
НЕДЕСТРУКТИВНОЕ ПРЕОБРАЗОВАНИЕ ПОСЛЕДОВАТЕЛЬНЫХ ПРОГРАММ В ПАРАЛЛЕЛЬНЫЕ ДЛЯ СУПЕР-ЭВМ

Васенин В.А., Крошилин А.Е., Крошилин В.Е., Рагулин А.Д., Роганов В.А.

- Механико-математический факультет МГУ
- Всероссийский НИИ по эксплуатации АЭС
- НИИ механики МГУ



История вопроса

Мечта прикладника - чтобы суперЭВМ считала ту же самую программу, но, скажем, в сто раз быстрее :-)

Недеструктивный подход удобен для всех основных сторон-участников процесса разработки и модификации ПО:

- менеджеров (минимизация рисков при разработке, улучшается документированность исходного кода)
- специалистов по предметной области (не возникает необходимости детального изучения технологии программирования суперЭВМ)
- специалистов по высокопроизводительным вычислениям (не требуется глубокого погружения в предметную область)

Особенности подхода

Под недеструктивностью здесь понимается отказ от переделывания исходной программы, которая могла создаваться и совершенствоваться годами, и имеет форму, привычную для специалистов соответствующей предметной области.

Вместо этого создается набор специализированных "примесей" (mix-ins), которые, взаимодействуя с исходной программой на разных уровнях и стадиях, превращают ее в высокопроизводительную программу для суперЭВМ.

Такое отделение "высокопроизводительного аспекта" от "содержательного аспекта" подобно тому, как оформление отделено в современных технологиях верстки от содержания, что позволяет, в частности, форматировать и масштабировать один и тот же контент для устройств отображения разного вида и разрешения.

В чем именно проявляются преимущества подхода

- поддержка и доработка оригинальной последовательной версии может продолжаться прежним путем, без суперЭВМ, и без задействования специалистов по высокопроизводительным вычислениям;
- может поддерживаться и отлаживаться сразу несколько параллельных версий для различных платформ;
- отладочные включения могут быть оформлены в виде примесей и добавляться по мере необходимости;
- валидация параллельных версий происходит простым путем сравнения результатов вычислений с результатами, полученными актуальной последовательной версией

Исходный код VS целевая платформа

Исходный код пишется в первую очередь для людей, и в идеале должен всего лишь четко и ясно отражать конструктивную постановку задачи.

Специалистам предметной области приходится изучать, таким образом, хотя бы один из языков программирования.

Между прямолинейно написанной программой и высокопроизводительной программой (то есть программой, прооптимизированной для конкретной целевой платформы) имеется существенная разница. Исходный код может довольно сильно отличаться, зато программа может ускориться на порядок и более. И это пока без всякого параллелизма!

Когда добавляется оптимизация для параллельной суперЭВМ, программа быстро начинает терять и без того уже не слишком изящную форму.

Отладка параллельной программы

Отладка параллельной программы на суперЭВМ также весьма увлекательна.

А если программа работает, то как узнать, правильно ли она работает? После перехода на параллельную версию (по известному закону) ресурсов на поддержку старой программы может уже не хватать.

Распараллеливающие примеси (mix-ins)

Тройка $\langle T, P, L \rangle$, где:

- T - микроправила трансформации кода
- P - новые программные блоки
- L - новые библиотеки для суперЭВМ

Содержательная часть проблемы распараллеливания

Подпространства эквивалентных программ

- Новые измерения "параллельного программирования" и "ускорителей"
- Степени свободы - основа потенциала масштабирования
- Вскрытие потенциала масштабирования - непростая задача
- Чем плохи крайности/стереотипы в НРС

Масштабируемый вызов, или "Супервызов"

Минимизация правок - минимальное изменение семантики кода

Абстракция супервызова в целом не меняет языковой модели вычислений и выглядит для прикладного программиста, как новая возможность параллельного запуска на счет произвольного количества *заданий*. Каждое задание можно разделить на произвольное число фрагментов, которые смогут выполняться параллельно.

