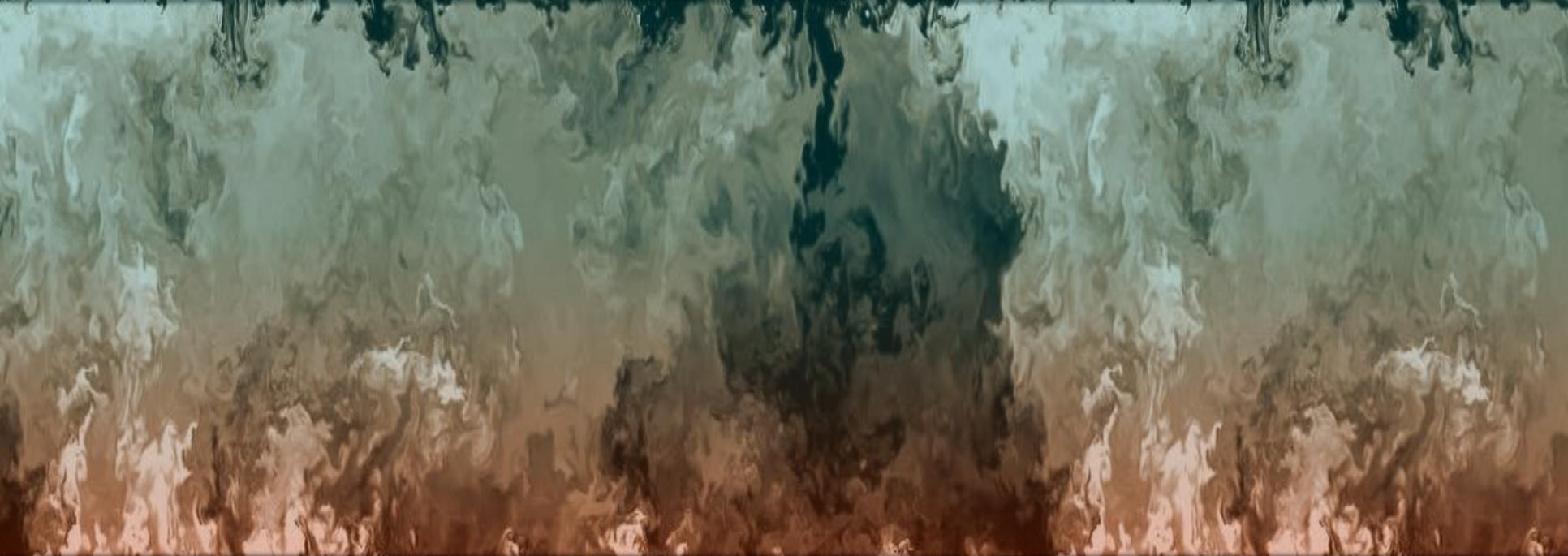


Суперкомпьютерное моделирование турбулентного течения несжимаемой жидкости при естественной и вынужденной конвекции



Сектор вычислительной аэроакустики
<http://caa.imamod.ru>
ИГМ им. М. В. Келдыша РАН

Программный комплекс STG-CFD&HT

- Система уравнений Навье-Стокса:

$$\begin{aligned}\nabla \cdot \mathbf{u} &= 0, \\ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} &= \frac{\text{Pr}}{\sqrt{\text{Ra}}} \nabla^2 \mathbf{u} - \nabla p + \mathbf{f}, \\ \frac{\partial T}{\partial t} + (\mathbf{u} \cdot \nabla) T &= \frac{1}{\sqrt{\text{Ra}}} \nabla^2 T.\end{aligned}$$

- Дискретная система для скорости-давления:

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \frac{3}{2} \mathbf{R}^n - \frac{1}{2} \mathbf{R}^{n-1} - Gp^{n+1},$$

$$M\mathbf{u}^{n+1} = 0,$$

$$\text{где } \mathbf{R}(\mathbf{u}) = -C(\mathbf{u})\mathbf{u} - D\mathbf{u} + \mathbf{f}.$$

- Проекционный метод дробного шага:

$$\text{Предиктор: } \mathbf{u}^p = \mathbf{u}^n + \Delta t \left(\frac{3}{2} \mathbf{R}^n - \frac{1}{2} \mathbf{R}^{n-1} \right)$$

$$\text{Проекция: } \mathbf{u}^{n+1} = \mathbf{u}^p - G\tilde{p}, \text{ где } \tilde{p} = \Delta t p^{n+1}$$

$$\text{Уравнение неразрывности: } M\mathbf{u}^{n+1} = M\mathbf{u}^p - GM\tilde{p} = 0$$

$$M\mathbf{u}^{n+1} = M\mathbf{u}^p - GM\tilde{p} = -M\Omega M^* \tilde{p} = \boxed{L\tilde{p} = M\mathbf{u}^p} \quad \text{Уравнение Пуассона}$$

Алгоритм шага по времени

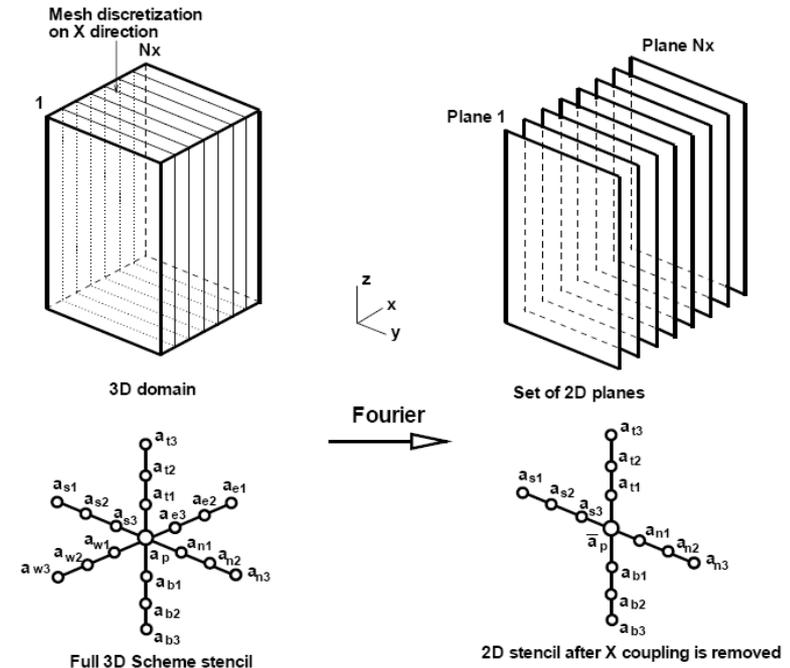
1. Предиктор поля скорости \mathbf{u}^p находится явно.
2. Коррекция \tilde{p} находится из уравнения Пуассона.
3. Вычисляется новое поле скорости \mathbf{u}^{n+1} .
4. Уравнение энергии решается в явном виде.

Алгоритм для несжимаемых течений на структурированных сетках

- Вычислительно-емкий препроцессор
- Гибкая конфигурация для широкого спектра вычислительных систем
- Базовый решатель для задач с одним периодическим направлением

Алгоритм решателя СЛАУ

1. FFT разделяет 3D задачу на набор независимых 2D подзадач (плоскостей).
2. Прямой метод на основе дополнений Шура для “плохих” низкочастотных плоскостей.
3. PCG для остальных плоскостей
4. IFFT восстанавливает 3D решение.



Разделение 3D подобласти на плоскости с помощью FFT

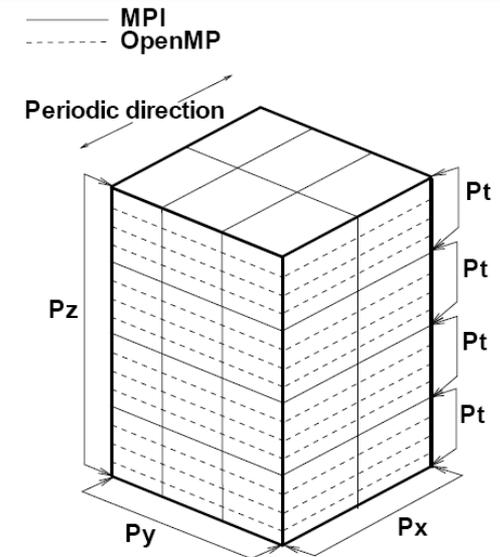
Двухуровневое MPI+OpenMP распараллеливание

MPI распараллеливание

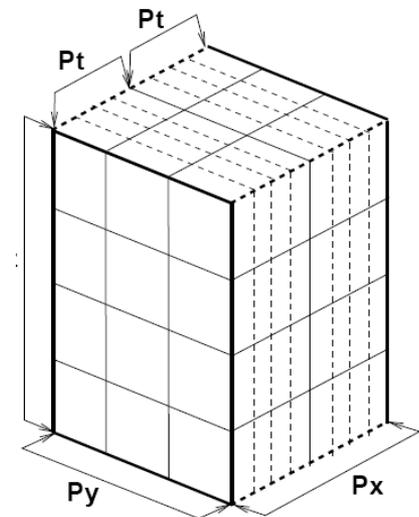
- Область разделяется на $P = P_x \times P_{yz}$ частей в 3-х направлениях
- P_x частей в периодическом направлении: FFT реплицируется внутри 1D групп.
- Каждая плоскость делится на P_{yz} частей

OpenMP распараллеливание

- Явная часть – декомпозиция циклов по узлам
- FFT (стадии 1,4): группы подвекторов делятся между нитями
- Наборы плоскостей разделяются между нитями (этапы 2, 3)
- I/O и MPI обмены выполняет главная нить



