

Математическое моделирование столкновения галактик на гибридных суперЭВМ с ускорителями Intel Xeon Phi *

И.М. Куликов^{1,2,3}, И.Г. Черных^{1,2}, В.Е. Ненашев³, Е.В. Катыхшева³

Институт вычислительной математики и математической геофизики
Сибирского отделения РАН¹,
Новосибирский государственный университет²,
Новосибирский государственный технический университет³

В статье представлены результаты суперкомпьютерного моделирования процесса столкновения галактик на гибридном суперкомпьютере RSC PetaStream, оснащённом ускорителями Intel Xeon Phi. Для описания взаимодействующих галактик используется математическая модель, основанная на уравнениях гравитационной газовой динамики для описания газовой компоненты галактик и на уравнениях для первых моментов бесстолкновительного уравнения Больцмана для описания звездной компоненты и темной материи. Модель взаимодействующих галактик дополнена "подсеточными" физическими процессами (в зарубежной литературе используется термин subgrid-physics) – процессом звездообразования, эффектом от взрыва молодых сверхновых, самосогласованным образованием молекулярного водорода и процессом охлаждения/нагрева. Использование такой модели позволило сформулировать единый параллельный вычислительный метод для решения используемых гиперболических уравнений. Численная модель была реализована в виде программного комплекса для гибридных суперЭВМ, оснащённых ускорителями Intel Xeon Phi. В рамках одного ускорителя было достигнуто 134-кратное ускорение и 92-процентная эффективность при использовании 64 ускорителей Intel Xeon Phi.

1. Введение

Столкновение галактик играет ключевую роль в их эволюции [1]. Наблюдательное и теоретическое изучение взаимодействующих галактик – незаменимый метод исследования их свойств и эволюции. Математическое моделирование играет более чем важную роль в теоретическом исследовании таких процессов. В последнее время все более актуальным становится учет "подсеточных" физических процессов во взаимодействующих галактиках – процессов химокинетики [2], звездообразования [3], эффектов от взрыва сверхновых звезд [4], и связанных с ними охлаждения/нагрева, которые оказывают значительное влияние на динамику галактик.

Одной из главных проблем моделирования взаимодействующих галактик является соотношение масштабов. Так масса одной галактики составляет $10^{13} M_{\odot}$ и размер 10 килопарсек, а масса рядовой звезды в галактике порядка $M_{\odot}$ и размер 10^{-7} парсек, что приводит к разрыву в 13 порядков для массы и 14 порядков для размера между астрофизическими объектами. Данное обстоятельство приводит к необходимости использования мощнейших из доступных суперЭВМ. На сегодняшний день самые производительные суперкомпьютеры построены на гибридной архитектуре на основе графических ускорителей и ускорителей Intel Xeon Phi. Так в июльской версии 2015 года списка Top500 первые два суперкомпьютера (и четыре из Top10) построены на этих технологиях. Очевидно, что на основе именно гибридных архитектур будет построен первый эксафлопсный суперкомпьютер.

Особенности архитектуры, сетевой инфраструктуры, организации памяти и вычисле-

*Работа поддержана грантами Российского фонда фундаментальных исследований 15-31-20150 мол-авед, 15-01-00508 и 14-01-31199, грантом Президента РФ МК – 6648.2015.9.

ний делают разработку программного обеспечения для таких суперЭВМ достаточно сложной научной задачей, которая требует детальной проработки от физической постановки задачи до инструментов разработки. В этом случае мы говорим о концепции со-дизайна высокопроизводительных вычислений [5]. В работе описана математическая модель взаимодействующих галактик с учетом различных "подсеточных" физических процессов и численный метод, используемый для ее решения. Приведены схема и анализ производительности программной реализации, а также продемонстрированы результаты вычислительных экспериментов по моделированию столкновения галактик.

2. Математическая модель взаимодействующих галактик

В этом разделе мы сформулируем основные принципы со-дизайна, используемые при построении математической модели взаимодействующих галактик, а также опишем полученную математическую модель.

2.1. Со-дизайн вычислительной модели для описания взаимодействующих галактик

В работе [5] были сформулированы шесть основных стадий со-дизайна вычислительных моделей физических процессов для суперЭВМ: физическая модель → математическая формализация → численный метод → структуры данных → архитектура суперЭВМ → средства параллельного программирования. В самом общем смысле основная стратегия при построении вычислительной модели – это максимизация независимых вычислений. С учётом того, что уравнения гравитационной газовой динамики ранее нами успешно решались с использованием сеточных методов [1, 7], то мы и в дальнейшем будем использовать этот подход к моделированию газовой компоненты. Тогда стратегию построения вычислительной модели мы можем сформулировать более точно – максимизация независимых однородных вычислений в ячейках регулярной сетки, разбивающую расчетную область. Фактически при использовании регулярной (на самом деле и равномерной) расчетной сетки мы формулируем структуры данных. В этом случае мы решаем сразу две проблемы:

1. За счёт использования регулярной структуры данных (в нашем случае трехмерный массив) мы можем достаточно легко использовать любую вычислительную архитектуру суперЭВМ, которая основана на декартовой топологии. В основе графических ускорителей, ускорителей Intel Xeon Phi, большинства классических суперкомпьютеров используется именно такая топология.
2. В случае регулярности и однородности вычислений мы минимизируем число используемых программных средств, предоставляемых средствами параллельного программирования. В этом случае нам достаточно инструмента распределения задач (и при необходимости данных, как в случае графических ускорителей и ускорителей Intel Xeon Phi в режиме offload) между нитями одного процесса и простейшего инструмента обмена данными между процессами.

