

ОБ ИСПОЛЬЗОВАНИИ ГРИД-ТЕХНОЛОГИЙ В ЗАДАЧАХ РЕАНАЛИЗА МОРСКИХ ТЕЧЕНИЙ

Д.В. Бородин, В.В. Фомин,
В.А. Иванов

Морской гидрофизический институт
НАН Украины
г. Севастополь, ул. Капитанская, 2
E-mail: borodin112@gmail.com

Изложена методология решения задачи ретроспективного анализа морских течений в среде грид. Используемая модель течений базируется на уравнениях мелкой воды, описывающих средние по глубине горизонтальные компоненты скорости течений, и уровень моря. Дано описание программной реализации и технических особенностей решения задачи в среде Украинского Национального Грида.

Введение. На современном этапе развития вычислительных технологий под термином грид подразумевается географически распределенная, согласованная, открытая и стандартизованная среда разделения вычислительных и информационных ресурсов [1]. В настоящее время грид-технологии начинают интенсивно использоваться для проведения фундаментальных и прикладных научных исследований, требующих предельно больших вычислительных ресурсов. В Украине существует Украинский национальный грид (УНГ) – сообщество научно-исследовательских институтов и высших учебных заведений, занимающихся разработкой и внедрением грид-технологий в науку, промышленность, финансовую, социальную, гуманитарную и другие сферы жизнедеятельности. Морской гидрофизический институт НАН Украины (МГИ) входит в число участников УНГ, и вычислительный кластер МГИ [2] является составной частью инфраструктуры УНГ. В настоящее время в МГИ ведутся исследования по внедрению грид-технологий в численное моделирование динамических процессов в Азово-Черноморском бассейне.

В данной работе рассматриваются вопросы, связанные с использованием грид-технологий для решения задач рет-

роспективного анализа морских течений. Ретроспективный анализ (далее реанализ) применяется для изучения статистических характеристик и климатических тенденций полей течений. Его суть заключается в следующем. Для заданной акватории имеем поля ветра из атмосферной модели за достаточно большой период времени T с некоторой дискретностью по пространству и времени. Используя эти данные в качестве форсинга, рассчитываем характеристики течений с помощью подходящей численной модели за тот же период с некоторой другой дискретностью по пространству и времени. Далее полученные массивы данных подвергаются статистической обработке точно так, как если бы это были результаты наблюдений за течениями. Понятно, что при больших T (несколько десятилетий) задача реанализа требует достаточно больших вычислительных ресурсов и использование грид становится актуальным.

Целью данной работы является разработка подходов и программных средств решения задач реанализа морских течений в грид-среде. Расчетная модель базируется на уравнениях мелкой воды, описывающих средние по глубине горизонтальные компоненты скорости течений $u(x,y,t)$, $v(x,y,t)$ и уровень моря $\eta(x,y,t)$. Здесь x , y – горизонтальные координаты. Модель течений представляет собой модифицированный вариант модуля баротропных движений Принстонской океанической модели (*Princeton Ocean Model*) [3], работающий в однопроцессорном режиме.

Методология решения задачи в среде грид. Для эффективного решения задачи в среде грид необходима ее декомпозиция – представление в виде набора независимых (не связанных между собой) подзадач. В общем случае, количество подзадач n определяется спецификой решаемой задачи и производительностью имеющихся вычислительных ресурсов.

При решении рассматриваемой задачи предлагается использовать декомпозицию по времени. Расчетный период $t \in [0, T]$ разбивается на n интервалов переменной ширины ΔT_i , где $1 \leq i \leq n$. На каждом интервале рассчитываются поля

течений и уровня моря по соответствующим этому интервалу полям ветра. Расчет на каждом интервале начинается с нулевых начальных условий ($u = v = \eta = 0$). Для исключения влияния начальных условий расчетные интервалы ΔT_i для $1 \leq i \leq n$ берутся с некоторым перекрытием σ так, что каждый предыдущий интервал захватывает начало следующего (рис. 1).

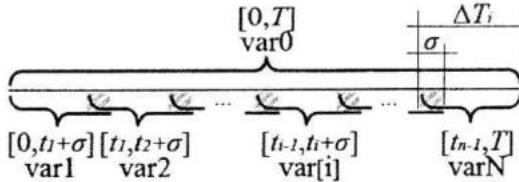


Рис. 1. Декомпозиция модели POM по периоду расчета

Данный подход основывается на том факте, что вид начальных условий влияет на баротропные течения и уровень моря только в первые несколько суток, в то время как ветер является доминирующим фактором на протяжении всего расчетного периода. Как показывает опыт авторов статьи, для Азовского моря приемлемая величина перекрытия интервалов составляет 3 суток.

