Gromacs - Feature #1209
Reduce tools verbosity
03/25/2013 05:49 PM - Mark Abraham

Status: Closed
Priority: Normal
Assignee: Teemu Murtola
Category: core library
Target version: 5.0

Description
I think the verbosity of CopyRight() is far too high. I think we're rather out of line with the way most *nix toolsets work, and I'm sure Sziarlı agrees.

The statements of authors, code copyright status and licensing details are unnecessary. The license does not restrict the use of the code. Ordinarily, the purpose of asserting copyright and authorship is to provide force to the license, but that is not applicable to output files. I don't think mdrun should write these, either.

Identifying the (suffixied) tool name, perhaps some compilation settings, and the code version has some genuine value to the user (and on the mailing list, if anyone provides stdout). I think this should be one or two lines for tools. mdrun warrants greater reporting of compilation settings.

I think printing all this other stuff contributes to users just scrolling to the bottom of anything we output, which decreases the value of everything we output that has genuine information value.

Related issues:
- Related to Gromacs - Bug #1267: Add a GMX_NO_CREDITS variable to not show cre... Closed 05/28/2013
- Related to Gromacs - Feature #1427: GMX_NO_CREDITS and structure of credit se... Rejected 01/23/2014

Associated revisions
Revision 6ff9572c - 07/19/2013 05:25 PM - Teemu Murtola

Uniform and less verbose startup for all binaries.

Now all binaries call gmx::ProgramInfo::init() as more or less the first thing. Removed set_program_name() as unnecessary, since the above call does all that. Make CommandLineModuleManager responsible of also printing any common startup header, and add a method to suppress this stderr output in unit tests.

Replace CopyRight() with a less verbose method and move the responsibility of calling it to parse_common_args().

Added a PCA_STANDALONE flag for parse_common_args() to know when to do this (can be removed once all programs are part of the wrapper binary).

Left the CopyRight() function still there, in case we want to add a -license command-line option or similar to print that information. Currently, it is not called from anywhere. Removed GMX_NO_CREDITS and some other unnecessary code from CopyRight().

Most stuff in contrib is broken by this. There may be more changes coming to the initialization sequence, and it will be simple to adapt those programs that people want to get working again, but it is not worth doing it more than once.

Part of #1209.

Change-Id: i5403dd259ab5f314ccee3283aac275a6c26d4818d
Don't print the list of program options by default when running the
Gromacs binaries. The command line printed as part of the startup
header should provide any diagnostic value there is to this output.
Replace the previous -quiet option (that disabled this output) with
a -verbose option that enables this output again. Left the mostly no-op
-quiet option there (there is still a special case of -quiet -h used by
gmx_tune_pme), because it will get new use shortly.
Removed the hidden -verb option whose value wasn't used.

Related to #1209.

Change-Id: I9730e1eb08a2ee35d26fd603f48e89f561dee50a

Add a -quiet option to the wrapper binary. This option suppresses all
startup headers and the gmx_thanx() call in the end. Restructure
CommandLineModuleManager::run() to avoid multiple paths that all need to
call gmx_thanx() and potentially also print the startup header.

Make -quiet suppress the startup header also in parse_common_args().
Reordering required by this causes invalid command-line option error
messages to be printed before the header gets printed, but this only
affects those few binaries that don't go through the wrapper binary.

Make gmx_print_version_info() callable through printBinaryInformation(),
and do this in the wrapper binary and in parse_common_args() in
response to -version.

Print the copyright in response -copyright in the wrapper binary. Easy
to add to parse_common_args() as well, but even better would be to merge
the remaining few programs into the wrapper binary.

Part of #1209.

Change-Id: l0c7dddc91065b12f347da12acd82047e2d94b44c
Revision d67173b3 - 07/21/2013 02:21 PM - Teemu Murtola

Uniform "generated by" information in output files.

Extend the printBinaryInformation() function to make it possible to write the lines in a particular comment format. Use the function to print out "Generated by" information in several output formats. Removed command_line() from statutil.h as it is no longer used. Replaced most uses of Program() with ShortProgram() where the actual path was not relevant.

Also use printBinaryInformation() to print out information into the mdrun log file. Again print the copyright header to the beginning of the log file.

Related to #1209.

Change-Id: Ibf32de318421a10d39f217517005fea408d19bf

Revision e426cc63 - 07/22/2013 06:47 AM - Teemu Murtola

Fix Doxygen warnings.

Ibf32de31 introduced some undocumented members into gmx::BinaryInformationSettings, which were causing Doxygen warnings. Jenkins does not catch these in legacyheaders/…

Related to #1209.

Change-Id: I3c5c0c9818966290d3b298422905604b136d01c3

Revision e45b119f - 07/30/2013 06:10 AM - Teemu Murtola

Handle wrapper binary options also for symlinks.

Now the -quiet, -version, and -copyright options also work when the binary is invoked through a symlink (so, e.g., g_angle -quiet suppresses the startup headers). They also work for the user tools, implemented using the single-module wrapper.

