Administration

Table of contents

1 Output location strict check.....2
2 Disabling Pig commands and operators.....2

Copyright © 2007 The Apache Software Foundation. All rights reserved.

1 Output location strict check

Pig scripts could contain multiple STORE statements. There are cases when one would like to avoid writing to the same output location. Pig provides admins/script writers with a property to check if multiple STORE statements make an attempt to write to the same output directory. And fail fast letting the user know of the same.

Specifically this makes sense for file-based output locations (HDFS, Local FS, S3..) to avoid Pig script from failing when multiple MR jobs write to the same location.

To enforce strict checking of output location, set **pig.location.check.strict=true**. See also Pig Properties on how to set this property.

2 Disabling Pig commands and operators

This is an admin feature providing ability to blacklist or/and whitelist certain commands and operations. Pig exposes a few of these that could be not very safe in a multitenant environment. For example, "sh" invokes shell commands, "set" allows users to change non-final configs. While these are tremendously useful in general, having an ability to disable would make Pig a safer platform. The goal is to allow administrators to be able to have more control over user scripts. Default behaviour would still be the same - no filters applied on commands and operators.

There are two properties you can use to control what users are able to do

- pig.blacklist
- pig.whitelist

2.1 Blacklisting

Set "pig.blacklist" to a comma-delimited set of operators and commands. For eg, **pig.blacklist=rm,kill,cross** would disable users from executing any of "rm", "kill" commands and "cross" operator.

2.2 Whitelisting

This is an even safer approach to disallowing functionality in Pig. Using this you will be able to disable all commands and operators that are not a part of the whitelist. For eg, **pig.whitelist=load,filter,store** will disallow every command and operator other than "load", "filter" and "store".

Copyright © 2007 The Apache Software Foundation. All rights reserved.

2.3 Note

There should not be any conflicts between blacklist and whitelist. Make sure to have them entirely distinct or Pig will complain.

Copyright © 2007 The Apache Software Foundation. All rights reserved.