Physical space, explicit part

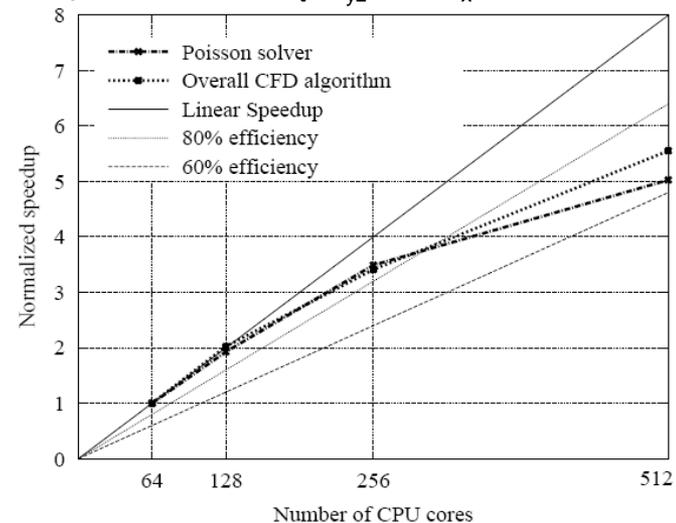


Fourier space, Poisson solver

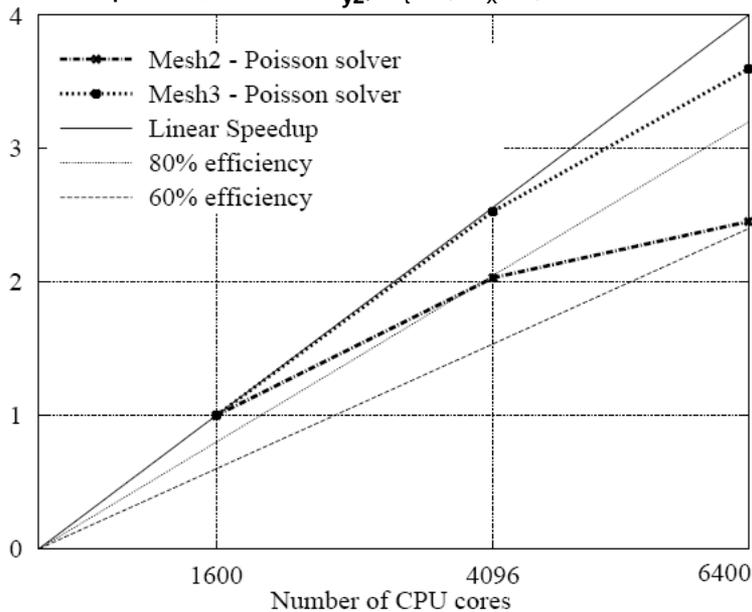
Тестирование распараллеливания кода STG-CFD&HT

- Общее число CPU ядер: $P = P_x \times P_{yz} \times P_t$
 - $P_t = 16$ – предел на системе с 16-ядерными узлами
 - $P_x = 8$ – предел из-за группового обмена
 - P_{yz} до ~800 из-за ограничений памяти и препроцесса
- оценка применимости: $800 \times 8 \times 16 = 102400$ CPU ядер

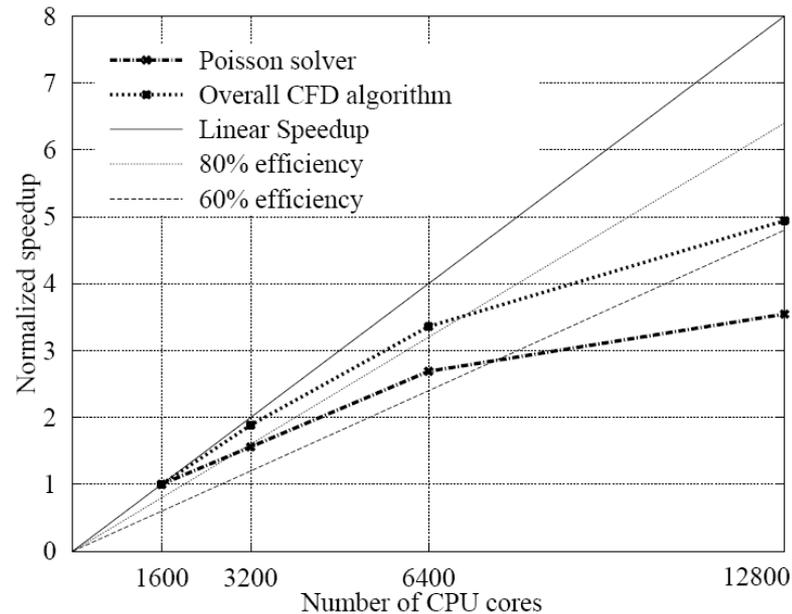
Ускорение, меняя P_t , $P_{yz}=64$, $P_x=1$, MVS100K



Ускорение, меняя P_{yz} , $P_t=8$, $P_x=1$, Ломоносов

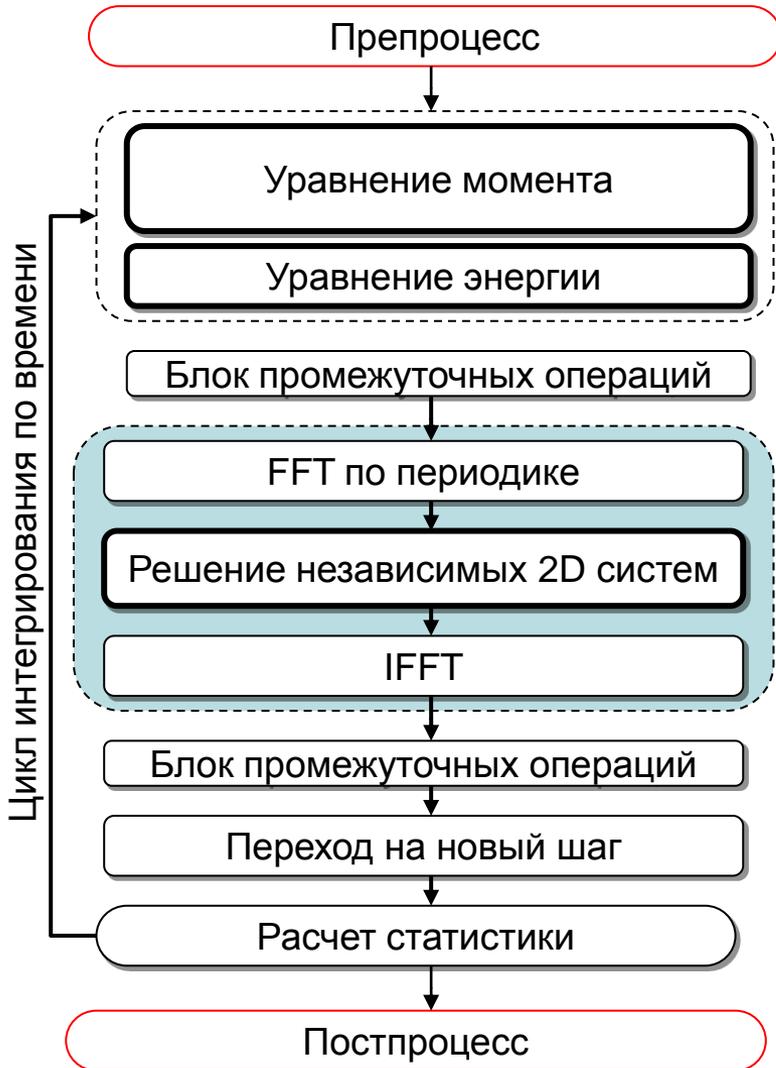


Сетка Mesh2 содержит 330M узлов, сетка Mesh3 – 1003M узлов



Ускорение, меняя P_x , $P_{yz}=200$, $P_t=8$, Ломоносов

Алгоритм и гетерогенная реализация на OpenCL



Базовые операции

T1. Линейный оператор

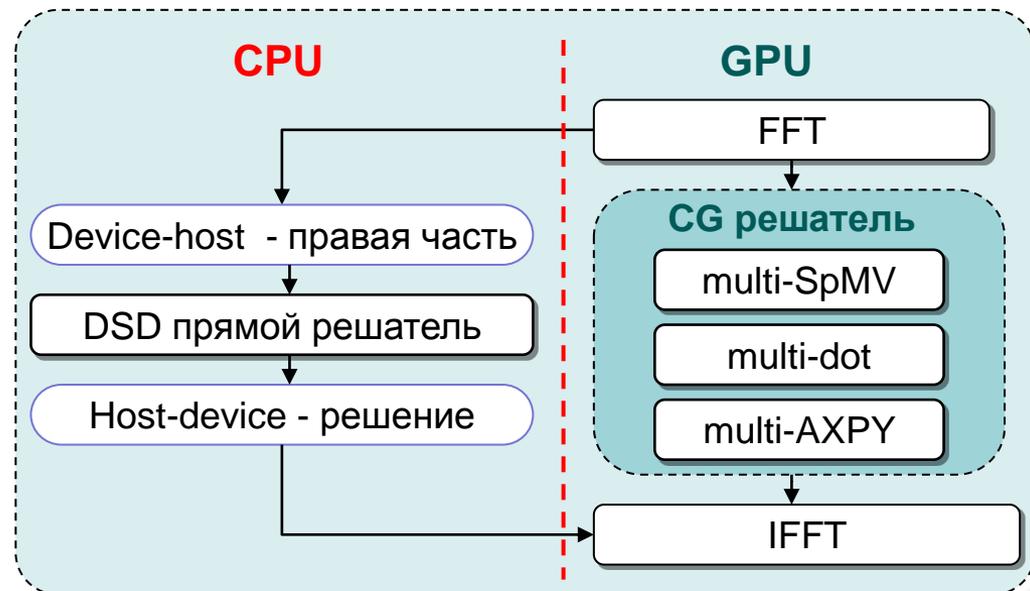
T2. Нелинейный

T3. $y = \sum_{i=1}^n a_i x_i + c$

T4. FFT, IFFT

T5. мульти-SpMV

T6. скалярное произведение



Представление оператора конвекции в виде SpMV

C – нелинейный, зависит от поля скоростей

Упрощение: сводим все к SpMV

Превращаем **C** в два SpMV, где выходной вектор первой операции является коэффициентами матрицы второй операции