Однако, в этом случае появляется нетривиальная задача построения численного метода, который не будет давать артефактов, связанных с осями координат, а также задача построения математической модели для описания бесстолкновительной компоненты.

Основными ингредиентами галактик являются столкновительная компонента, моделирующая газовую составляющую галактик и равномерно распределенную в ней пыль, и бесстолкновительная компонента, моделирующая звездную компоненту и темную материю. Для описания бесстолкновительной компоненты будем использовать уравнения для первых моментов бесстолкновительного уравнения Больцмана. Такой подход был успешно опробован на моделировании эволюции [6] и столкновении галактик [7], а также он поз-

воляет термодинамически согласованно описать процессы фазового перехода между компонентами, который происходит при звездообразовании и взрыве сверхновых. В связи с этим, модель взаимодействующих галактик формулируется в виде системы гиперболических уравнений как для описания газовой компоненты, так и для описания бесстолкновительной компоненты. Такая формулировка позволяет нам использовать согласованный подход к фазовому переходу, необходимого для моделирования процесса звездообразования и взрыва сверхновых звезд, единый численный метод высокого порядка точности, способ задания структур данных и параллельного алгоритма. Для учета химических реакций будем рассматривать уравнения многокомпонентной односкоростной газовой динамики. С учётом того, что в нашей модели мы используем только процесс образования/распада молекулярного водорода, то мы можем построить аналитическое решения для изменения концентрации водорода в каждой ячейке расчетной области.

2.2. Уравнения многокомпонентной многофазной гидродинамики

Для описания газовой компоненты будем использовать систему уравнений односкоростной компонентной гравитационной газовой динамики, записанную в эйлеровых координатах:

$$\begin{aligned}\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) &= \mathcal{S} - \mathcal{D}, \\ \frac{\partial \rho_H}{\partial t} + \nabla \cdot (\rho_H \vec{u}) &= -s_{H,H_2} + \mathcal{S} \frac{\rho_H}{\rho} - \mathcal{D}^H \frac{\rho_H}{\rho}, \\ \frac{\partial \rho_{H_2}}{\partial t} + \nabla \cdot (\rho_{H_2} \vec{u}) &= s_{H,H_2} + \mathcal{S} \frac{\rho_{H_2}}{\rho} - \mathcal{D}^{\rho_{H_2}} \frac{\rho_{H_2}}{\rho}, \\ \frac{\partial \rho \vec{u}}{\partial t} + \nabla \cdot (\rho \vec{u} \vec{u}) &= -\nabla p - \rho \nabla(\Phi) + \vec{v} \mathcal{S} - \vec{u} \mathcal{D}, \\ \frac{\partial \rho E}{\partial t} + \nabla \cdot (\rho E \vec{u}) &= -\nabla \cdot (p \vec{v}) - (\rho \nabla(\Phi), \vec{u}) - \Lambda + \Gamma - \varepsilon \frac{\mathcal{D}}{\rho}, \\ \frac{\partial \rho \varepsilon}{\partial t} + \nabla \cdot (\rho \varepsilon \vec{u}) &= -(\gamma - 1) \rho \varepsilon \nabla \cdot \vec{u} - \Lambda + \Gamma - \varepsilon \frac{\mathcal{D}}{\rho}, \\ \rho E &= \frac{1}{2} \rho \vec{u}^2 + \rho \varepsilon, \\ p &= (\gamma - 1) \rho \varepsilon,\end{aligned}$$

Для описания бесстолкновительной компоненты будем использовать систему уравнений для первых моментов бесстолкновительного уравнения Больцмана, записанную также в эйлеровых координатах:

$$\begin{aligned}\frac{\partial n}{\partial t} + \nabla \cdot (n \vec{v}) &= \mathcal{D} - \mathcal{S}, \\ \frac{\partial n \vec{v}}{\partial t} + \nabla \cdot (n \vec{v} \vec{v}) &= -\nabla \Pi - n \nabla(\Phi) + \vec{u} \mathcal{D} - \vec{v} \mathcal{S}, \\ \frac{\partial \rho W}{\partial t} + \nabla \cdot (\rho W \vec{v}) &= -\nabla \cdot (\Pi \vec{v}) - (n \nabla(\Phi), \vec{v}) - \Gamma + \varepsilon \frac{\mathcal{D}}{\rho}, \\ \frac{\partial \Pi_{\xi\xi}}{\partial t} + \nabla \cdot (\Pi_{\xi\xi} \vec{v}) &= -2 \Pi \nabla \cdot \vec{u} - \Gamma + \varepsilon \frac{\mathcal{D}}{3\rho}, \\ \rho W &= \frac{1}{2} \rho \vec{v}^2 + \frac{\Pi_{xx} + \Pi_{yy} + \Pi_{zz}}{2},\end{aligned}$$