Например, вместо вызова подпрограммы

```
CALL sthd102n
```

можно написать

```
CALL scall(sthd102n,12)
```

Пример использования SCALL

```
        subroutine calc1(isc,nsc)
        [...]
c Instead of original cycle:
c     do i = 1, NM
c
c     ...
c     end do
c we do partial computation:
        integer ib(2)
        call ibounds(isc,nsc,ib, 1,NM)
        is = 0
        do i = ib(1),ib(2)
            do j = 1,NM
                do k = 1,NM
                    is = is + j*k
                    rm(k,j,i) = rm(k,j,i)+(135.*j)*(23.*k)
                end do
            end do
        end do
        call slock(lck)
        summa = summa + is
        call sunlock(lck)
        end

        [...]
        call scallw(calc1,0)
```

Распараллеливание примера

```
scall: +0 hot threads
      26460000
Saving SSTAT into /tmp/stest.0.sstat file
calc1_: {128,127,127,147,127} us,   30 times
```

```
scall: +1 hot threads
Core 1 activated...
      26460000
Saving SSTAT into /tmp/stest.0.sstat file
calc1_: {69,69,83,88,67} us,   30 times
```

Первое применение: модель воспроизводства научного потенциала России

Изначально модель создавалась в среде моделирования AnyLogic (язык программирования Java). Заказчику нужно было запустить ее на суперЭВМ, чтобы перейти от размеров Москвы к России (~140.000.000 агентов)

Был реализован набор микроправил, преобразующих используемые конструкции Java в C++, что дало ускорение примерно в 10 раз, а также существенную экономию оперативной памяти.

Подмножество используемого интерфейса среды AnyLogic было реализовано на базе API ADEVS, который был частично распараллелен на MPI (используемое подмножество функционала). После этого перенос на суперЭВМ был завершен (программа была успешно запущена на суперЭВМ "Чебышев" и "Ломоносов")

Трехмерный теплогидродинамический расчетный код улучшенной оценки "БАГИРА"

Страны-заказчики АЭС с ВВЭР (приоритет - безопасность!)

- Беларусь
- Болгария
- Бангладеш
- Индия
- Иран
- Китай
- Словакия
- Турция

Устройство АЭС с ВВЭР



Упрощенная схема устройства АЭС с ВВЭР

Моделирование АЭС с ВВЭР

- моделирование работы существующих и проектируемых перспективных АЭС;
- качественная тренировка и обучение персонала на имитационных моделях (симуляторах) АЭС в режиме реального времени;
- выработка рекомендаций по предотвращению и минимизации последствий всевозможных нештатных ситуаций;
- упреждающее моделирование для принятия решений по предотвращению и выходу из нештатных и чрезвычайных ситуаций.

Расчетный код улучшенной оценки БАГИРА

- смесь состоит из жидкой и газообразной (пар и неконденсирующийся газ) фаз;
- смесь неравновесна по температуре;
- смесь неравновесна по скорости;
- моделируются нестационарные, трёхмерные течения смеси;
- учитывается турбулентный тепло- и массоперенос.

Строятся неявные по давлению и объёмной скорости конечно-разностные уравнения, учитывается множество факторов (различные взаимодействия и граничные условия)

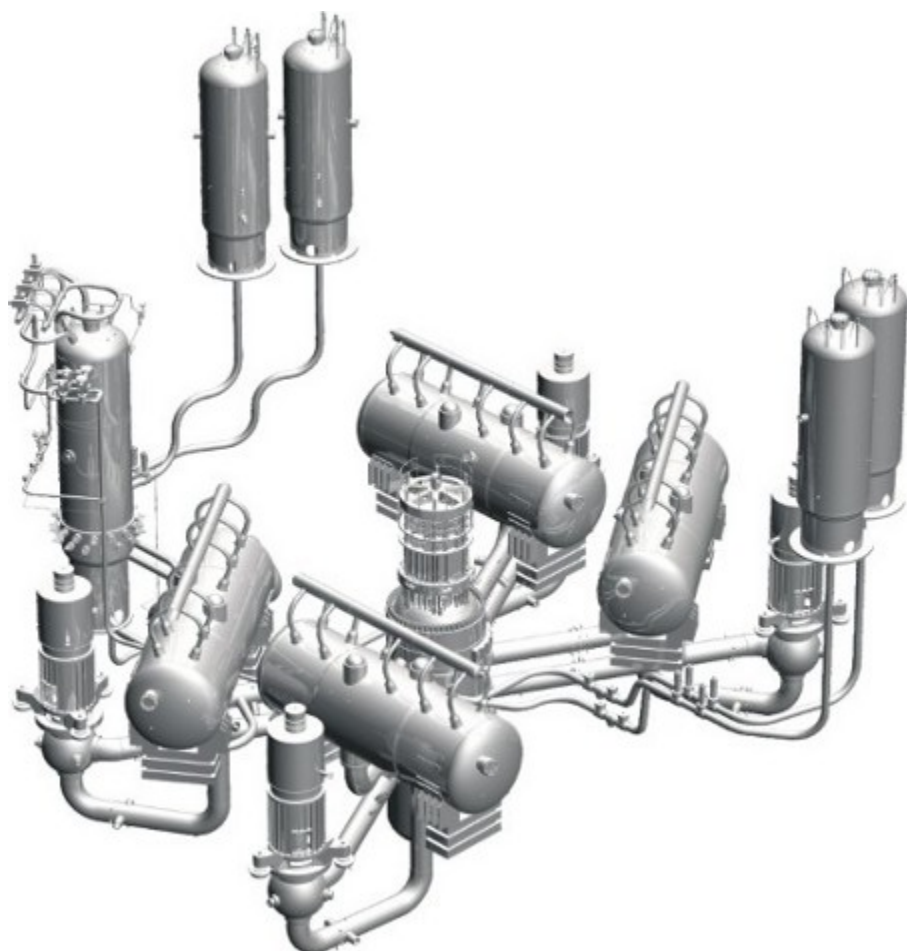
Счетное ядро содержит ~25.000 строк на языке Фортран.