Operators

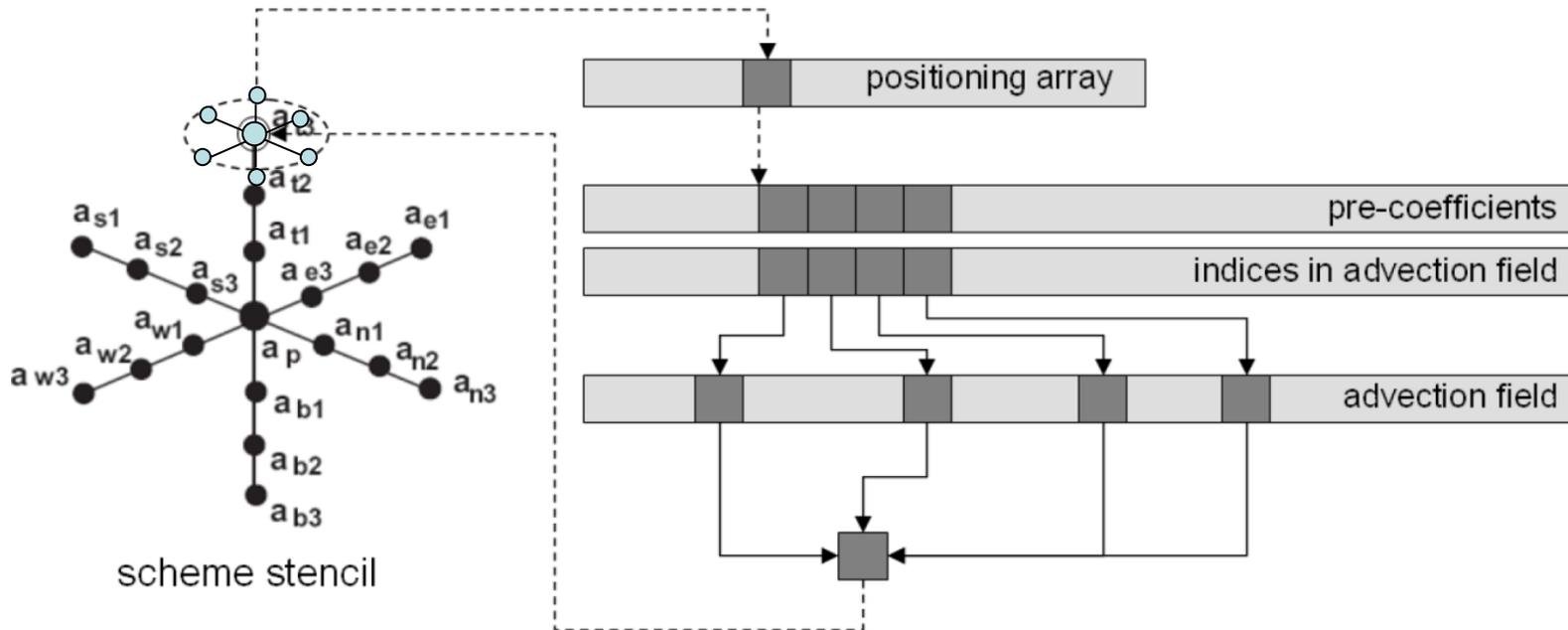
C – convection

D – diffusion

G – gradient

M – divergence

...



Подход к программной реализации

- Вычислительный алгоритм сведен к трем стандартным операциям линейной алгебры, совместимым с потоковой обработкой: **SpMV, dot, axpy**

Препроцесс:

обработка сеточных данных, построение дискретизации, расчет геометрии контрольных объемов, генерация матриц

- Объектно-ориентированный подход, удобство использования, сложные структуры данных,



Интегрирование по времени:

“Алгебраический” подход, простые структуры данных, ориентированные на максимальную производительность вычислений

- Подбор оптимальных форматов матриц для разных этапов алгоритма: **COO, CSR, ELLPACK, sliced ELLPACK...**

*G. Oyarzun, R. Borrell, A. Gorobets, A. Oliva. "Portable implementation model for CFD simulations. Application to hybrid CPU/GPU supercomputers". International Journal of Computational Fluid Dynamics.

Главная проблема - SpMV

И опять все усложняется...

- Different matrix formats can be used depending on matrix structure
- Fused SpMV – result of one SpMV is a matrix for another
- Generalized SpMV – one matrix, multiple vectors
- Communication and computation overlap
- Device-oriented data format adaptation
- RCM band reduction

ELLPACK

6			7						
9		1							
		3							5
	4			2					
7			8			1			
2		5							6
	1			7			9		
			4		8			3	

CPU sliced ELLPACK (S=4)

Val

6	7	9	1	3	5	4	2	7	8	1	2	5	6	1	7	9	4	8	3
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Col

0	3	0	2	2	7	1	4	0	3	5	0	2	7	1	4	6	3	5	7
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

start_slice

0	8	24
---	---	----

GPU sliced ELLPACK (S=4)

Val

6	9	3	4	7	1	5	2	7	2	1	4	8	5	7	8	1	6	9	3
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

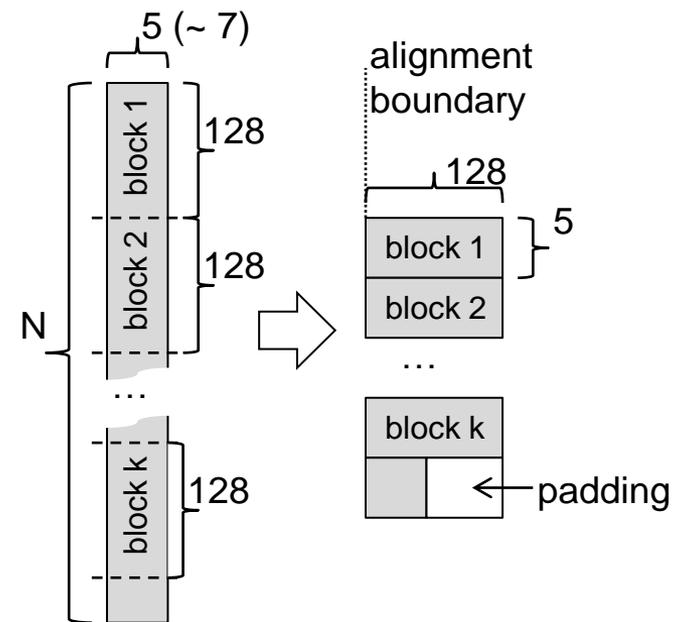
Col

0	0	2	1	3	2	7	4	0	0	1	3	3	2	4	5	5	7	6	7
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

start_slice

0	8	24
---	---	----

Трансформация структуры данных для формата ELLPACK



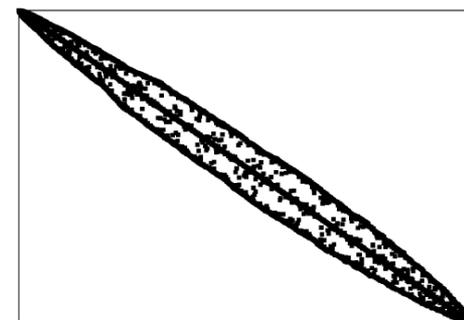
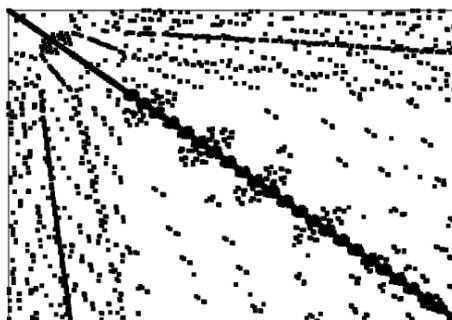
G. Oyarzun, R. Borrell, A. Gorobets, A. Oliva,

MPI-CUDA sparse matrix–vector multiplication for the conjugate gradient method with an approximate inverse preconditioner

