

## fhicl-cpp - Feature #8529

### Cross-platform end-of-line format

04/29/2015 12:43 PM - Gianluca Petrillo

<b>Status:</b>	Closed	<b>Start date:</b>	04/29/2015
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Kyle Knoepfel	<b>% Done:</b>	100%
<b>Category:</b>		<b>Estimated time:</b>	4.00 hours
<b>Target version:</b>	1.15.00	<b>Spent time:</b>	6.00 hours
<b>Description</b>			
Now that Mac platform is supported, I have started receiving reports of cryptic errors like:  Failed to parse the configuration file 'hack_evd.fcl' with exception ---- Malformed #include directive: BEGIN #include "evdservices_microboone.fcl" at line 1 of file ./hack_evd.fcl ---- Malformed #include directive: END  It was due to the fact that the termination of the #include line was "\x0A\x0D" and the parser was not recognizing it. I suspect that that FHiCL file was created on a different platform (I actually think "\x0A\x0D" is DOS/Windows format) and copied it to a different machine... or whatever. Also, it seems it's a problem specific to the #include directive or, at least, not general.  The point is, to find what's the problem there is sure source of grief for most people.			

#### History

##### #1 - 05/04/2015 11:36 AM - Kyle Knoepfel

- Tracker changed from Bug to Feature
- Status changed from New to Accepted
- Estimated time set to 2.00 h

We will take a look and fix it.

##### #2 - 05/04/2015 01:03 PM - Kyle Knoepfel

- Status changed from Accepted to Feedback

Gianluca, we are unable to reproduce this error. Can you send us a sample FHiCL file with which you are having problems.

##### #3 - 05/04/2015 01:55 PM - Gianluca Petrillo

- File DumpWires\_dos.fcl added
- File DumpWires\_mac.fcl added
- File DumpWires\_unix.fcl added

Mirabile dictu, I can.

The following test is run on a SLF6 machine (uboonegpvm06.fnal.gov).  
I have saved the same file with KDE editor, kate, using three different End-Of-Line styles.  
The first line of the FHiCL file is just a hash symbol...

Running on a UNIX-style EOL:

```
$ hexdump -C DumpWires_unix.fcl | head -n 1
00000000 23 0a 23 20 46 69 6c 65 3a 20 20 20 20 20 44 75 |#.# File: Du|
$ lar --debug-config test.cfg -c job/DumpWires_unix.fcl
** ART_DEBUG_CONFIG is defined: config debug output to file test.cfg **
Art has completed and will exit with status 1.
```

As you can see, hexdump shows that the termination line (second character in the file) is "\x0A".  
The parser is happy.

Running on a Mac-style EOL:

```
$ hexdump -C DumpWires_mac.fcl | head -n 1
00000000  23 0d 23 20 46 69 6c 65 3a 20 20 20 20 20 44 75 |#.# File:  Du|
$ lar --debug-config test.cfg -c job/DumpWires_mac.fcl
Failed to parse the configuration file 'job/DumpWires_mac.fcl' with exception
---- Parse error BEGIN
  Local lookup error
  ---- Can't find key BEGIN
    microboone_geometry_helper (at part "microboone_geometry_helper")
  ---- Can't find key END
    at line 1, character 375, of file "./job/DumpWires_mac.fcl"

    ExptGeoHelperInterface: @local::microboone_geometry_helper # defined in geometry_microboon Geometry:
                            @local::microboone_geo # defined in geometry_microboon} # physics: [ ana ] "
daq"ata"sed); in the comments, defaults are reportede)

    ^
---- Parse error END

Art has completed and will exit with status 7002.
```

Here the termination is "\x0D", that FHiCL parser hates but only in the #include line.

If I place a UNIX termination line before the #include line, or if the #include line is the first in the file, I will get a different error (Malformed #include directory). If I add also a UNIX terminator at the end of the #include line, it becomes content.

Finally, running on a DOS-style EOL:

```
$ hexdump -C DumpWires_dos.fcl | head -n 1
00000000  23 0d 0a 23 20 46 69 6c 65 3a 20 20 20 20 20 44 |#.# File:  D|
$ lar --debug-config test.cfg -c job/DumpWires_dos.fcl
** ART_DEBUG_CONFIG is defined: config debug output to file test.cfg **
Art has completed and will exit with status 1.
```

Here the termination is "\x0D\x0A", and everything is well.

I suspect that "\0d" does not break a line, but since "\0d" is still considered a blank character and FHiCL is mostly able to be written on a single line, that typically does not matter. I am unsure how comment parsing enters this explanation though.

Also note that #include lines are already special in that they can't accept a comment at their end (I am tempted to call this a bug).

The original configuration files I used are attached and, in case the uploading messes with EOL terminations, also available in /uboone/data/users/petrillo/LArSoft/develop/prof/job (accessible from any MicroBooNE GPVM, e.g. uboonegpvm06.fnal.gov).

#### #4 - 05/06/2015 10:29 AM - Kyle Knoepfel

- Status changed from Feedback to Assigned

- Estimated time changed from 2.00 h to 4.00 h

Thank you for the configuration files. We now understand what the problem is. In order to insert "#include" files into the final FHiCL file, we depend on std::getline, whose default delimiter is '\n'. For the UNIX and DOS format, both include the '\n' to signify the end of a line (DOS files also include '\r' before EOL). However, for older-style Mac file formats, the newline is signified only by the carriage return '\r'. The fix is relatively straightforward and almost fully implemented.

#### #5 - 05/06/2015 10:29 AM - Kyle Knoepfel

- Assignee set to Kyle Knoepfel

#### #6 - 05/07/2015 01:37 PM - Kyle Knoepfel

- Status changed from Assigned to Resolved

- % Done changed from 0 to 100

The cetlib includer functionality now supports file formats where newlines are specified only by carriage returns.

Implemented with [cetlib:01c9a1746a7a75141f4c8ad52e8eb44edc2c5d26](https://github.com/celestiaorg/cetlib/pull/1746a7a75141f4c8ad52e8eb44edc2c5d26).

#7 - 07/01/2015 06:24 PM - Christopher Green

- Status changed from Resolved to Closed

- Target version set to 1.15.00

**Files**

---

DumpWires_dos.fcl	2.23 KB	05/04/2015	Gianluca Petrillo
DumpWires_unix.fcl	2.14 KB	05/04/2015	Gianluca Petrillo
DumpWires_mac.fcl	2.14 KB	05/04/2015	Gianluca Petrillo