Extended the command-line parser to support parsing only recognized options and removing those from the command line while leaving the rest untouched.

Related to #685 and #1209.

Change-Id: I740f70386d89694246c3e25ba0a1c1c4df17dc6b
Based on discussion in earlier changes, made the copyright on by default. It is now possible to suppress it while keeping other parts of the startup header with -nocopyright.

Simplify the copyright printing code and make the output more in line with the rest of the startup header (left-aligned etc.). Move the copyright to the beginning of the header to have all diagnostic information available under it; that makes it easier for the user (and for us in bug reports) to spot all the relevant information.

Also adjust the wording of the license text; LGPL does not automatically place a user program that calls GROMACS under the same license, so changed "This program" to "GROMACS". Made a similar clarification to the "Written by" text.

Part of #1209.

Change-Id: If104074995db891702dd9eb410fe811408add4ad

History

#1 - 04/04/2013 06:55 PM - Teemu Murtola
This should be straightforward to do, in particular after #685 is more thoroughly implemented (i.e., after there is only one or a few places in the code where this information gets printed).

#2 - 04/26/2013 05:46 AM - Teemu Murtola
- Status changed from New to Accepted

#3 - 06/29/2013 06:45 AM - Teemu Murtola
- Assignee changed from Mark Abraham to Teemu Murtola

I started making the initialization and finalization of the binaries more uniform now that #685 is mostly implemented. Can do this as part of that as well. Requires converting the remaining main() functions to C++, which in turn requires #1294 to avoid compiler warnings.

#4 - 06/29/2013 09:20 PM - Teemu Murtola
- Status changed from Accepted to In Progress

Found a way to remove some include file dependencies to avoid solving #1294 right now. But that is going to bit as soon as someone wants to introduce more C++ stuff in mdrun.
#5 - 07/02/2013 03:13 PM - Peter Kasson
My suggestion (after commenting on Teemu's CL) would be that we have CopyRight() verbosity on by default for mdrun and logs but not for tools. Does that make sense? Teemu, is that even what you're currently implementing?

#6 - 07/02/2013 07:07 PM - Teemu Murtola
My current implementation (https://gerrit.gromacs.org/#/c/2472/ and https://gerrit.gromacs.org/#/c/2473/) don't currently print the authors/license anywhere. I would prefer to first clean up the code that potentially needs to print this information (including the code that writes output files, so that also this output would go through a single function), and then decide on what to print where. I will anyways implement a -quiet option for the wrapper binary that suppresses all this output, and possibly some others as well; if there is more control over what really gets printed, then it is less important what gets printed out by default. But explicit requests to cite some papers will probably be more efficient than a generic list of authors. I don't think that we should change the default visibility of those "Please cite" notices.

#7 - 07/05/2013 01:14 PM - Teemu Murtola
- Status changed from In Progress to Fix uploaded

Changes up to https://gerrit.gromacs.org/#/c/2484/ should resolve the issues; it should be easy to adapt the code after that to any default while still giving the user an option to print extra information or be totally quiet. I have a few more in-progress refactoring changes that are not directly related, but make the default header a bit more useful, but marked the issue to "Fix uploaded" nonetheless.

#8 - 07/05/2013 07:04 PM - Erik Lindahl
Hi!
Haven't had time to download the patch (and my C++ still isn't good enough to make it obvious without compiling...), but in general I very much support this idea. We can continue working on the details, but I think Teemu has the balance pretty much right. I would like something like:

- Remove most of our current output. It's better to only write a few things, but the stuff we do write we really expect the user to read!
- Since "mdrun" is our main program, I like the idea to have more output there, although it might be more important with some text saying citation information and documentation is available in the log, rather than the copyright?
- In particular when we have this, I would move all paper-citation-requests to the log file (but if people think it's important we could have a short info to read the logfile for algorithm references)
- In general I would like a "-q" option that silences everything except for critical errors, and a "-v" option that gives me verbose output :-)

However, rather than getting in the habit of nitpicking about details I would suggest pushing Teemu's code in, and then we can adjust things later if we want to - there's still quite a bit of time until the next major release.

It's of course important that good algorithms in the tools get cited, but I don't think it has a huge effect to write something on stdout when people write the paper 6 months later. Instead, perhaps we should add information on our wiki where we summarize the algorithms used, what they do, and suggest good ways to cite them (possibly including endnote and bibtex-format entries). My experience is that people want to cite things, but it can be a drag to look up the references for standard in lots of different places!

#9 - 12/31/2013 07:01 AM - Teemu Murtola
- Project changed from Next-generation analysis tools to Gromacs
- Subject changed from reduce tools verbosity to Reduce tools verbosity
- Category set to core library
- Status changed from Fix uploaded to Resolved

#10 - 01/10/2014 11:37 PM - Rossen Apostolov
- Status changed from Resolved to Closed

Closing too, please reopen if needed.

#11 - 01/23/2014 04:48 PM - Teemu Murtola
- Related to Feature #1427: GMX_NO_CREDITS and structure of credit section added