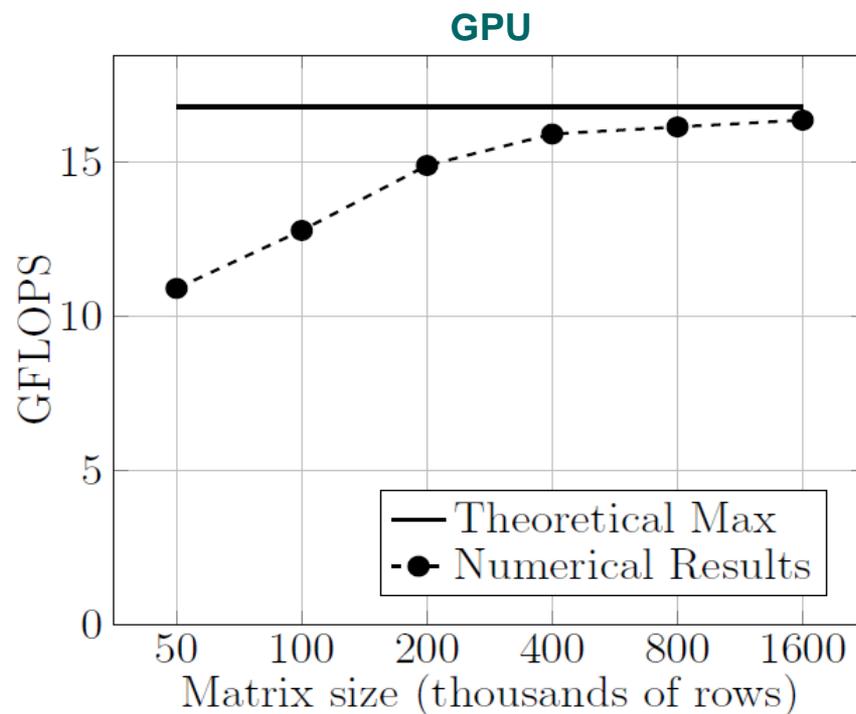
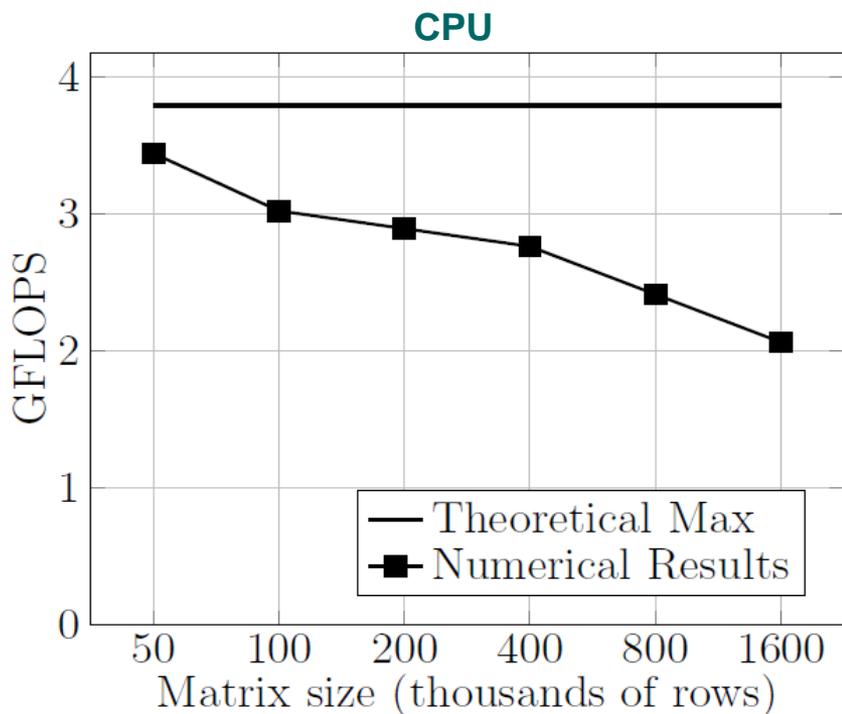
Computers and Fluids, Volume 92, 20 March 2014, Pages 244–252

Производительность SpMV на одиночном устройстве

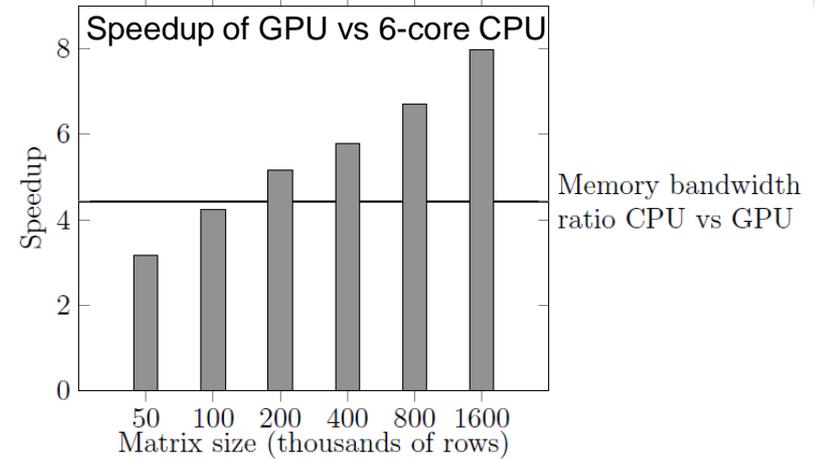
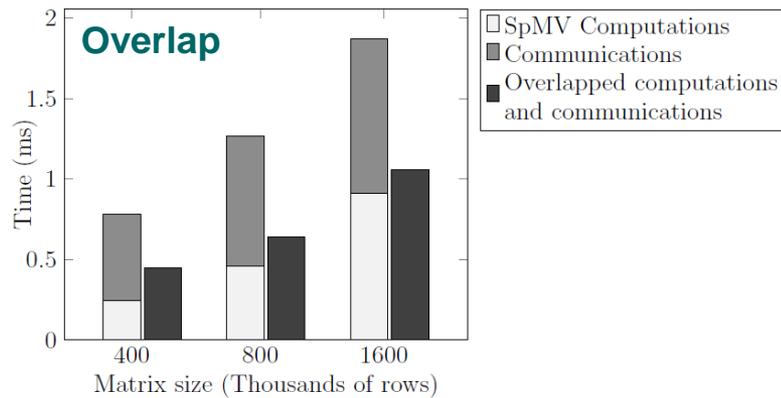
- Тесты для формата ELLPACK
- CPU Intel Xeon E5649 (6 ядер)
- GPU NVIDIA M2090
- Максимальная теоретически достижимая производительность на CPU 9.3% от пика, на GPU 2.5%



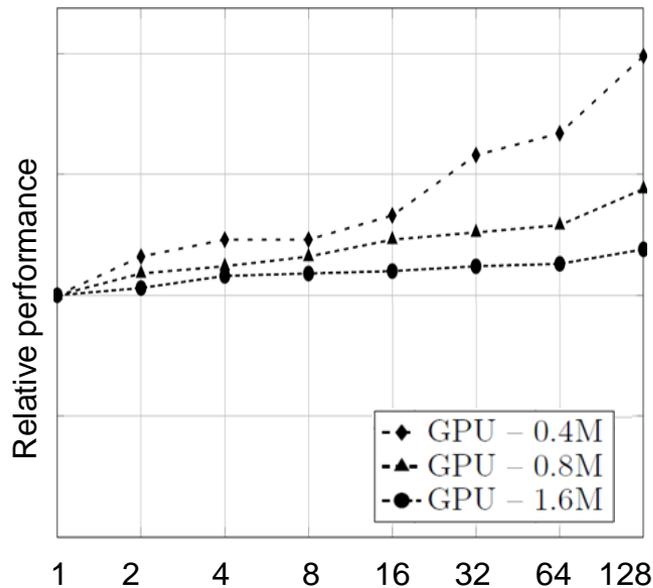
Переупорядочивание Катхилла — Макки для оптимизации доступа к памяти



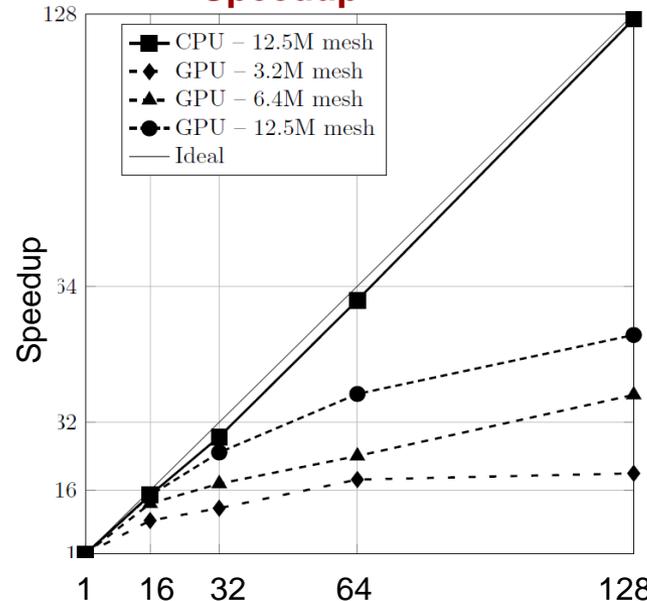
Производительность и параллельная эффективность



