GROMACS - Bug #2086

Crashes/memory problems with selections

12/02/2016 09:26 PM - Chris Neale

Status: Closed
Priority: Low
Assignee: Teemu Murtola
Category: selections
Target version: 2016.2
Affected version - extra info: all earlier versions with selection code
Affected version: 5.1.2

Difficulty: uncategorized

Description
I am not sure if this is an error in gmx distance or simply an indication that the program doesn't fail gracefully when I do something idiotic.

I have attached a .pdb file and a script for running gmx distance commands that works on some selections but gives a memory allocation error on other selections. I believe that all selections should be valid.

More information is available here:
https://mailman-1.sys.kth.se/pipermail/gromacs.org_gmx-users/2016-November/109623.html

Associated revisions

Revision c848f4a4 - 12/07/2016 02:03 AM - Teemu Murtola
Fix use of position variables with plus/merge

If a selection contained a position variable (e.g., 'com of ...') that was used more than once, and at least one of those uses was with plus/merge, there were out-of-bounds memory writes. This was caused by the internal position structure not getting fully initialized. Incomplete initialization happens in all contexts with such variables, but only plus/merge (and possibly permute) actually use the values that remained uninitialized, which caused them to incorrectly compute the amount of memory required to store the result.

Fixes part of #2086.

Change-Id: I016e796db268a11d557309935c02cbd1bc79a83c

Revision b7817e2d - 12/08/2016 05:21 PM - Teemu Murtola
Fix possible memory error with long selections

If a selection was more than 1000 characters long and there was a whitespace exactly at the 1000 point, a buffer overflow could occur. Replaced the buffer with std::string, simplifying the code significantly. Update the generated code to use a newer flex, which also removes the need for some suppressions.

Should fix #2086.

Change-Id: i56513bcf5ee99f05ce144461740d0f868be10186

History

#1 - 12/03/2016 09:53 AM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #2086.
Uploader: Teemu Murtola (teemu.murtola@gmail.com)
Change-Id: i56513bcf5ee99f05ce144461740d0f868be10186
Gerrit URL: https://gerrit.gromacs.org/6354

#2 - 12/03/2016 09:17 PM - Teemu Murtola
- Category changed from analysis tools to selections
This looks like a rare memory error in the selection code, where a buffer can be overrun by one (a zero gets written past the end of a buffer) if your selections is longer than about 1000 characters, and you happen to have a word break at exactly that 1000 character mark (or at higher multiples of 1000).

#3 - 12/04/2016 07:04 AM - Chris Neale
just wondering if you checked your fix against my second group of selection attempts (included in the script originally uploaded) that utilized the -sf option in place of the -select option. In the case of -sf, I saw segfaults, though again this could be a case of my not using the tool correctly.

Thank you,
Chris.

#4 - 12/04/2016 07:31 AM - Teemu Murtola
I did not have time to test it against anything, except a simple case that produced a memory error around the call stack mentioned in the original email. Are you saying it doesn't work, or just wondering? This issue does not mention anything about a segfault.

#5 - 12/04/2016 08:19 AM - Chris Neale
My upload contained a .pdb file and a bash script to run the program. That bash script showed first a success, then a failure (both with -select). Then I showed segfaults with the -sf usage. That could be due to my error, though it seemed like a reasonable usage to me based on the docs.

I did not re-test any code after your revisions.

Thank you.

#6 - 12/04/2016 09:13 AM - Teemu Murtola
Your report (in the email) says that you get "the same kind of crash" with both -select and -sf. The symptoms would be easily explained if the failure is related to the length of the selection, not to anything in gmx distance. My fix is in code that it's executed identically on both paths. I can't easily check the attachment when 95% of the time I'm looking at these issues from a mobile client. There can be other issues than just the one I fixed (your email actually seems to have different selections for -select and -sf), but that one certainly fixes one memory access problem that is easily triggered by your long selection for the -select case.

#7 - 12/05/2016 08:31 PM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #2086
Uploader: Teemu Murtola (teemu.murtola@gmail.com)
Change-Id: gromacs~release-2016~I016e796db268a11d557309935c02cbd1bc79a83c
Gerrit URL: https://gerrit.gromacs.org/6359

#8 - 12/05/2016 08:35 PM - Teemu Murtola
There was another issue that was triggered if you used a variable with a position value (like com of ...) with plus or merge more than once. That is fixed by the second linked change.

In the future, it would be good not to hide essential information about the nature of the bug report into comments in a script within a compressed attachment.

#9 - 12/05/2016 08:37 PM - Chris Neale
OK, sorry. Thanks for fixing it.

