resplit_1.0 软件包使用说明

欧念森

2015 年 7 月

目录

1 关于 resplit_1.0 软件包

重启 (restart) 是海洋模式所必需的一种技术。因为海洋模式进行气候模拟时，往往需要几星期甚至几个月的计算时间。且目前广泛使用的海洋模式多远程运行于大型超级计算机上，而超级计算机的复杂性，更增加了模式运行意外中断的概率。为保证模式中断时，已经得出的计算结果不至于浪费掉，就需要模式具备重启技术，使得故障恢复后, 可以使用中断前的结果继续计算。

目前预报中心所使用的 NEMO 模式，也具备了重启技术。然而，由于并行计算相对与串行计算的复杂性，目前 NEMO 在中断后重启时，所使用的 CPU 数必需与中断前的一致。但是由于中断前后提交的作业情况不一样，可用的 CPU 数也就往往不一样。若中断后可用的 CPU 数小于中断前模式使用的 CPU 数，NEMO 将无法重启。因此，若能以变 CPU 数重启 NEMO，那么重启 NEMO 时，如果可用的 CPU 数较多，就可以使用比中断前更多的 CPU 运行，加快模式运行速度；如果可用的 CPU 数较少，就可以使用比中断前更少的 CPU 马上运行，而不必等待其他作业结束后再运行模式。这样将会给 NEMO 的运行带来很大的灵活性，提高工作效率。

resplit_1.0 是为了实现 NEMO 变 CPU 数重启而编写的软件包，使得 NEMO 可以根据当前超级计算机上的可用 CPU 资源，灵活选用不同于中断前的 CPU 数重新启动。

2 使用要求

2.1 软件要求

resplit_1.0 软件包不依赖于其他的 NEMO 软件包，只需要系统安装有 Fortran 编译器和对应的 NetCDF 库即可。这些也是 NEMO 模式所需的。故只要所在的系统环境能正常编译运行 NEMO 模式，就能编译运行 NEMO 软件包。

resplit_1.0 的开发和测试，是在国家超级计算天津中心上完成的，所使用的编译器为英特尔 Fortran 编译器。

2.2 数据要求

resplit_1.0 的算法原理是将原来 NEMO 输出的 N（N 为原来使用的 CPU 个数）个 restart 场，先拼接为一个全球场，然后再重新分割为 M（M 为准备使用的新 CPU 数）个 restart 场。因此，除了像正常重启时所需要的 N 个 restart 场文件外，resplit_1.0 还需要以下数据：

1. 将全球场分割成 N 个子区域的分割信息文件（各个子区域的经纬度、起止格点坐标在全球网格中的位置等，由 resplit_1.0 包中的 create_domain_info 子程序从 N 个 restart 场文件中提取得到，且只需要提取一次)。

2. 将全球场分割成 M 个子区域的分割信息文件（关于如何得到这一信息文件，见下文)。

3. 当前分辨率下全球网格点的经纬度坐标信息数据 mesh_mask.nc。由于 NEMO 采用了剔除陆地块的技术来减少 CPU 的使用量，而被剔除的陆地块虽然没有海洋变量信息，却依然有经纬度信息。当由 N 分割变为 M 分割时，部分原先被剔除的陆地点有可能又需要重新纳入计算格点，因此 mesh_mask.nc 是必需的（注意由于 NEMO 将海洋变量在陆地点上设为缺测值，因此海洋变量不存在像经纬度坐标那样在剔除陆地块时信息丢失的问题)。

对于如何将全球区域分割成若干个子区域，NEMO 有专门的软件包（MPP_PREP-1.0)。为了保证区域划分的准确，resplit_1.0 要求上面第 2 步的 M 值是某个已经跑过的 CPU 数，并且其 M 个 restart 文件仍然存在（当其分割信息文件已经提取出来后，则不再需要 M 个 restart 场文件)。鉴于 NEMO 的 restart 场数据量往往达几百 G，建议用户先用一些可能常用的 CPU 数跑出 restart 场，再用 create_domain_info 子程序提取出各个 CPU 数的分割信息文件（大小为几十 K）后，便可将这些临时的 restart 场删除以节省空间。

目前，resplit_1.0 已经提取出全球 $4322 \times 3059 \times 75(x \times y \times z)$ 分辨率下，$N = 480, 600, 720, 840, 960$，5 种 CPU 数所对应的分割信息文件。用户可以在这几种 CPU 数中自由切换，而不必再去提取分割信息文件。

3 使用步骤

假设在 NEMO4322 × 3059 × 75 分辨率下, 原来使用 N 个 CPU 运行，当使用上次的 restart 场重启时，希望 M 个 CPU 重启运行。当 N 和 M 取不同值时，具体步骤稍有不同。以下举两例说明。下文的 $root_dir 指代 resplit_1.0 软件包目录 /WORK/home/qhyc1/zhangyu/ouniansen/resplit_1.0

3.1 例 1

N=960, M=600.