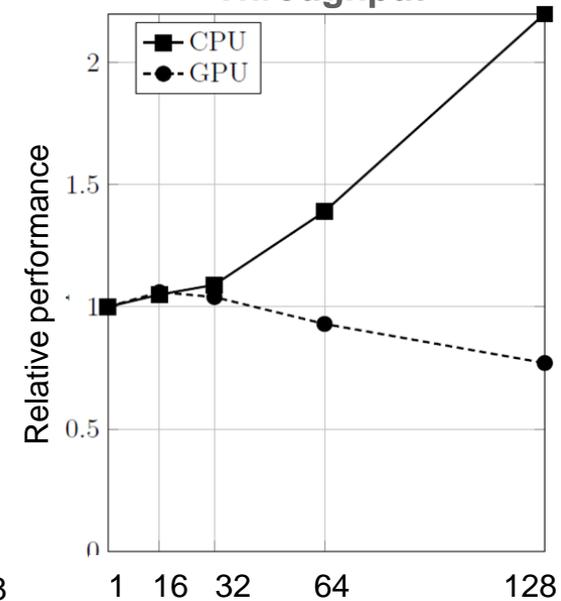
Scalability on GPUs



Speedup

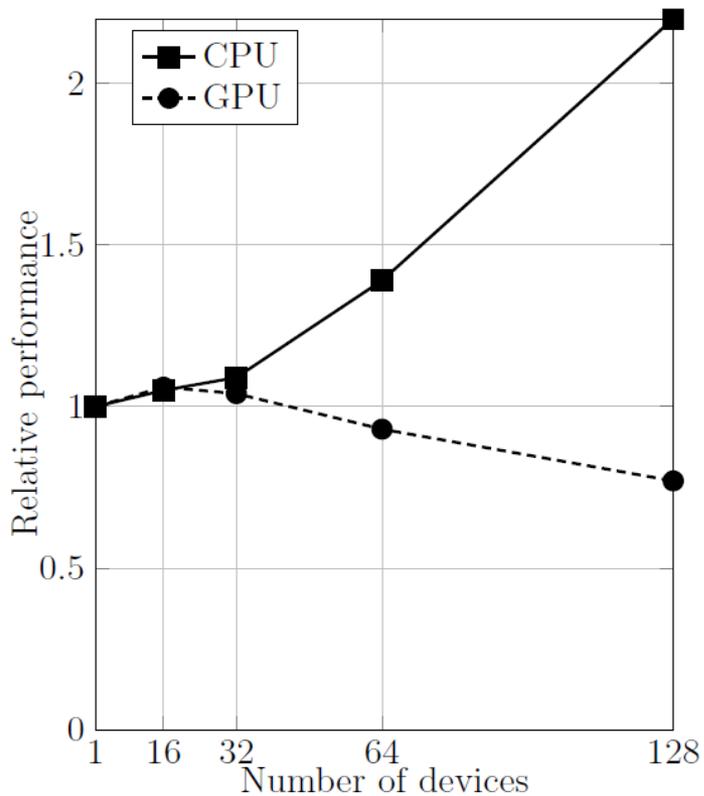


Throughput

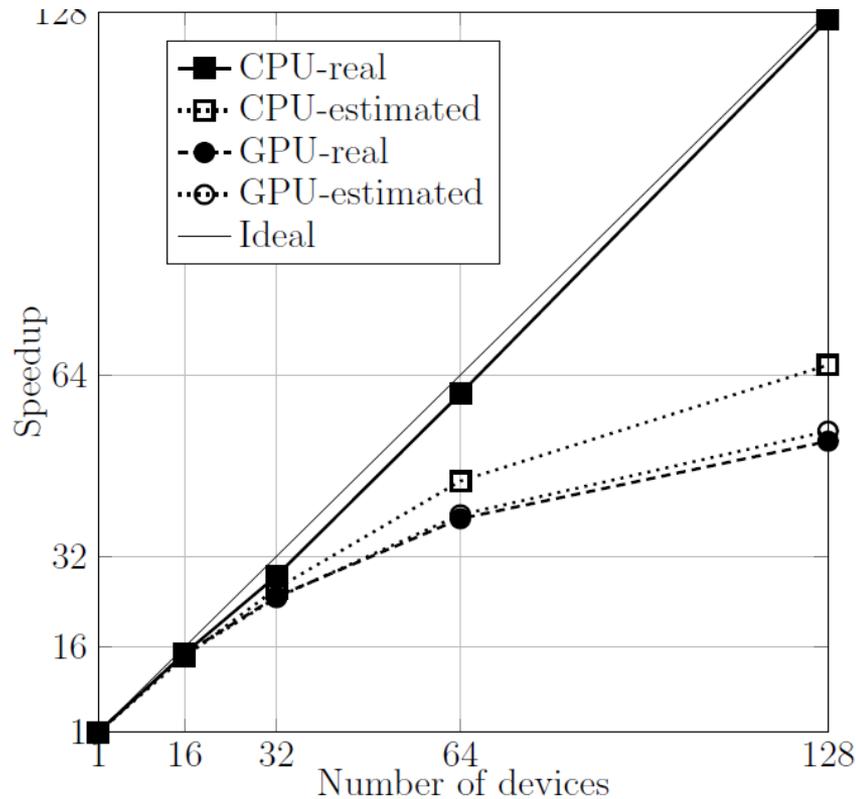


*G. Oyarzun, R. Borrell, A. Gorobets, A. Oliva. "Portable implementation model for CFD simulations. Application to hybrid CPU/GPU supercomputers". International Journal of Computational Fluid Dynamics.

Multi-CPU vs multi-GPU

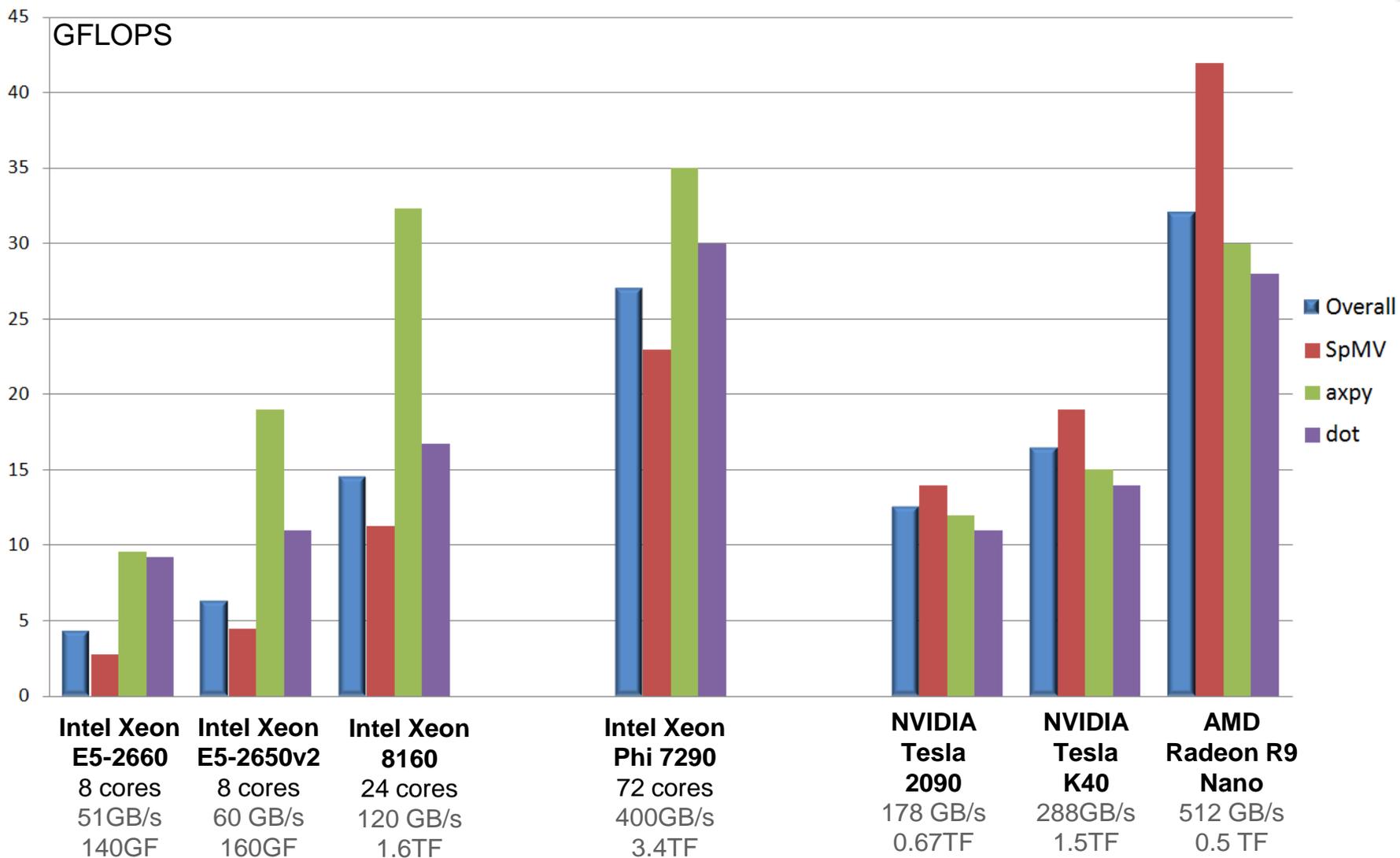


Относительная производительность устройства



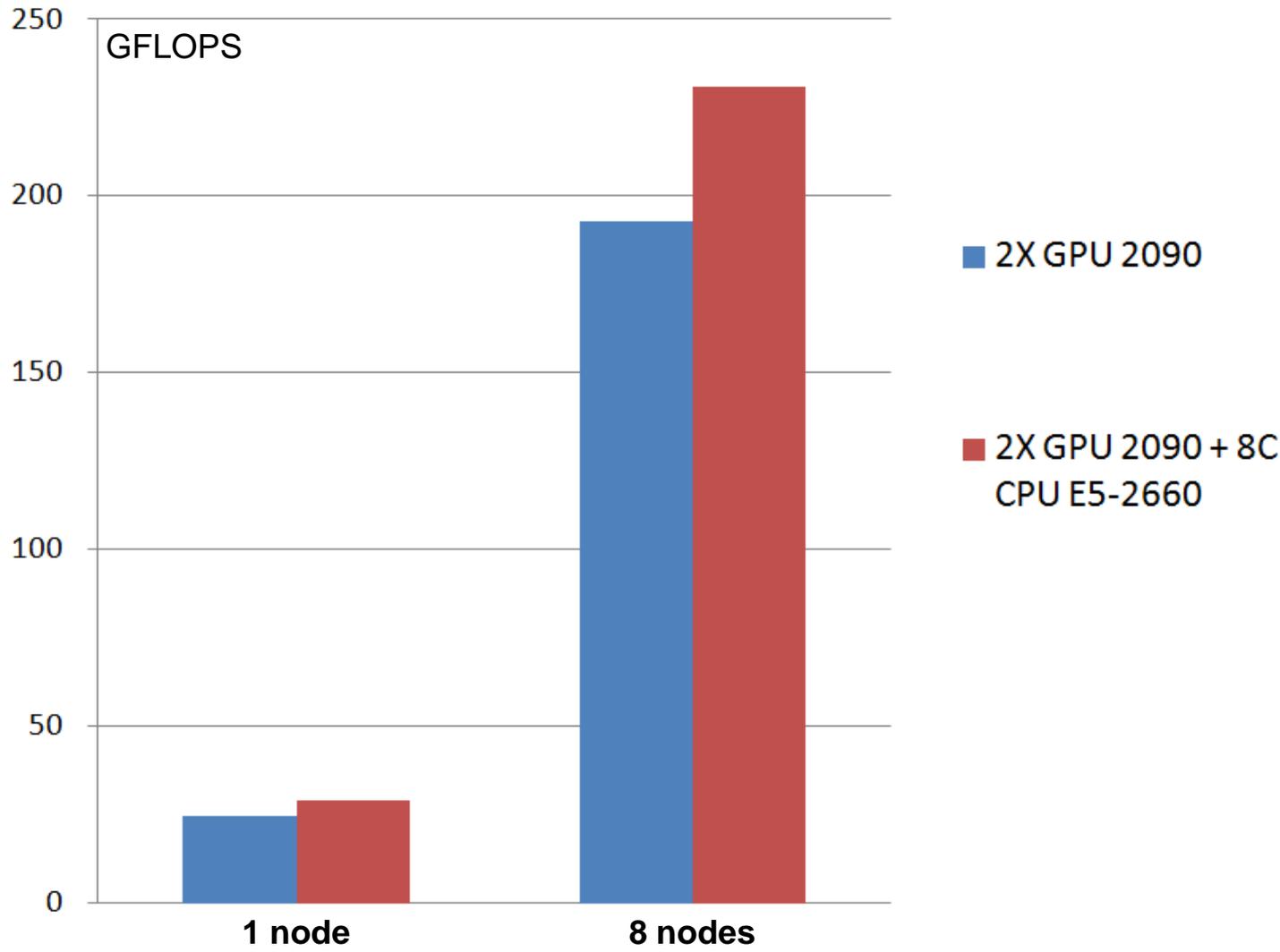
Оценка ускорения при отсутствии кэш-эффекта