#10 - 12/07/2016 05:01 AM - Mark Abraham
Teemu's second fix certainly fixes the segfault arising from the implementation of "plus".

Chris, I didn't observe any problem with your "residues 22 to 40" command in analyse.sh, using either release-2016 branch or 5.1.2. I got plausible -oxyz output in both cases. Are you still able to reproduce the issue? With release-2016 branch? Does Teemu's first fix resolve the issue with really long selection strings? Thanks

#11 - 12/11/2016 01:23 PM - Teemu Murtola
- Subject changed from gmx distance has possible memory problems to Crashes/memroy problems with selections
- Status changed from Fix uploaded to Resolved

#12 - 12/12/2016 08:24 PM - Chris Neale
Probably I did not download the correct source or something, but I still have problems.

I obtained the new code today like this:
git clone git://git.gromacs.org/gromacs.git

Here is the run command and the output:

```bash
$ gmx distance -s my.pdb -f my.pdb -select "$inpstr" -oxyz output_22-40.xvg
:-) GROMACS - gmx distance, 2017-dev-20161210-dd11612 (-:
```

GROMACS is written by:

<table>
<thead>
<tr>
<th>Name</th>
<th>Name</th>
<th>Name</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emile Apol</td>
<td>Rossen Apostolov</td>
<td>Herman J.C. Berendsen</td>
<td>Par Bjelkmar</td>
</tr>
<tr>
<td>Aldert van Buuren</td>
<td>Jiri Kraus</td>
<td>Anton Feenstra</td>
<td>Gerrit Groenhof</td>
</tr>
<tr>
<td>Christoph Junghans</td>
<td>Anca Hamuraru</td>
<td>Vincent Hindriksen</td>
<td>Dimitrios Karkouli</td>
</tr>
<tr>
<td>Peter Kasson</td>
<td>Jiri Kraus</td>
<td>Carsten Kutzner</td>
<td>Per Larsson</td>
</tr>
<tr>
<td>Justin A. Lemkul</td>
<td>Magnus Lundborg</td>
<td>Pieter Meulenhoff</td>
<td>Erik Marklund</td>
</tr>
<tr>
<td>Teemu Murtoila</td>
<td>Szilard Fall</td>
<td>Sander Pronk</td>
<td>Roland Schulz</td>
</tr>
<tr>
<td>Alexey Shvetsov</td>
<td>Michael Shirts</td>
<td>Alfons Sijbers</td>
<td>Peter Tieleman</td>
</tr>
<tr>
<td>Teemu Virolainen</td>
<td>Christian Wennberg</td>
<td>Maarten Wolf</td>
<td></td>
</tr>
</tbody>
</table>

and the project leaders:

<table>
<thead>
<tr>
<th>Name</th>
<th>Name</th>
<th>Name</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mark Abraham</td>
<td>Berk Hess</td>
<td>Erik Lindahl</td>
<td>David van der Spel</td>
</tr>
</tbody>
</table>

Copyright (c) 1991–2000, University of Groningen, The Netherlands.
Copyright (c) 2001–2015, The GROMACS development team at Uppsala University, Stockholm University and the Royal Institute of Technology, Sweden.
check out http://www.gromacs.org for more information.

GROMACS is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 2.1 of the License, or (at your option) any later version.

GROMACS: gmx distance, version 2017-dev-20161210-dd11612
Executable: /scratch/cneale/xe/GROMACS/exec/gromacs_git/2016_Dec12/serial/bin//gmx
Data prefix: /scratch/cneale/xe/GROMACS/exec/gromacs_git/2016_Dec12/serial
Working dir: /scratch/cneale/work/2016/December2016/2/gmxDistanceErrorFiles
Command line:
```bash
gmx distance -s my.pdb -f my.pdb -select 'com of (residue 22 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 23 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 24 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 25 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 26 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 27 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 28 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 29 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 30 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 31 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 32 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 33 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 34 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 35 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 36 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 37 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 38 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 39 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue 40 and name CA) plus com of (resname POPC or resname POPS) plus com of (residue POPC or resname POPS) -oxyz output_22-40.xvg
```

*** glibc detected *** gmx: realloc(): invalid next size: 0x00000000022c8450 ***

Thank you,
Chris.

#13 - 12/12/2016 08:27 PM - Teemu Murtola
The fix has not been merged to the master (2017) branch yet.

#14 - 12/12/2016 09:02 PM - Chris Neale
Are there instructions online about how I can get the fixed code branch?

I went here: https://gerrit.gromacs.org/#/c/635

Under the Download pulldown arrow, I downloaded the contents of "Archive -> tgz"
same error.

