CDP后metastore事件消息工厂类有变动，从ExtendedJSONMessageFactory类改为使用GzipJSONMessageEncoder类，这可能说明ExtendedJSONMessageFactory类是有问题的才会修改，catalog 事件状态ERROR，报错非法字符正是ExtendedJSONMessageFactory类抛出的，这可能是一个bug。



参考文档：

https://support.oracle.com/knowledge/Oracle%20Cloud/2575085\_1.html

https://github.com/square/okhttp/issues/1026

https://docs.cloudera.com/cdp-private-cloud-upgrade/latest/upgrade-cdh/topics/ug\_hive\_configuration\_changes.html

https://github.com/square/okhttp/issues/635

https://github.com/codehaus/jackson/blob/master/src/mapper/java/org/codehaus/jackson/map/ObjectMapper.java

<https://github.com/apache/hive/blob/master/standalone-metastore/metastore-server/src/main/java/org/apache/hadoop/hive/metastore/messaging/json/JSONMessageDeserializer.java#L63>

<https://github.com/FasterXML/jackson-databind/blob/2.17/src/main/java/com/fasterxml/jackson/databind/ObjectMapper.java>

<https://issues.apache.org/jira/browse/HIVE-16219>

<https://stackoverflow.com/questions/42621547/jackson-error-illegal-character-only-regular-white-space-allowed-when-parsi>

<https://github.com/FasterXML/jackson-core/issues/633>

<https://copyprogramming.com/howto/jackson-error-illegal-character-only-regular-white-space-allowed-when-parsing-json>

<https://www.googlecloudcommunity.com/gc/Data-Analytics/Deserialize-BigQuery-records-with-Avro-with-JsonParseException/m-p/551517>

<https://blog.csdn.net/qq_44079295/article/details/127532101>

https://www.jianshu.com/p/df37eb5f2169