Уравнение Пуассона для обеих компонент записывается в виде:

$$\Delta \Phi = 4\pi G(\rho + n),$$

где p – давление газа, ρ_H – плотность атомарного водорода, ρ_{H_2} – плотность молекулярного водорода, s_{H,H_2} – скорость образования молекулярного водорода из атомарного, $\rho = \rho_H + \rho_{H_2}$ – плотность смеси газа, n – плотность бесстолкновительной компоненты, $\vec{u}$ – скорость газовой компоненты, $\vec{v}$ – скорость бесстолкновительной компоненты, ρE – плотность полной механической энергии газа, ρW – плотность полной механической энергии бесстолкновительной компоненты, Φ – гравитационный потенциал, ε – плотность внутренней энергии газа, γ – эффективный показатель адиабаты, $\Pi_{\xi\xi} = (\Pi_{xx}, \Pi_{yy}, \Pi_{zz})$ – диагональный тензор дисперсии скоростей бесстолкновительной компоненты, $\mathcal{S}$ – скорость образования сверхновых звезд, $\mathcal{D}$ – скорость звездообразования, Λ – функция охлаждения, Γ – функция нагревания от взрыва сверхновых звезд.

2.3. Модель ”подсеточной” физики

Мы остановимся на следующих, на наш взгляд, основных ”подсеточных” процессах, играющих основную роль при столкновении галактик.

2.3.1. Образование молекулярного водорода

Молекулы водорода в межгалактическом пространстве формируются на поверхности частиц и диссоциируют космическим излучением. Предполагая, что плотность газа пропорциональна плотности частиц из-за хорошего перемешивания частиц и газа в нашей модели, концентрация молекулярного водорода определяется следующим выражением [8]:

$$\frac{dn_{H_2}}{dt} = R_{gr}(T)n_H(n_H + 2n_{H_2}) - (\xi_H + \xi_{diss}(N_{H_2}, A_V))n_{H_2},$$

где n_{H_2} и n_H – концентрация молекулярного и атомарного водорода, N_{H_2} – столбцевая плотность молекулярного водорода.

2.3.2. Процесс звездообразования и образование сверхновых звезд

Необходимое условие процесса звездообразования сформулируем в виде [9]:

$$T < 10^4 K \quad \nabla \cdot \vec{u} < 0 \quad \rho > 1.64 \frac{M_\odot}{pc^{-3}}$$

тогда скорость звездообразования может быть записана в виде:

$$\mathcal{D} = \frac{dn}{dt} = \mathcal{C} \frac{\rho}{\tau_{dyn}} = \mathcal{C} \rho^{3/2} \sqrt{\frac{32G}{3\pi}}$$

где $\mathcal{C} = 0.034$ – коэффициент эффективности звездообразования. Скорость образования сверхновых звезд записывается в виде [10]:

$$\mathcal{S} = \Delta^{SN} = \frac{d\rho}{dt} = \beta \mathcal{C} \frac{n}{\tau_{dyn}} = \beta \mathcal{C} n^{3/2} \sqrt{\frac{32G}{3\pi}}$$

где $\beta = 0.1$ – коэффициент взрыва молодых звезд. При взрыве одной звезды солнечной массы выделяется энергия 10^{51} эрг, в этом случае функция нагревания за счет взрыва сверхновых может быть записана в виде:

$$\Gamma = 10^{51} \frac{M^{SN}}{M_\odot} erg$$

где M^{SN} – масса сверхновых звезд в локальном объеме.

2.3.3. Функция охлаждения

Галактический газ, разогретый за счет столкновения до температуры $\sim 10^4 - 10^8 K$ описывается функцией охлаждения [11]:

$$\Lambda \simeq 10^{-22} n_H^2 \text{cm}^{-3} \text{erg}$$

где n_H – концентрация атомарного водорода.

3. Численный метод решения

Для решения гидродинамических уравнений был использован оригинальный численный метод, основанный на комбинации operator splitting подхода, метода Годунова, схемы Рое и кусочно-параболической реконструкции на локальном шаблоне. Такой метод объединяет все достоинства перечисленных методов и обладает высокой степенью параллелизации, на чем остановимся подробнее в следующем разделе. Для решения уравнения Пуассона используется метод, основанный на быстром преобразовании Фурье.

3.1. Эйлеров этап решения гидродинамических уравнений

На эйлеровом этапе численного метода решаются уравнения без учета адвективных членов и функций, моделирующих "подсеточные" процессы. Для аппроксимации пространственных производных используется решение задачи линеаризованной задачи Римана. Для этого на всех интерфейсах между ячейками (L – обозначение левой ячейки, R – обозначение правой ячейки) происходит осреднение скорости и давления по формулам:

$$\begin{aligned} \rho &= \frac{\rho_L^{3/2} + \rho_R^{3/2}}{\sqrt{\rho_L} + \sqrt{\rho_R}} & n &= \frac{n_L^{3/2} + n_R^{3/2}}{\sqrt{n_L} + \sqrt{n_R}} \\ p &= \frac{p_L \sqrt{\rho_L} + p_R \sqrt{\rho_R}}{\sqrt{\rho_L} + \sqrt{\rho_R}} & \Pi &= \frac{\Pi_L \sqrt{n_L} + \Pi_R \sqrt{n_R}}{\sqrt{n_L} + \sqrt{n_R}} \end{aligned}$$