Производительность на одиночном устройстве



Mesh 1M cells

Гетерогенный режим CPU+GPU



Mesh 1M cells per node

MONT-BLANC

- **Fused GPU+CPU ARM:**

CPU Coretex A15

2 cores, 1.7 Ghz, 12.8 GB/s, 6.8 GFLOPS

GPU Mali T604

4 cores, 533Mhz, 12.8GB/s, 21 GFLOPS

Physically shared memory

- **10Gb Ethernet network**

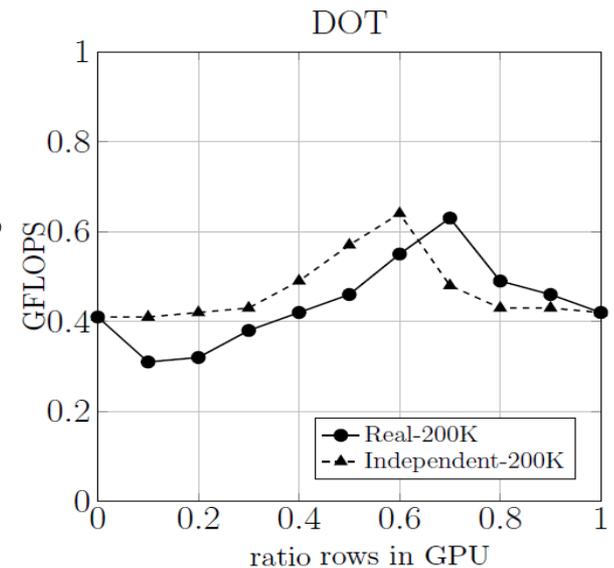
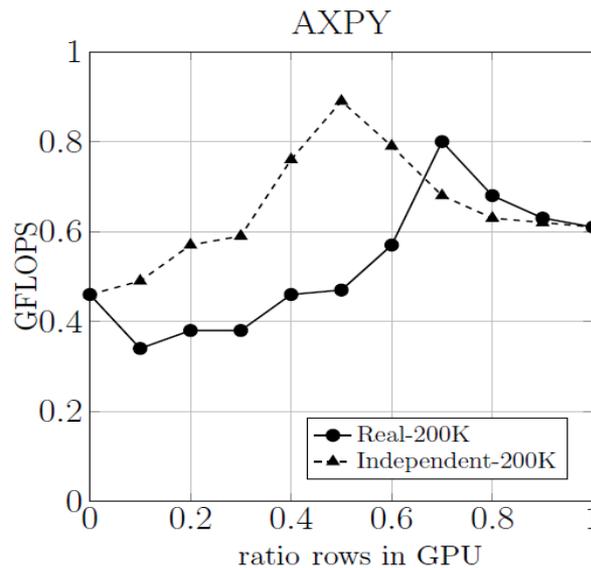
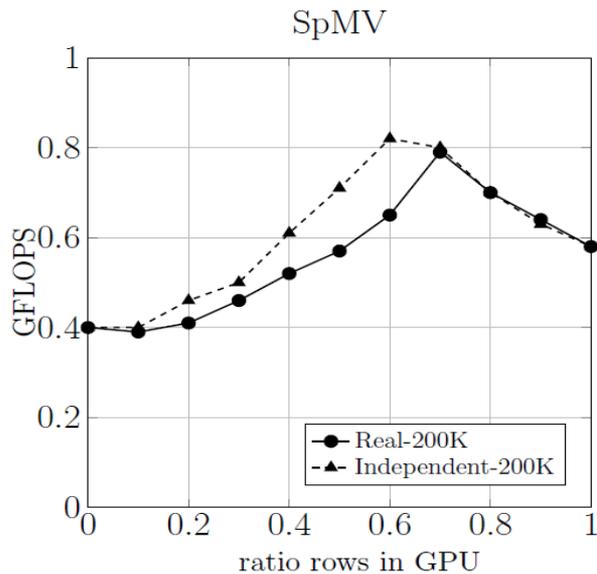
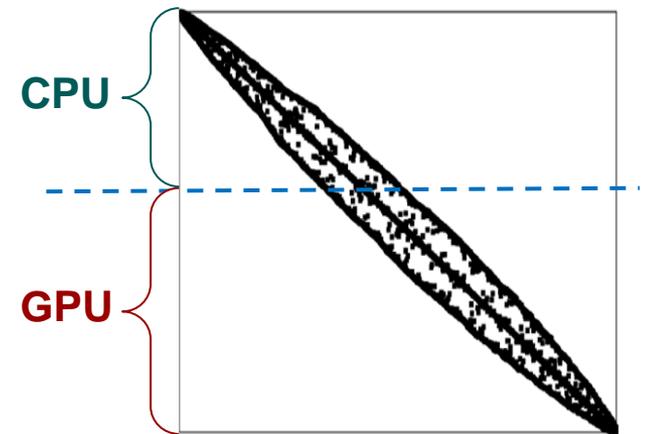
- **1080 compute cards, 2160 CPU cores and 1080 GPUs**



*G.Oyarzun, R. Borrell, A. Gorobets, F. Mantovani, A. Oliva. "Efficient CFD code implementation for the ARM-based Mont-Blanc architecture". Future Generation Computer Systems. In press. <https://doi.org/10.1016/j.future.2017.09.029>

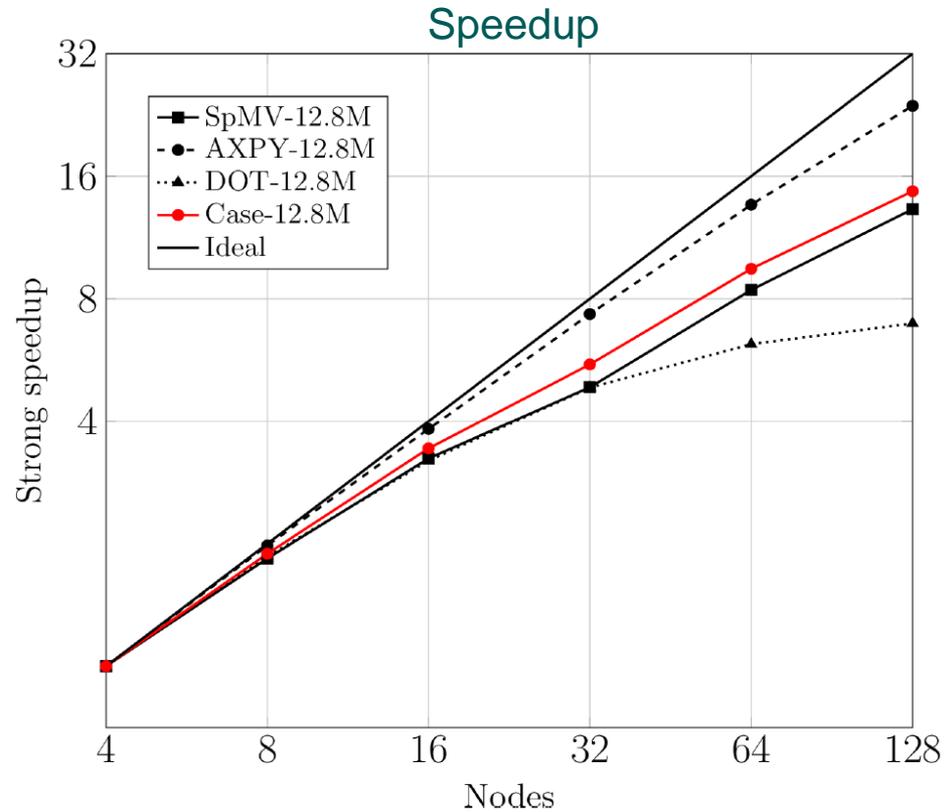
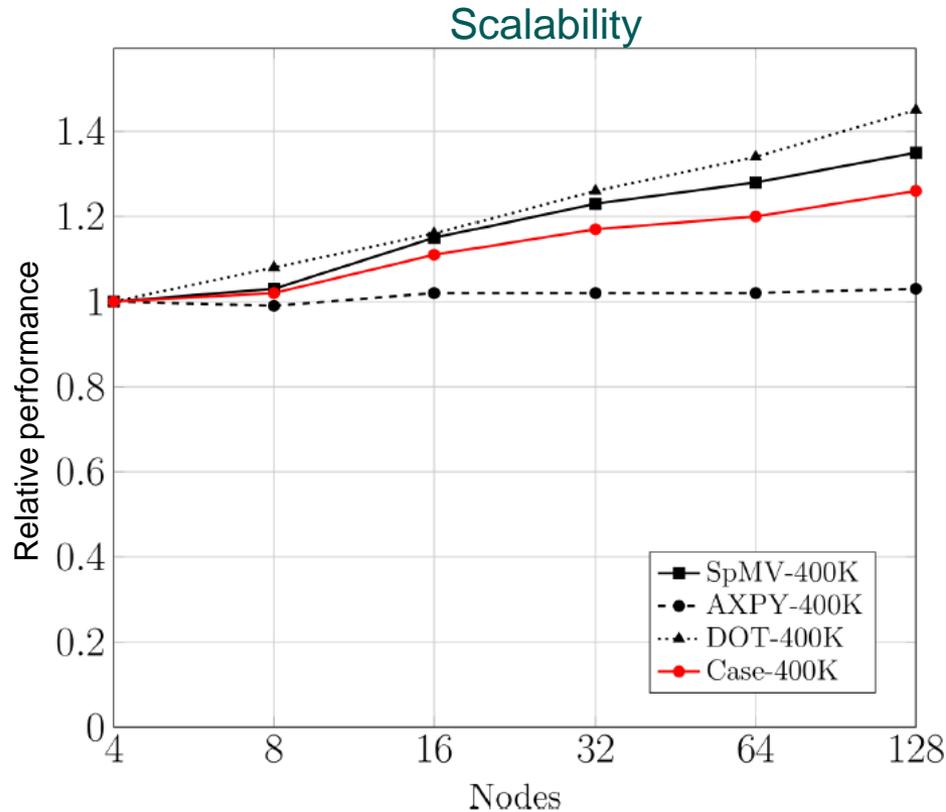
Parallel performance on ARM machine