Программная реализация для среды УНГ. Алгоритм расчетов сводится к запуску в грид-среде несколько копий модели течений. Таким образом, вместо одной «длинной» задачи имеем n «коротких» подзадач, каждая из которых является независимой однопроцессорной копией модели (на практике $n \sim 5-20$). В операционной системе Linux отдельная копия модели течений запускается командой `./pom2d.exe var`, где `pom2d.exe` – исполняемый файл, `var[i]` – имя файла конфигурации, содержащего настроочные параметры: даты начала и конца расчетов; интервал времени через который сохраняются расчетные данные; интервал времени через который создаются резервные копии выходных файлов; параметры расчетной сетки; пути к файлам, содержащим рельеф дна и поля ветра. Результаты расчетов сохраняются в каталогах `output.var[i]` и `restart.var[i]`. В `output.var[i]` накапливаются рассчитанные поля u , v , η , а каталоги `restart.var[i]` содержат информацию, которая исполь-

зуется в качестве начальных данных при возобновлении расчетов.

В УНГ используется программное обеспечение промежуточного уровня ARC проекта Nordugrid [4], для описания заданий в котором используется язык xRSL [5]. Простейший скрипт для запуска отдельной копии модели в среде Nordugrid-ARC выглядит следующим образом:

```
& (executable="pom2d.exe")
  (arguments="var")
  (inputfiles= ("input.dat" ""))
    ("var0" ""))
  (outputfiles=
    ("output.var0/*")
    ("output.var0/")
    ("restart.var0/*")
    ("restart.var0/"))
  (stdout="out_var0")
  (stderr="err_var0")
  (jobname="pom2d")
```

Скрипт передается команде `ngsub`, которая формирует и ставит в очередь грид-задание. Каждому заданию ставится в соответствие ссылка вида: `gsiftp://somecluster:2811/jobs/K`, где K – индивидуальный номер задания. По ссылке проверяется состояние задания (команда `ngstat`) и по готовности скачивается результат (команда `ngget`).

Для запуска серии подобных заданий необходимо:

1. Для каждого конфигурационного файла `var[i]` сгенерировать отдельный xRSL-скрипт и передать его команде `ngsub`.
2. Если в грид-очереди, помимо заданий модели, находятся другие грид-задания, то необходимо отслеживать статус выполнения первых, поскольку сами ссылки не несут никакой смысловой нагрузки.
3. Если часть заданий завершилась аварийно, то необходимо для каждой i -го задания проанализировать каталог `restart.var[i]` и возобновить расчеты с последней просчитанной даты.

Для реализации перечисленных выше шагов было разработано на языке C++ специализированное программное обеспечение, включающее (см. блок-схему на рис. 2):