Такой способ осреднения связан с необходимостью более аккуратного учета границы газ-вакуум, чем как это сделано в оригинальной схеме Рое [12]. При этом будем предполагать, что в рассматриваемых ячейках решение является кусочно-параболической функцией. Процедура построения локальных парабол подробно описана в классической работе [13]. Эта процедура нам понадобится также в следующем этапе метода. Опустив подробные выкладки для вывода схемы с помощью метода Годунова (численный метод и его верификация подробно описаны в работе [14]) решение задачи Римана для эйлерова этапа решения уравнений для газовой компоненты записываются в виде:

$$\begin{aligned} U &= \frac{u_L(-\lambda t) + u_R(\lambda t)}{2} + \frac{p_L(-\lambda t) - p_R(\lambda t)}{2} \sqrt{\frac{(\sqrt{\rho_L} + \sqrt{\rho_R})^2}{\frac{\gamma_L + \gamma_R}{2} (\rho_L^{3/2} + \rho_R^{3/2}) (p_L \sqrt{\rho_L} + p_R \sqrt{\rho_R})}} \\ P &= \frac{p_L(-\lambda t) + p_R(\lambda t)}{2} + \frac{u_L(-\lambda t) - u_R(\lambda t)}{2} \sqrt{\frac{\frac{\gamma_L + \gamma_R}{2} (\rho_L^{3/2} + \rho_R^{3/2}) (p_L \sqrt{\rho_L} + p_R \sqrt{\rho_R})}{(\sqrt{\rho_L} + \sqrt{\rho_R})^2}}, \end{aligned}$$

для бесстолкновительной компоненты суть:

$$\begin{aligned} V &= \frac{v_L(-\mu t) + v_R(\mu t)}{2} + \frac{\Pi_L(-\mu t) - \Pi_R(\mu t)}{2} \sqrt{\frac{(\sqrt{n_L} + \sqrt{n_R})^2}{3(n_L^{3/2} + n_R^{3/2})(\Pi_L \sqrt{n_L} + \Pi_R \sqrt{n_R})}} \\ \Pi &= \frac{\Pi_L(-\mu t) + \Pi_R(\mu t)}{2} + \frac{v_L(-\mu t) - v_R(\mu t)}{2} \sqrt{\frac{3(n_L^{3/2} + n_R^{3/2})(\Pi_L \sqrt{n_L} + \Pi_R \sqrt{n_R})}{(\sqrt{n_L} + \sqrt{n_R})^2}} \end{aligned}$$

где

$$\lambda = \sqrt{\frac{\frac{\gamma_L + \gamma_R}{2}(p_L \sqrt{\rho_L} + p_R \sqrt{\rho_R})}{\rho_L^{3/2} + \rho_R^{3/2}}} \quad \mu = \sqrt{\frac{3(\Pi_L \sqrt{n_L} + \Pi_R \sqrt{n_R})}{n_L^{3/2} + n_R^{3/2}}}$$

$$q_L(-\nu t) = q_i^R - \frac{\nu t}{2h} \left(\Delta q_i - q_i^6 \left(1 - \frac{2\nu t}{3h} \right) \right) \quad q_R(\nu t) = q_i^L + \frac{\nu t}{2h} \left(\Delta q_i + q_i^6 \left(1 - \frac{2\nu t}{3h} \right) \right)$$

Процедура построения параболы и параметров $q_i^R, q_i^L, \Delta q_i, q_i^6$ подробно приведена в работе [13], а также в работах [15, 16], где кусочно-параболический метод на локальном шаблоне был использован для полной системы уравнений газовой динамики и магнитной газовой динамики.

3.2. Лагранжев этап решения гидродинамических уравнений

На лагранжевом этапе происходит адвективный перенос гидродинамических параметров и все уравнения на лагранжевом этапе имеют вид:

$$\frac{\partial f}{\partial t} + \nabla \cdot (f \vec{v}) = 0$$

где f может быть плотностью ρ, n , импульсом $\rho \vec{u}, n \vec{v}$, плотностью общей механической $\rho E, nW$ или внутренней $\rho \epsilon, \Pi_{\xi\xi}$ энергий газа или бесстолкновительной компоненты. Для решения уравнений используется аналогичный, что и на эйлеровом этапе подход. Для вычисления потока $F = f \vec{v}$ при $\lambda = |\vec{v}|$ используется формула:

$$F = v \times \begin{cases} f_L(-\lambda t), v \geq 0 \\ f_R(\lambda t), v < 0 \end{cases}$$

где $f_L(-\lambda t)$ и $f_R(\lambda t)$ – кусочно-параболические функции для величины f и скорость на интерфейсе между ячейками вычисляется по усреднению Рое:

$$v = \frac{v_L \sqrt{\rho_L} + v_R \sqrt{\rho_R}}{\sqrt{\rho_L} + \sqrt{\rho_R}}$$

Для построения кусочно-параболического решения используется аналогичная эйлеровому этапу процедура. На заключительном этапе происходит учет "подсеточных" процессов с помощью метода Эйлера.

3.3. Использование переопределенной системы уравнений

На завершающем этапе решения гидродинамических уравнений происходит корректировка решения. В случае границы газ-вакуум с использованием формулы [17]:

$$|\vec{v}| = \sqrt{2(E - \epsilon)}, (E - \vec{v}^2/2)/E \geq 10^{-3}$$

в остальной области происходит корректировка, которая гарантирует неубывание энтропии [18]:

$$|\rho \epsilon| = \left(\rho E - \frac{\rho \vec{v}^2}{2} \right), (E - \vec{v}^2/2)/E < 10^{-3}$$

Такая модификация обеспечивает детальный баланс энергий и гарантирует неубывание энтропии.

