create heuristic for c_skipNeutralAtoms

03/16/2020 09:50 PM - Szilárd Páll

<table>
<thead>
<tr>
<th>Status:</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>Normal</td>
</tr>
<tr>
<td>Assignee:</td>
<td>Jonathan Vincent</td>
</tr>
<tr>
<td>Category:</td>
<td></td>
</tr>
<tr>
<td>Target version:</td>
<td></td>
</tr>
<tr>
<td>Difficulty:</td>
<td>uncategorized</td>
</tr>
</tbody>
</table>

Description

 Skipping zero-charge atoms can have significant performance benefit (measured ~10-15% spread+gather time a while ago before the new spline recompute). The initial implementation hardcoded the c_skipNeutralAtoms=false.

History

#1 - 03/16/2020 09:53 PM - Szilárd Páll
- Description updated

#2 - 03/17/2020 11:22 AM - Jonathan Vincent
- Assignee set to Jonathan Vincent

So what is a good set of input to test this on?

We have been using water boxes previously which will not work as there are no zero charge atoms.

Would ADH and similar work?

What we have right now are

- villin -> 5,032 atoms
- rnase dodec -> 16,816 atoms
- JAC (DHFR) -> 23,558 atoms
- ADH dodec -> 95,561 atoms
- Prace A -> 141,577 atoms
- Cellulose -> 408,609 atoms
- STMV -> 1,066,628 atoms

Do you have any feeling on what will affect the performance between skipping the neutral atoms or not in terms of input?

#3 - 03/17/2020 12:12 PM - Szilárd Páll

I don't think we need to test all of those inputs. We should think of what are the conditions that have an impact on performance and pick accordingly.

For this parameter I think the aspects that matters are:
- input size / size relative to hardware "size" and uarch
- fraction of zero charge atoms
- atom order (DD sort?)

Have I missed something?