- Automatic dynamic load balancing between CPU and GPU
- Balancing based on separate timing of CPU and GPU is far from optimal as devices compete for bandwidth
- Performance may vary runtime



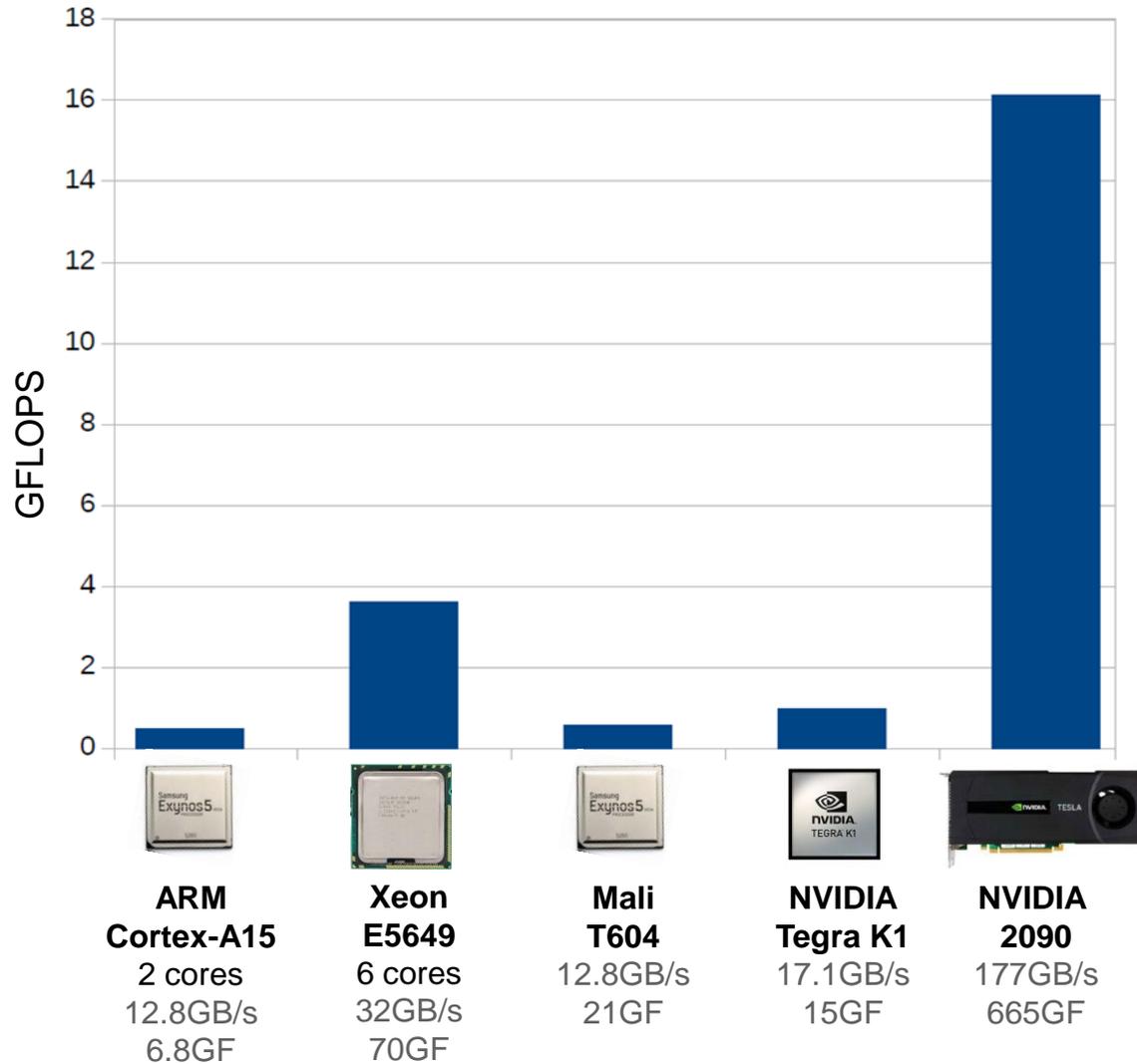
Parallel performance on ARM machine

MPI-parallel execution on multiple nodes



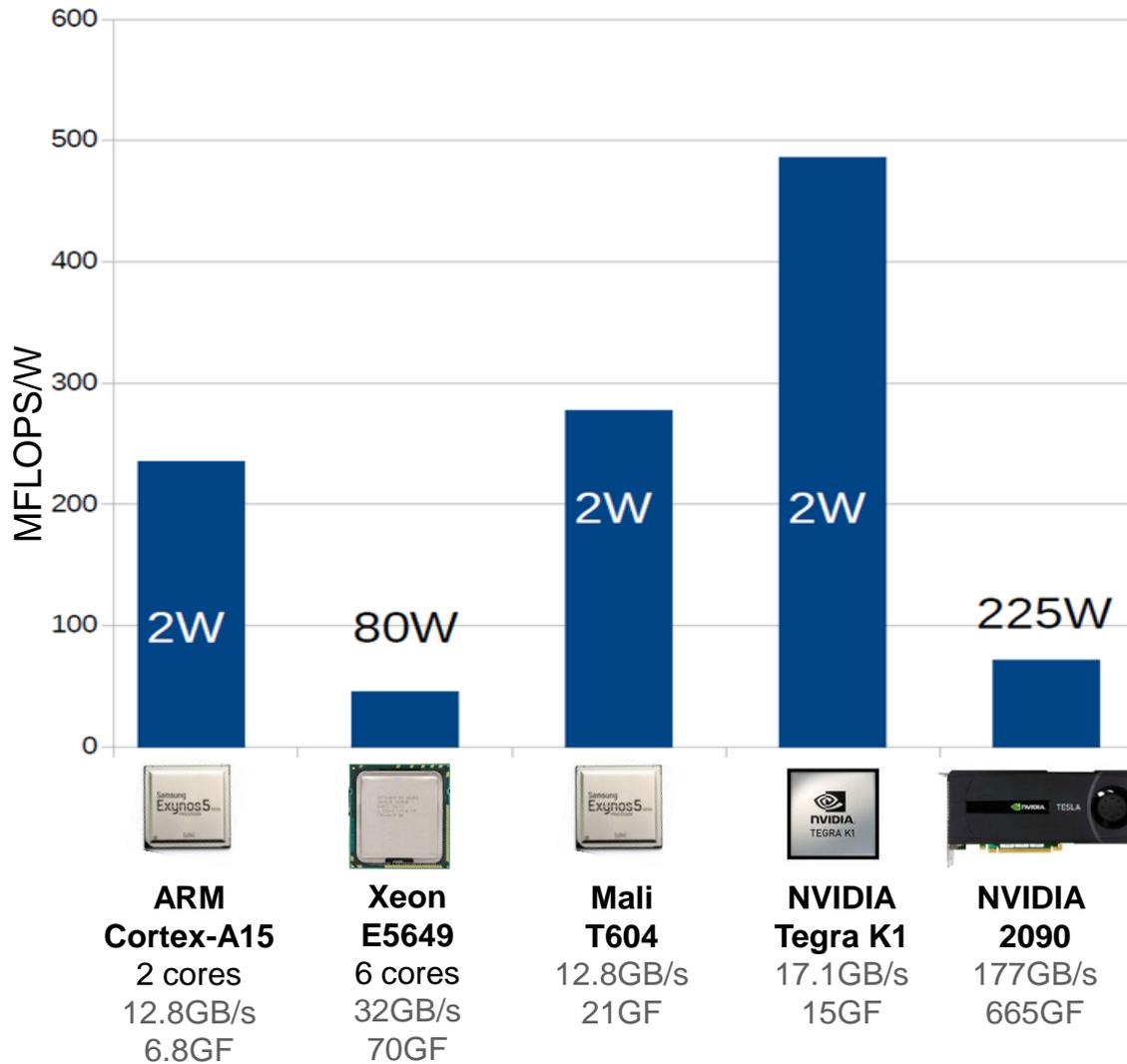
*G.Oyarzun, R. Borrell, A. Gorobets, F. Mantovani, A. Oliva. "Efficient CFD code implementation for the ARM-based Mont-Blanc architecture". Future Generation Computer Systems. In press. <https://doi.org/10.1016/j.future.2017.09.029>

Comparison of performance on single devices



*G.Oyarzun, R. Borrell, A. Gorobets, F. Mantovani, A. Oliva. "Efficient CFD code implementation for the ARM-based Mont-Blanc architecture". Future Generation Computer Systems. In press. <https://doi.org/10.1016/j.future.2017.09.029>

Comparison of performance on single devices



*G.Oyarzun, R. Borrell, A. Gorobets, F. Mantovani, A. Oliva. "Efficient CFD code implementation for the ARM-based Mont-Blanc architecture". Future Generation Computer Systems. In press. <https://doi.org/10.1016/j.future.2017.09.029>