3.4. Решение уравнения Пуассона

После решения гидродинамических уравнений необходимо восстановить гравитационный потенциал по плотностям газовой и бесстолкновительной компонент. Для этого мы будем использовать 27-точечный шаблон для аппроксимации уравнения Пуассона. Это связано с обеспечением максимальной инвариантности решения относительно поворота. Алгоритм решения уравнения Пуассона будет состоять из нескольких этапов.

Этап 1. Постановка граничных условий для уравнения Пуассона. Для постановки граничных условий для гравитационного потенциала на границе области D будем использовать первые члены мультипольного разложения – статический, осевой и центробежный моменты инерции. Когда значения потенциала на границе области определены, то они подставляются в 27-точечный шаблон для определения итоговой плотности $\rho_{i,k,l} + n_{i,k,l}$ на границе D .

Этап 2. Преобразование плотности в пространстве гармоник. Итоговая плотность представляется в виде суперпозиции по собственным функциям оператора Лапласа:

$$\rho_{i,k,l} + n_{i,k,l} = \sum_{jmn} \sigma_{jmn} \exp \left(\frac{\vec{i}\pi j}{I} + \frac{\vec{i}\pi km}{K} + \frac{\vec{i}\pi ln}{L} \right)$$

где I, K, L – число ячеек по каждой координате, $\vec{i}$ – мнимая единица. Для этого используется быстрое преобразование Фурье.

Этап 3. Решение уравнения Пуассона в пространстве гармоник. Мы предполагаем, что потенциал также представлен в виде суперпозиции по собственным функциям оператора Лапласа и может быть вычислен по следующей формуле из амплитуд гармоник плотности:

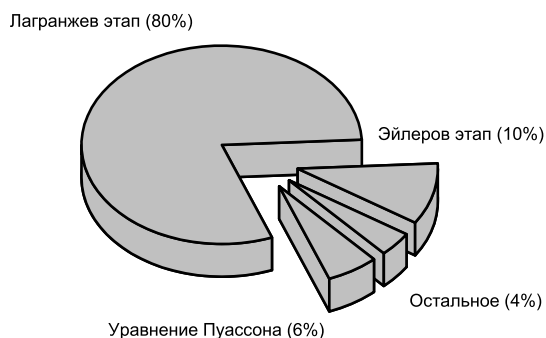
$$\phi_{jmn} = \frac{\frac{2}{3}\pi h^2 \sigma_{jmn}}{1 - \left(1 - \frac{2\sin^2 \frac{\pi j}{I}}{3}\right) \left(1 - \frac{2\sin^2 \frac{\pi m}{K}}{3}\right) \left(1 - \frac{2\sin^2 \frac{\pi n}{L}}{3}\right)}$$

После чего необходимо проделать обратное быстрое преобразование Фурье гармоник потенциала в функциональное пространство гармоник.

4. Описание параллельной реализации

Со-дизайн [5] физико-математической модели, численного метода и структур данных позволяет использовать геометрическую декомпозицию расчетной области с одним слоем перекрытия подобластей. Такую возможность мы имеем за счёт построения парабол на предыдущем шаге, что требует только локального взаимодействия между ячейками. На рисунке 1 приведены процентные соотношения между этапами. Для решения уравнения Пуассона, в основе которого быстрое преобразование Фурье для суперЭВМ с распределенной памятью, была использована библиотека FFTW [19]. В основе этой библиотеки лежит процедура ALLTOALL, которая "транспонирует" трехмерный массив, перераспределяя значительные объемы памяти между всеми процессами. Безусловно, это дорогая сетевая операция, которая требует отказа от всего алгоритма в случае использования сколь либо значительного количества вычислителей. Однако, эта процедура в случае использования сетевой инфраструктуры InfiniBand не занимает критическое время и, по всей видимости, оптимизирована на низком сетевом уровне [20].

Основными этапами вычислительной схемы являются эйлеров и лагранжевы этапы. Мы сосредоточимся именно на этих этапах, как на наиболее затратных. Отдельно остановимся на процедуре вычисления шага по времени, исходя из условия Куранта. В случае использования графических ускорителей данная процедура была реализована только на CPU [7] (также было сделано и в коде GAMER [21]). Причина этого – отсутствие эффективной низкоуровневой реализации редуцирующей операцией *min* в технологии CUDA. В то время как



(а)

Рис. 1. Процентное соотношение между этапами численного метода

в OpenMP такая операция эффективно реализована. Стоимость этой процедуры составляет порядка одного процента от общего времени вычислений и практически не влияет на эффективность параллельной реализации. Однако, при увеличении количества графических ядер до нескольких тысяч и стократного ускорения в рамках одного графического процессора суммарно всех остальных процедур, может возникнуть курьёзная ситуация, когда процедура вычисления шага по времени будет выполняться дольше всех остальных. При том, что авторами уже было достигнуто 55-кратное ускорение в рамках одного GPU [7] и количество графических ядер в одном ускорителе увеличивается, то такая ситуация может быть достигнута в ближайшие пару лет. Стоит отметить, что такая проблема в принципе невозможна на ускорителях Intel Xeon Phi.