由于 960, 600 对应的分割信息文件已事先提取好了, (见 $root_dir/domain_info/中的 N0960.nc 和 N0600.nc)，故只需以下步骤即可将 960 个 restart 场重新分割为 600 个 restart 场:

1. 进入 $root_dir 目录，修改 namelist_main 文件，如下:

```
&param
  old_dir  = "/WORK/home/qhyc1/zhangyu/ouniansen/EXP_960/",
 new_dir = "/WORK/home/qhyc1/zhangyu/ouniansen/test/EXP_960_to_600/",
 old_ncpu  = 960,
 new_ncpu  = 600,
rst_prefix = "ORCAR12_00000720_restart_",
/
```

   其中，old_dir 是中断前的 restart 场目录，在本例中，该目录含有 960 个 restart 文件及 960 个海冰 restart 文件，NEMO 将要在这些 restart 场上启动。new_dir 则是 resplit_1.0 的输出目录, 用来存放重新分割后的 restart 场 (若该目录不存在, resplit_1.0 会自动创建)。old_ncpu 是中断前使用的 CPU 个数, new_ncpu 是重启时准备使用的 CPU 个数。rst_prefix 是 restart 文件的前缀，resplit_1.0 根据这一前缀确定 restart 场的具体文件名。

2. 运行脚本 yhbatch_main.sh，等待脚本在作业系统中运行完成后 (通常需要一个多小时)，即可到 new_dir 中查看分割结果。

3.2 例 2

N=360, M=1080.

由于 360, 1080 对应的分割信息文件在 $root_dir/domain_info/中并不存在，故需要按以下步骤先生产分割信息文件。

首先生成分割信息文件 N0360.nc:

1. 进入 $root_dir 目录，修改 namelist_create_domain_info 文件，如下:

```
&param
  rst_dir  = "/WORK/home/qhyc1/zhangyu/ouniansen/EXP_360/",
rst_prefix = "ORCAR12_00000720_restart_",
    ncpu  = 360,
/
```

   其中，rst_dir 是 N=360 时的 restart 场目录，ncpu 是对应的 CPU 数。

2. 运行脚本 yhbatch_create_domain_info.sh，等待脚本在作业系统中运行完成后(通常需要几分钟)，resplit_1.0 将在 $root_dir/domain_info/中生成 N0360.nc 文件。

其次生成分割信息文件 N1080.nc。由于 N=1080 对应的 restart 场并不存在，因此需要另外使用 1080 个 CPU，在当前分辨率下，积分一小段时间，得到相应的 restart 场，再用以上方法生产分割信息文件 N1080.nc。得到 N1080.nc 后，刚才临时跑出的 restart 场即可删除。

至此，两个分割信息文件均已得到，即可按例 1 的步骤进行分割。

3.3 其他例子

若模式分辨率不再是 resplit_1.0 中 $root_dir/domain_info/目录下 mesh_mask.nc 对应的分辨率，那么则需要将新的 mesh_mask.nc 拷贝至 $root_dir/domain_info/下，再删除 $root_dir/domain_info/中所有的信息分割文件，参照例 2 重新生成新分辨率下的信息分割文件，再参照例 1 完成 restart 场的再分割。

4 重新编译

目前，resplit_1.0 已经在天河一号编译好，见 $root_dir/bin/下的两个二进制文件 main 和 create_domain_info。若用户希望重新编译这两个二进制文件，可进入 $root_dir/utils/, 分别运行 compile_main.sh 和 compile_create_domain_info.sh 这两个脚本。

5 其他

关于 resplit_1.0 的详细算法、代码结构、附带工具程序等，见另一文档《resplit_1.0 软件包开发手册》。

resplit_1.0 软件包是作者受国家海洋环境预报中心委托开发，该软件包所有权属于国家海洋环境预报中心。用户关于 resplit_1.0 的使用、可靠性等问题，可联系国家海洋环境预报中心的张宇博士（zhangy@nmefc.gov.cn）；关于具体算法、代码细节等问题，可联系作者欧念森（ouyuyuan@lasg.iap.ac.cn）。