

#### git\_comments:

1. \* Splits given hadoop glob path by commas. \* e.g. splitGlob("/a,/b") -> ["/a", "/b"] \*  
splitGlob("/a/{c,d}") -> ["/a/c", "/a/d"]
2. \* Licensed to Metamarkets Group Inc. (Metamarkets) under one \* or more contributor license agreements. See the NOTICE file \* distributed with this work for additional information \* regarding copyright ownership. Metamarkets 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.
3. Note: This class has been created to workaround <https://issues.apache.org/jira/browse/MAPREDUCE-5061>
4. \* Licensed to Metamarkets Group Inc. (Metamarkets) under one \* or more contributor license agreements. See the NOTICE file \* distributed with this work for additional information \* regarding copyright ownership. Metamarkets 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.
5. \*

#### git\_commits:

1. **summary:** HadoopGlobPathSplitter implementation to split hadoop glob paths  
**message:** HadoopGlobPathSplitter implementation to split hadoop glob paths This can be safely reverted once <https://issues.apache.org/jira/browse/MAPREDUCE-5061> is fixed

#### github\_issues:

#### github\_issues\_comments:

#### github\_pulls:

#### github\_pulls\_comments:

#### github\_pulls\_reviews:

#### jira\_issues:

1. **summary:** MultipleInputs can not manage path with glob characters like {a,b}  
**description:** Hello, The org.apache.hadoop.mapreduce.lib.input.MultipleInputs expects a configuration with : \* the parameter mapreduce.input.multipleinputs.dir.mappers in the format <path 1>,<mapper class 1>;<path 2>,<mapper class 2> \* the parameter mapreduce.input.multipleinputs.dir.formats in the format <path 1>,<input format class 1>;<path 2>,<input format class 2> The usage of the comma as separator unable the use of the glob characters {a,b} in the path (like : [hdfs://localhost/user/foo/data201{2,3}](https://hdfs://localhost/user/foo/data201{2,3}) ). I suggest to change the separator to another non-problematic one (maybe =). I patched the MultipleInputs code locally and found an other related issue : when the path are created, the string representing the path is not unescape (like it's done in the org.apache.hadoop.mapreduce.lib.input.FileInputFormat.getInputPaths method) which makes an invalid path in case of glob characters {a,b} Thank you

#### jira\_issues\_comments:

1. Maybe = is not a good idea for separator as that would break the backward compatibility, so a smart regexp seems to be more appropriate.

2. FileInputFormat.java contains // This method escapes commas in the glob pattern of the given paths.  
private static String[] getPathStrings(String commaSeparatedPaths) {...} we can make that method public and possibly do following in the MultipleInputs.addInputPath(..) for (String p : FileInputFormat.getPathStrings(path)) { //add the path p , it will not have commas }
3. spoke too soon, FileInputFormat.getPathStrings(..) does not work if path contained braces.
4. What do you guys think about an escape approach (e.g. "{a\\,b}")?