Использование равномерной сетки в декартовых координатах для решения уравнений гидродинамики позволяет использовать произвольную декартову топологию для декомпозиции расчетной области. Такая организация вычислений имеет потенциально бесконечную масштабируемость. В нашей программной реализации используется многоуровневая одномерная декомпозиция расчетной области. По одной координате внешнее одномерное разрезание происходит средствами технологии MPI, внутри каждой подобласти разрезание происходит средствами OpenMP, адаптированного для MIC-архитектур (рис. 2). Такой подход использовался также и в первой версии программного кода AstroPhi [22] с учетом использования offload режима. Такая декомпозиция связана с топологией и архитектурой

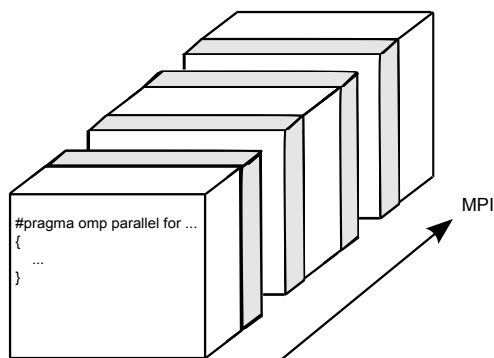


Рис. 2. Декомпозиция расчетной области. Серым цветом обозначены слои перекрытия подобластей

гибридного суперЭВМ RSC PetaStream, который был использован для вычислительных экспериментов. Далее приведем исследования производительности программной реализации.

4.1. Исследование ускорения кода в рамках одного MIC

Проводилось исследование ускорения программной реализации на сетке 512^3 . Для этого замерялось время каждого этапа (Euler, Lagrange) численного метода, в секундах, а затем вычислялась их сумма при различном числе используемых логических ядер. Ускорение P вычислялось по формуле 1:

$$P = \frac{Total_1}{Total_K} \quad (1)$$

где $Total_1$ – время вычислений на одном логическом ядре, $Total_K$ – время вычислений при использовании K логических ядер.

Также была сделана оценка реальной производительности (S_{real}). Для оценки пиковой производительности S_{peak} ускорителя Intel Xeon Phi была использована формула 2:

$$S_{peak} = Units \times Frequency \times Cores = 74,28 GFLOPS \quad (2)$$

где $Units = 1 FLOP$ – число скалярных устройств, $Frequency = 1,238 GHz$ – тактовая частота, $Cores = 60$ – число физических ядер. Для нахождения эффективности использования ускорителя S_{ratio} будем использовать формулу 3

$$S_{ratio} = \frac{S_{real}}{S_{peak}} \quad (3)$$

Эта характеристика производительности с одной стороны несколько странная, так как пиковый уровень производительности Intel Xeon Phi это величина порядка одного терафлопса. Однако, такая производительность в основном достигается векторным расширением, которые мы пока не используем. Результаты исследования ускорения на сетке 512^3 приведены на рисунке 3. Таким образом, было получено 134-кратное ускорение (масштабируемость в сильном смысле) в рамках одного ускорителя Intel Xeon Phi. Количество гигафлопсов соотносятся с результатами, полученными в книге [23]. А результат порядка чуть менее 29 GFLOPS это примерно 40 % от пиковой, что достаточно высокий результат.

4.2. Исследование масштабируемости кода

Проводилось исследование масштабируемости кода AstroPhi на расчетной сетке $512p \times 512 \times 512$ при использовании 4 логических ядер на каждый ускоритель, где p – число используемых ускорителей. Таким образом, на каждый ускоритель приходится размер подоласти 512^3 . Для исследования масштабируемости замерялось время каждого этапа (Euler, Lagrange) численного метода, в секундах, а затем вычислялась их сумма при различном