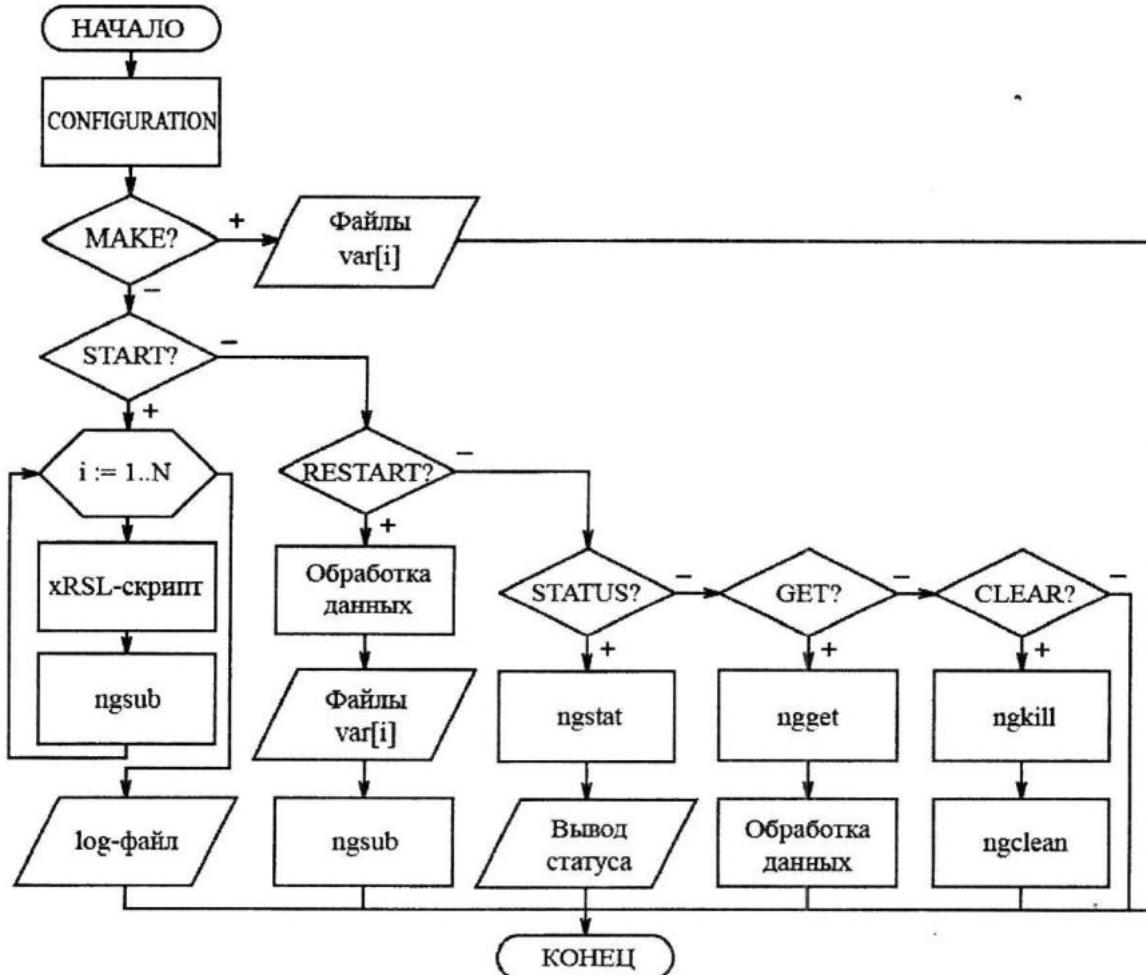


Рис. 2. Блок-схема разработанного программного средства

- модуль пользовательских настроек – «Configuration»;
- модуль формирования конфигурационных файлов $var[i]$ – «Make»;
- модуль генерации xRSL-скриптов и запуска их в грид-среде – «Start»;
- модуль проверки статуса сформированных грид-заданий – «Status»;
- модуль получения результатов выполненных грид-заданий – «Get»;
- модуль отмены запущенных грид-заданий – «Clear»;
- модуль запуска подзадач с последней посчитанной датой – «Restart».

В модуле «Configuration» анализируется конфигурационный файл и инициализируются соответствующие переменные. Данные, указанные в конфигурационном файле определяют: количество заданий; имена временных файлов и каталогов, а также имена лог-файлов; ме-

стоположение входных данных – input.dat и файлов конфигураций подзадач – $var[i]$; уровень отладки поставленных в очередь грид-заданий; количество запрашиваемых ресурсов в гриде и пр.

Модуль «Start» генерирует для каждого задания xRSL-скрипт, ставит его в очередь командой `ngsub`, заполняет лог-файл ссылками на поставленные в очередь задания. Модуль «Status» выдает в стандартный поток вывода информацию о состоянии заданий по лог-файлу с ссылками заданий.

Модуль «Get» выполняет закачку всех завершившихся заданий и последующую обработку информации: сбор и сортировку выходных файлов, файлов стандартных потоков вывода и вывода ошибок, файлов возобновления счета. Модуль «Clear» отменяет выполнение грид-заданий, указанных в лог-файле.

Модуль «Restart» на основе анализа содержимого каталогов restart.var[i] создает новые конфигурационные файлы var[i] по последней просчитанной дате.

Заключение. Разработанное программное обеспечение дает возможность на основе численной модели морской циркуляции решать задачу реанализа морских течений в среде УНГ, что может существенно сократить общее время расчетов по сравнению с традиционными вычислительными технологиями. Данный подход успешно применяется в МГИ для моделирования динамики течений в Азово-Черноморском бассейне. Предложенная методология может быть применена к другим типам задач, например к задаче реанализа ветрового волнения.

Работа выполнена в рамках Государственной целевой научно-технической программы внедрения и использования грид-технологий на 2009 – 2013 годы и Программы информатизации НАН Украины на 2011 г.

СПИСОК ЛИТЕРАТУРЫ

1. *Foster I., Kesselman C., Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations // International Journal of High Performance Computing Applications – 2001. – 15(3). – P. 200 – 222.*
2. *Фомин В.В., Бородин Д.В., Иванов В.А. Вычислительный кластер Морского гидрофизического института НАН Украины // Системы контроля окружающей среды. – Севастополь: МГИ НАН Украины, 2010. – С. 121 – 123.*
3. *Blumberg A.F., Mellor G.L. A description of three dimensional coastal ocean circulation model, in Three-Dimensional Coast Ocean Models. – 1987. – P. 1–16.*
4. *Advanced Resource Connector – <http://www.nordugrid.org>*
5. *Extended Resource Specification Language – <http://www.nordugrid.org/documents/xrsl.pdf>*