Модель на 5000 ячеек



Парогенератор

Полномасштабная модель ~ 30000 ячеек



Способ распараллеливания

- Использование Супервызова
- Расщепление переменных (Common блоки!)
- Адаптеры для подключения внешних параллельных алгоритмов для СЛАУ
- Тестирование внешних параллельных алгоритмов на реальных данных
- Адаптивный алгоритм для решения СЛАУ с разреженной матрицей

Замеры производительности

Левая колонка - это засечки времени каждого 10-го ШМ (шага моделирования; можно считать что это показания часов в момент каждого 10-го вызова решателя СЛАУ), правая - время одного решения СЛАУ. Без оптимизации - это, по сути, оригинальный способ сборки.

Без оптимизации, оригинальный sthd082n:

1331369889226944 +	9933
1331369890086842 +	9935
1331369890896834 +	9962

=> ~81-85 миллисекунд/ШМ

Без оптимизации, UMFPACK:

1331369822588624 +	673
1331369823338717 +	673
1331369824076257 +	671

=> ~74 миллисекунды/ШМ

Текущие замеры производительности

С сильной оптимизацией, оригинальный sthd082n:

1331370069822617 +	4844
1331370069885706 +	4854
1331370069948735 +	4821

=> ~6.3 миллисекунды/ШМ

С сильной оптимизацией, UMFRACK:

1331370225989056 +	530
1331370226009451 +	530
1331370226029731 +	531

=> ~2 миллисекунды/ШМ

МАТРИЦА

- Теория: прямые и итерационные методы
- Практика: смешанные методы
- Неклассический подход: полупрямые методы

Библиотека SMatrix

- Имеет унифицированный интерфейс для различных библиотек
- Автоматически выбирает оптимальную библиотеку для матриц

Адаптивный алгоритм - скачок производительности

```
1331581231722264 +      80
Starting matrix inversion in background...
1331581231738731 +      86
Inversion done!!!
1331581231755454 +      62
1331581231771711 +      81
```

Итого ~1.6 миллисекунды на 1 ШМ (с распаралл. 10-го модуля)

Прямой метод UMFPACK требует ~480 микросекунд, PARDISO порядка ~250, но он капризен и дорог.

Платформы

- Большие суперЭВМ (МГУ "Ломоносов")
- Компактные суперЭВМ (РОСАТОМ)



КС-ЭВМ АПК-1 (ФГУП «РЯЦ ВНИИЭФ»)