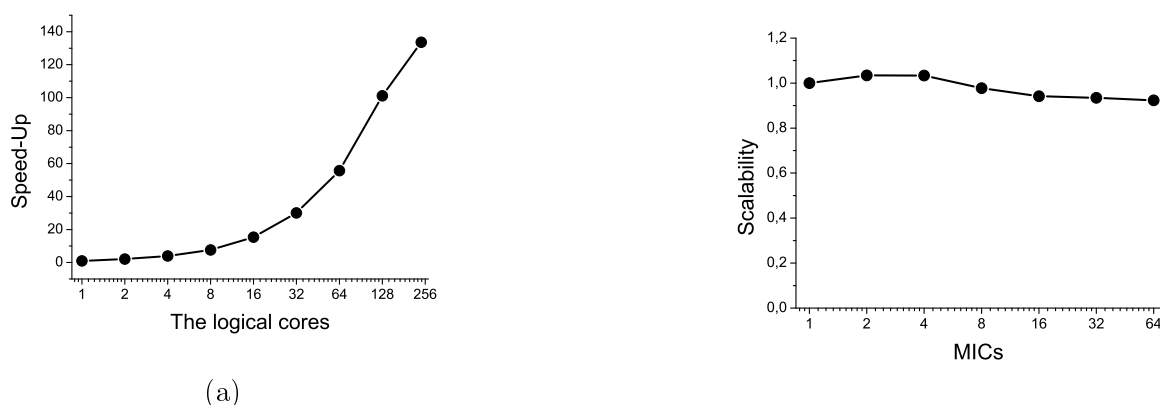


Рис. 3. Ускорение (a) и масштабируемость (b) программной реализации

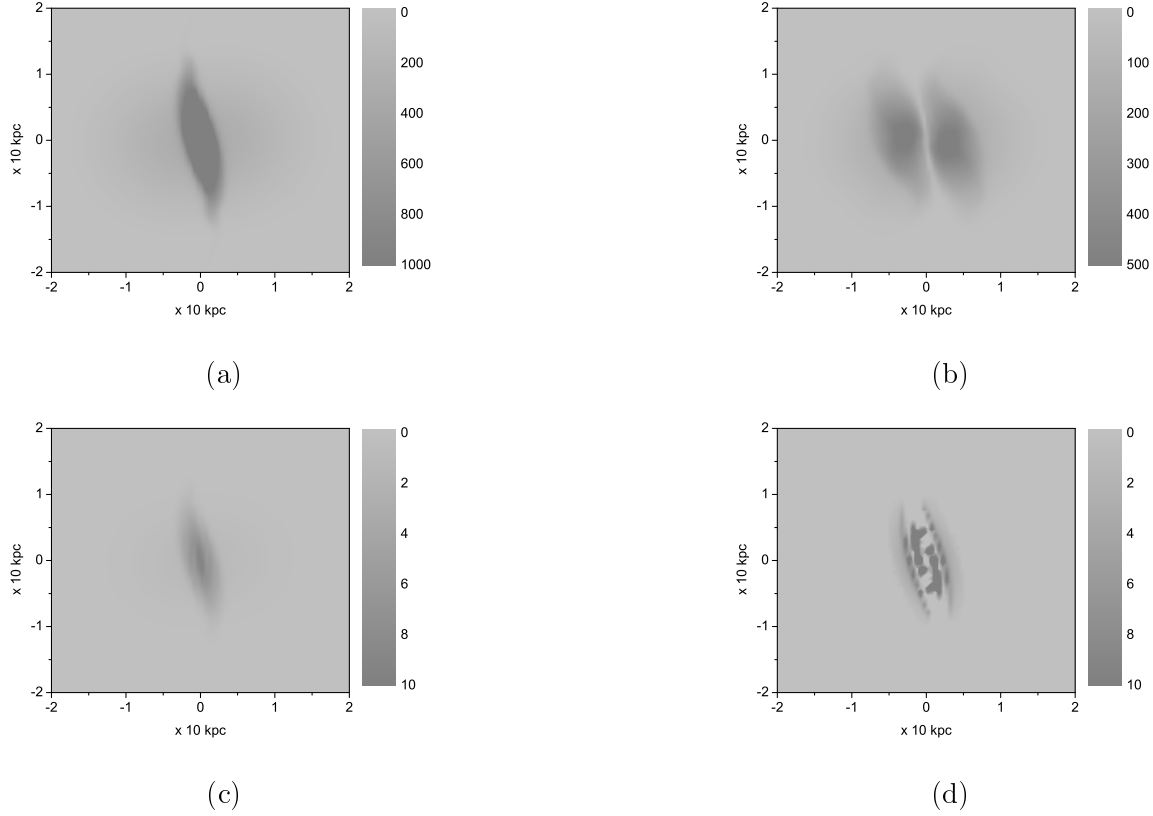


Рис. 4. Результаты моделирования на момент времени $t = 120$ миллионов лет. Столбцевая плотность в $M_{\odot}pc^{-2}$ газовой – (а) и бесстолкновительной – (b) компонент, молекулярного водорода – (с). Средняя скорость процесса звездообразования в $M_{\odot}pc^{-2}myr^{-1}$ – (d)

числе используемых ускорителей Intel Xeon Phi. Масштабируемость T вычислялось по формуле

$$T = \frac{Total_1}{Total_p} \quad (4)$$

где $Total_1$ – время вычислений на одном ускорителе при использовании одного ускорителя, $Total_p$ – время вычислений на одном ускорителе при использовании p ускорителей. Результаты исследований масштабируемости приведены на рисунке 3. Таким образом, при использовании всех 64 ускорителей была получена 92 % эффективность, что является достаточно высоким результатом.

5. Результаты вычислительных экспериментов

Будем моделировать столкновение двух галактик с массой $M = 10^{13}M_{\odot}$,двигающихся навстречу друг другу со скоростью $v_{cr} = 800$ км/сек, каждая из которых задана двумя самогравитирующими сферическими облаками для описания газовой и бесстолкновительной компоненты с равновесным начальным распределением плотности, давления/тензора дисперсии скоростей. Галактики вращаются в противоположные стороны с дифференциальным вращением:

$$v_{\phi} = \sqrt{r \frac{\partial \Phi}{\partial r}}$$

Начальное расстояние между центрами $30kpc$. На рисунке 4 представлены результаты распределения столбцевой плотности газовой и звездной компоненты. После процесса столкновения за фронтом ударной волны происходит активный рост скорости звездообразования,

а также на месте будущей галактики образуется молекулярный водород.

6. Заключение

Проведены вычислительные эксперименты по исследованию производительности программной реализации на кластере RSC PetaStream. Было получено 134-кратное ускорение (масштабируемость в сильном смысле) в рамках одного ускорителя Intel Xeon Phi, 92-процентная эффективность (масштабируемость в слабом смысле) при использовании 64 ускорителей Intel Xeon Phi, и 40-процентная производительность от пиковой при использовании скалярных вычислений на ускорителе Intel Xeon Phi. Приведены результаты суперкомпьютерных вычислительных экспериментов по столкновению дисковых галактик. Смоделированы области с активным звездообразованием и образованием молекулярного водорода, которые получаются в результате столкновения галактик.