Thank you,
Chris.

#15 - 12/13/2016 07:32 AM - Teemu Murtola
That should work, although if you have a git repository already, using the checkout link under Download could be more useful. But your link is broken, so we don't know which fix you downloaded and whether that version contains both fixes or not. Easiest is just to checkout the release-2016 version from git, since both fixes are already merged there. Builds from git also have an unambiguous version string the identifies the exact commit used.

#16 - 01/20/2017 02:05 PM - Mark Abraham
Is there any update on whether Chris's issue is resolved yet?

#17 - 01/20/2017 03:44 PM - Chris Neale
Either it's not fixed or the download button for that code patch didn't work. Teemu said above that what I did should have given the new code, so if that is true then the proposed fix did not work. Not sure that he meant by "your link is broken"... copy and paste https://gerrit.gromacs.org/#/c/635 worked for me just now. Teemu's instructions about git were insufficient for me to get the code that way (I know nothing about git).

#18 - 01/20/2017 04:34 PM - Aleksei Iupinov
Chris Neale wrote:

Either it's not fixed or the download button for that code patch didn't work. Teemu said above that what I did should have given the new code, so if that is true then the proposed fix did not work. Not sure that he meant by "your link is broken"... copy and paste https://gerrit.gromacs.org/#/c/635 worked for me just now. Teemu's instructions about git were insufficient for me to get the code that way (I know nothing about git).

Ah, he meant that your link links one character in the end, so it links to some ancient abandoned change instead. The actual link is https://gerrit.gromacs.org/#/c/6354, try it.

#19 - 01/20/2017 04:36 PM - Aleksei Iupinov

your link links one character

*your link misses one character, sorry for dyslexia.

#20 - 01/20/2017 05:01 PM - Mark Abraham
Aleksei Iupinov wrote:

Chris Neale wrote:

Either it's not fixed or the download button for that code patch didn't work. Teemu said above that what I did should have given the new code, so if that is true then the proposed fix did not work. Not sure that he meant by "your link is broken"... copy and paste https://gerrit.gromacs.org/#/c/635 worked for me just now. Teemu's instructions about git were insufficient for me to get the code that way (I know nothing about git).

Ah, he meant that your link links one character in the end, so it links to some ancient abandoned change instead. The actual link is https://gerrit.gromacs.org/#/c/6354, try it.

Looks like Teemu already reported he tried it on Chris's inputs and it was good. So we'll regard it as fixed unless Chris finds further problems some time.

#21 - 01/20/2017 05:01 PM - Mark Abraham
- Status changed from Resolved to Closed

#22 - 01/20/2017 07:28 PM - Chris Neale
Thank you Aleksei. I checked and that was a copy/paste error on my part. What I actually had downloaded was 6359 (as suggested in above post #7). I didn't realize that there was another subsequent separate fix (6354). To summarize, either one of these code changes fixes the second error I reported, but neither fixes the first error. Perhaps they need to be combined or something. I give more details below but I am fine with Mark's suggestion to let it drop.

The following command:

```
inpstr=$(for((i=22;i<=40;i++)); do echo -n "com of (residue $i and name CA) plus com of (resname POPC or resname POPS) plus "; done|sed "s/plus $//")
gmx distance -s my.pdb -f my.pdb -select "$inpstr" -xyz output_22-40.xvg
```

02/21/2020
still gives the error:

*** glibc detected *** /scratch/cneale/exe/GROMACS/exec/gromacs_git2/serial/bin/gmx: realloc(): invalid next size: 0x00000000022beba0 ***

with either gerrit 6354 (b7817e2.tar.gz) or gerrit 6359 (c848f4a.tar.gz)

However, both gerrit 6354 or gerrit 6359 do fix the second error that I reported, which is a segfault on:

```bash
{ 
  echo "bilayer=com of (rename POPC or rename POPS);" 
  for((i=2;i<=20;i++)); do 
    echo "residue $i and name CA plus bilayer;" 
  done 
} > input.selection

gmx distance -s my.pdb -f my.pdb -sf input.selection -oxyz outputB_2-20.xvg
```

#23 - 01/20/2017 07:40 PM - Chris Neale

My mistake. My above post no. 22 is wrong.

both gerrit 6354 and 6359 fix the second error. Only gerrit 6354 fixes the first error.

I agree that the fix is good.

Thank you Teemu for the fix and Mark and Aleksi for helping me apply it.

#24 - 01/23/2017 04:39 PM - Mark Abraham

OK great. Thanks Chris!

Files

| gmxDistanceErrorFiles.tgz | 564 KB | 12/02/2016 | Chris Neale |