- NVidia GPU
- Xeon Phi

Xeon Phi

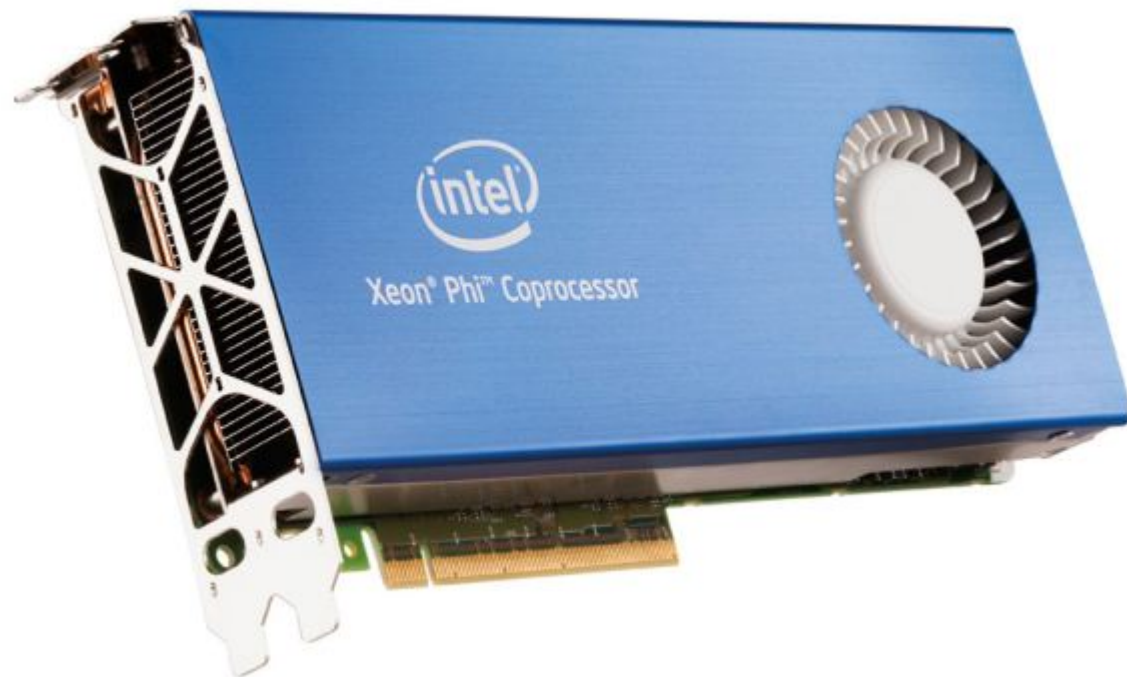
Со-процессор архитектуры Many Integrated Core, более 50 процессорных ядер, производительность *1 TFLOPs*.



Intel Knights Corner Chip

Последний Top500 и Xeon Phi

Top500 от 12.11.12: в семи суперкомпьютерах используется



60-ядер, 22-нм, 3-D Tri-Gate, 8 ГБ GDDR5, ОС Linux, 225 Вт, 1 терафлопс DP, \$2649

СПАСИБО ЗА ВНИМАНИЕ !

- V.A.Vasenin, M.A.Krivchikov, V.E.Kroshilin, A.E.Kroshilin, V.A.Roganov, "Scalable Three-Dimensional Thermal-Hydraulic Best-Estimate Code BAGIRA", Proceedings of ICAPP '12, June 24-28, Chicago, USA, 2012, paper №12375, pp. 2042-2050
- Васенин В. А., Кривчиков М. А., Крошилин А. Е., Крошилин В. Е., Рагулин А. Д., Роганов В. А. Распараллеливание расчетного кода улучшенной оценки "БАГИРА" для моделирования трехмерной теплогидродинамики многофазных сред в составе полномасштабной суперкомпьютерной модели "Виртуальная АЭС". "Программная Инженерия" №6 за 2012 год
- Макаров В.Л., Бахтизин А.Р., Васенин В.А., Роганов В.А., Трифонов И.А., Средства суперкомпьютерных систем для работы с агентно-ориентированными моделями. "Программная Инженерия", №3 за 2011 год
- В.А. Васенин, А.Е. Крошилин, В.Е. Крошилин, В.А. Роганов, М.А. Кривчиков, А.Д. Рагулин, РАСПАРАЛЛЕЛИВАНИЕ РАСЧЕТНОГО

КОДА ТРЕХМЕРНОЙ ТЕПЛОГИДРОДИНАМИКИ МНОГОФАЗНЫХ СРЕД В СОСТАВЕ ПОЛНОМАСШТАБНОЙ СУПЕРКОМПЬЮТЕРНОЙ МОДЕЛИ "ВИРТУАЛЬНАЯ АЭС", МГУ, Ломоносовские чтения, 2012

- Васенин В.А., Кривчиков М.А., Крошилин А.Е., Крошилин В.Е., Рагулин А.Д., Роганов В.А., "Улучшенное моделирование трехмерных теплогидродинамических процессов в узлах атомных электростанций на компактных суперкомпьютерных установках", XVII Байкальская Всероссийская конференция ИМТ-2012 "Информационные и математические технологии в науке и управлении", байкальская сессия, 3-9 июля 2012г