Авторы выражают благодарность коллегам из организаций Intel (Николаю Местеру и Дмитрию Петунину) и RSC Group (Александру Московскому, Алексею Шмелеву, Владимиру Миронову, Павлу Лавренко и Борису Гагаринову), за предоставление доступа к кластеру RSC PetaStream и подробные консультации по его использованию.

Литература

1. Tutukov A., Lazareva G., Kulikov I. Gas Dynamics of a Central Collision of Two Galaxies: Merger, Disruption, Passage, and the Formation of a New Galaxy // *Astronomy Reports*. – 2011. – V. 55, I. 9. – P. 770-783.
2. Combes F., Melchior A. Chemodynamical evolution of interacting galaxies // *Astrophysics and Space Science*. – 2002. – V. 281, I. 1-2. – P. 383-387.
3. Schweizer F. Merger-Induced Starbursts // *Astrophysics and Space Science Library*. – 2005. – V. 329. – P. 143-152.
4. Sol Alonso M., Lambas D., Tissera P., Coldwell G. Active galactic nuclei and galaxy interactions // *Monthly Notices of the Royal Astronomical Society*. – 2007. – V. 375, I. 3. – P. 1017-1024.
5. Glinskiy B., Kulikov I., Snytnikov A., Romanenko A., Chernykh I., Vshivkov V. Co-design of Parallel Numerical Methods for Plasma Physics and Astrophysics // *Supercomputing frontiers and innovations*. – 2014. – V. 1, I. 3. – P. 88-98.
6. Mitchell N., Vorobyov E., Hensler G. Collisionless Stellar Hydrodynamics as an Efficient Alternative to N-body Methods // *Monthly Notices of the Royal Astronomical Society*. – 2013. – V. 428. – P. 2674-2687.
7. Kulikov I. GPUPEGAS: A New GPU-accelerated Hydrodynamic Code for Numerical Simulations of Interacting Galaxies // *The Astrophysical Journal Supplements Series*. – 2014. – V. 214, I. 12. – P. 1-12.
8. Bergin E., Hartmann L., Raymond J., Ballesteros-Paredes J. Molecular Cloud Formation behind Shock Waves // *The Astrophysical Journal*. – 2004. – V. 612. – P. 921-939.
9. Katz N., Weinberg D., Hernquist L. Cosmological simulations with TreeSPH // *The Astrophysical Journal Supplement Series*. – 1996. – V. 105. – P. 19-35.
10. Springel V., Hernquist L. Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation // *Monthly Notices of the Royal Astronomical Society*. – 2003. – V. 339, I. 2. – P. 289-311.

11. Sutherland R., Dopita M. Cooling functions for low-density astrophysical plasmas // The Astrophysical Journal Supplement Series. – 1993. – V. 88. – P. 253-327.
12. Roe P. Approximate Riemann solvers, parameter vectors, and difference solvers // Journal of Computational Physics. – 1997. – V. 135. – P. 250-258.
13. Collela P., Woodward P. The Piecewise Parabolic Method (PPM) Gas-Dynamical simulations // Journal of Computational Physics. – 1984. – V. 54. – P. 174-201.
14. Kulikov I., Vorobyov E. Using the PPML approach for constructing a low-dissipation, operator-splitting scheme for numerical simulations of hydrodynamic flows // New Astronomy. – 2015. (in progress)
15. Popov M., Ustyugov S. Piecewise parabolic method on local stencil for gasdynamic simulations // Computational Mathematics and Mathematical Physics. – 2007. – V. 47, I. 12. – P. 1970-1989.
16. Popov M., Ustyugov S. Piecewise parabolic method on a local stencil for ideal magnetohydrodynamics // Computational Mathematics and Mathematical Physics. – 2008. – V. 48, I. 3. – P. 477-499.
17. Vshivkov V., Lazareva G., Snytnikov A., Kulikov I., Tutukov A. Computational methods for ill-posed problems of gravitational gasodynamics // Journal of Inverse and Ill-posed Problems. – 2011. – V. 19, I. 1. – P. 151-166
18. Godunov S., Kulikov I. Computation of Discontinuous Solutions of Fluid Dynamics Equations with Entropy Nondecrease Guarantee // Computational Mathematics and Mathematical Physics. – 2014. – V. 54, I. 6. – P. 1012-1024.
19. Frigo M., Johnson S. The Design and Implementation of FFTW3 // Proceedings of the IEEE. – 2005. – V. 93, I. 2. – P. 216-231.
20. Kalinkin A., Laevsky Y., Gololobov S. 2D Fast Poisson Solver for High-Performance Computing // Lecture Notes in Computer Science. – 2009. – V. 5698. – P. 112-120.
21. Schive H., Tsai Y., Chiueh T. GAMER: a GPU-accelerated Adaptive-Mesh-Refinement Code for Astrophysics // The Astrophysical Journal. – 2010. – V. 186. – P. 457-484.
22. Kulikov I.M., Chernykh I.G., Snytnikov A.V., Glinskiy B.M., Tutukov A.V. AstroPhi: A code for complex simulation of dynamics of astrophysical objects using hybrid supercomputers // Computer Physics Communications. – 2015. – V. 186. P. 71-80.
23. Jeffers J., Reinders J. Intel Xeon Phi Coprocessor High Performance Programming. Elsevier. – 2013. 409 p.