File transport handler

The file transport handler allows the reading and writing to arbitrary files

Options

- **filename**: The path to the file
- **directory**: (since 2023/06/07): read recursive all files inside this directory.
- **path**: (since 2024/02/28): When using directory, a pattern can be used. Default is "*"
- **append**: append values to end of file. If set to false the values are overwritten. (default false)
- **createdir**: When writing to a file, creates the file directory if it doesn't exist. (default false)
- **writedelaysize**: Uses a buffer with specified size which has to be filled before data is written to the file (default 0)
- **preload**: Loads the whole file into memory when the query is started (default false, incompatible with **delayopenin**)
- **delayopenin**: When reading from a file, delays opening until reading the first tuple. Can be used to monitor log files which do not exist upon starting the query (default false)
- **delayopenout**: When writing to a file, delays opening (and, if specified, creating the directory) until writing the first tuple (default false, incompatible with **preload**)

Example

PQL

File Transport Handler

```
input = ACCESS({source='Source',
wrapper='GenericPull',
transport='File',
protocol='CSV',
dataHandler='Tuple',
options=[['filename', '/src/odysseus/input.csv']],
schema=[
['id', 'Double'],
['data', 'String']]
})

output = SENDER({sink='Sink',
wrapper='GenericPush',
transport='File',
protocol='CSV',
dataHandler='Tuple',
options=[['filename', '/src/odysseus/output.csv']]
}, input)
```

Example with directory:

```
/// Data from tankerkoenig.de, example to stream all historic price data
#PARSER PQL
#ADDQUERY
in = CSVFILESOURCE({
          schema = [
            ['date', 'STARTTIMESTAMPSTRING'],
            ['station_uuid','STRING'],
            ['diesel','FLOAT'],
            ['e5','FLOAT'],
            ['e10','FLOAT'],
            ['dieselchange','BOOLEAN'],
            ['e5change','BOOLEAN'],
            ['e10change','BOOLEAN']
          ],
          READFIRSTLINE = false,
          directory = 'D:/tankerkoenig-data/prices/',
          DATEFORMAT = 'yyyy-MM-dd HH:mm:ssX',
          source = 'Tanker-Prices'
        }
      )
```

Remark: CSVFILESOURCE is a short hand for ACCESS as above. Typically, it is better to use this form, as options are provided as parameters.

CQL

File Transport Handler

```
CREATE STREAM source (id Double, data STRING)
    WRAPPER 'GenericPush'
    PROTOCOL 'CSV'
    TRANSPORT 'File'
    DATAHANDLER 'Tuple'
    OPTIONS ( 'filename' '/src/odysseus/input.csv')


CREATE SINK sink (id Double, data STRING)
    WRAPPER 'GenericPush'
    PROTOCOL 'CSV'
    TRANSPORT 'File'
    DATAHANDLER 'Tuple'
    OPTIONS ( 'filename' '/src/odysseus/output.csv')
```