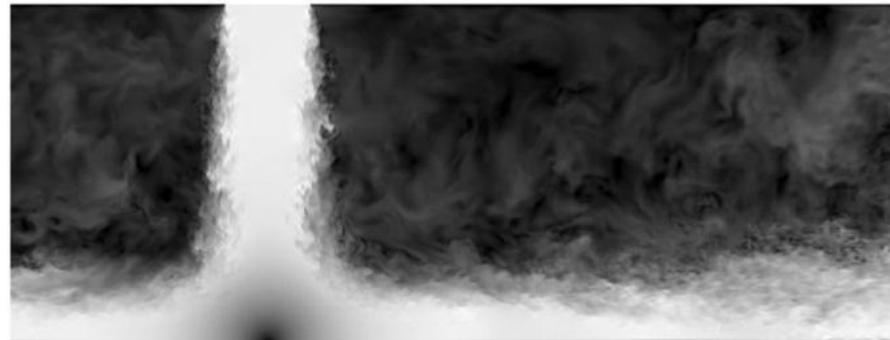
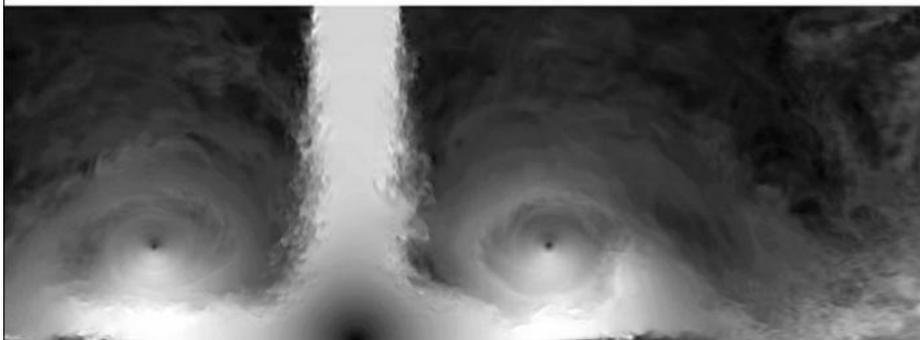
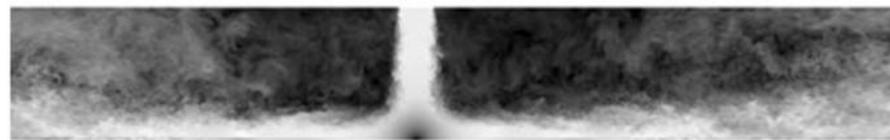
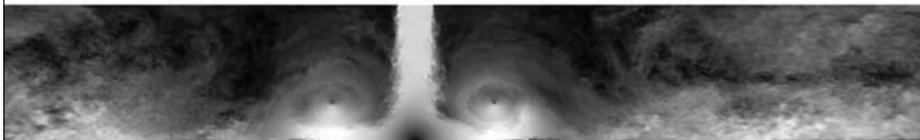
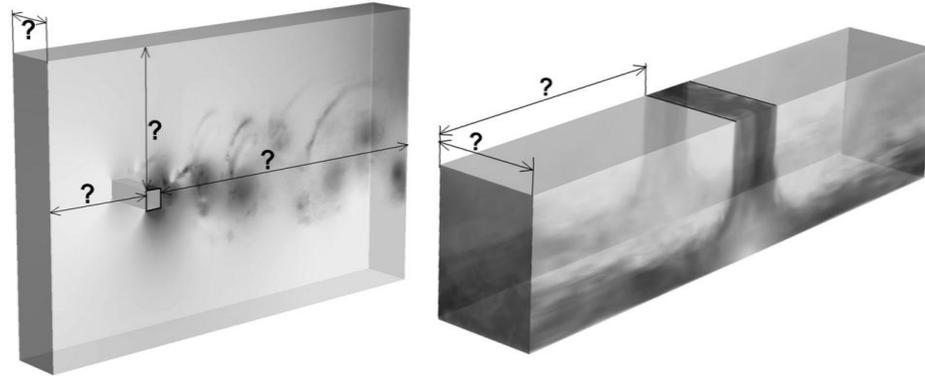
Условия работы на системе коллективного доступа

- Существенное время ожидания задачи в очереди при каждом запуске;
- ограниченный квант, то есть максимальное время, доступное для одного запуска;
- возможность остановки задачи в произвольный момент времени (по истечению кванта, в случае сбоя или внезапной профилактики);
- ограниченная дисковая квота;
- относительно низкая надежность файловой системы временного хранения данных.

Технология выполнения расчета

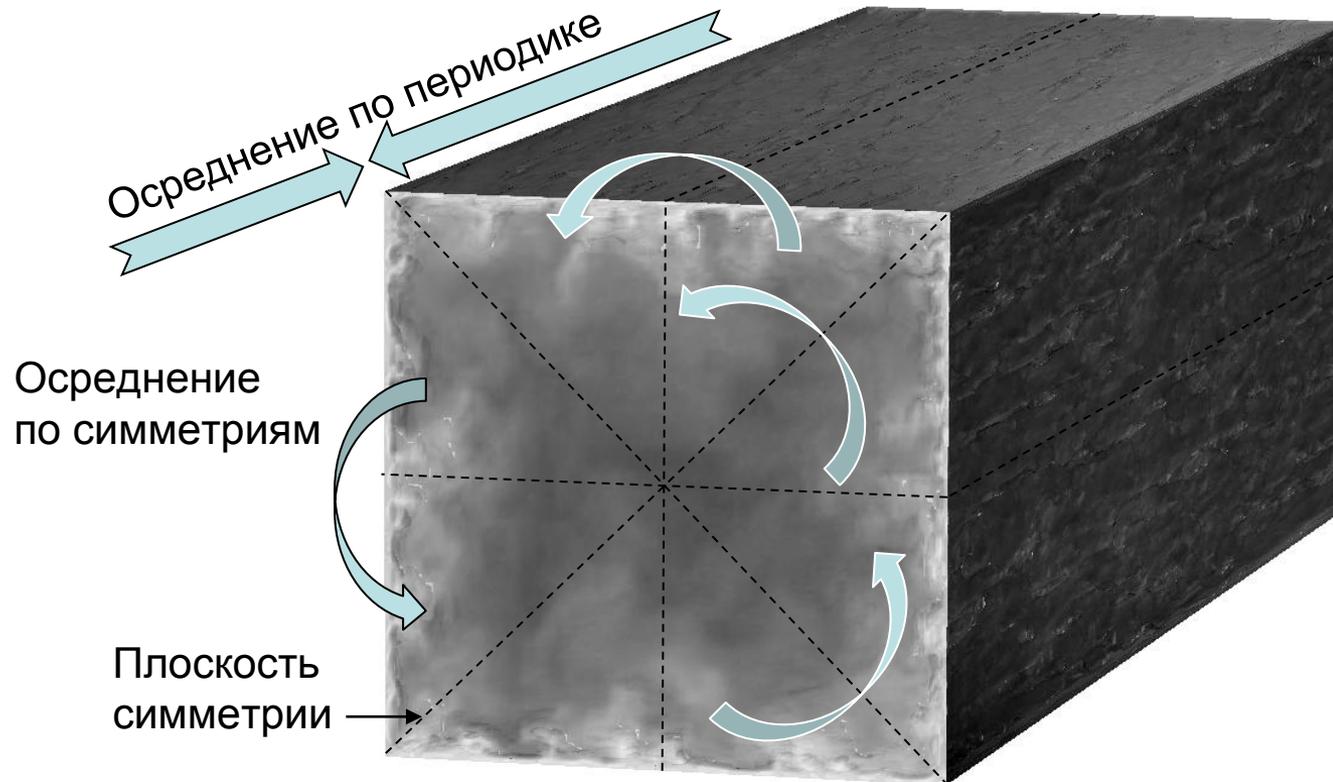
Выполняется два шага сгущения сетки

- чтобы исключить некорректность в постановке расчета;
- чтобы настроить ключевые параметры, влияющие на вычислительную стоимость: размер расчетной области, величина шага по времени, конфигурация решателя СЛАУ, невязки итерационных процессов, и т.д.;
- чтобы ускорить выход на установившийся статистически однородный режим течения, используя поля с грубой сетки в качестве начального распределения;



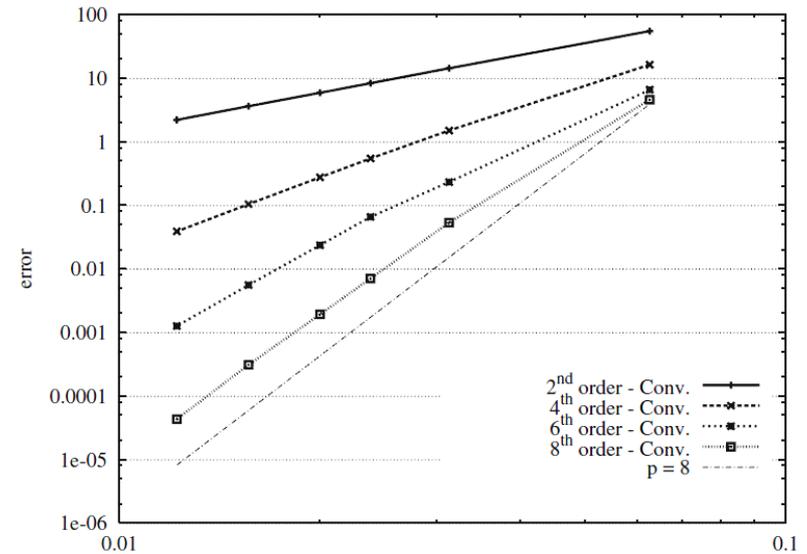
Развитие течения и установившееся течение для отбойной струи в канале

- Все пространственные симметрии должны использоваться для повышения качества статистики течения



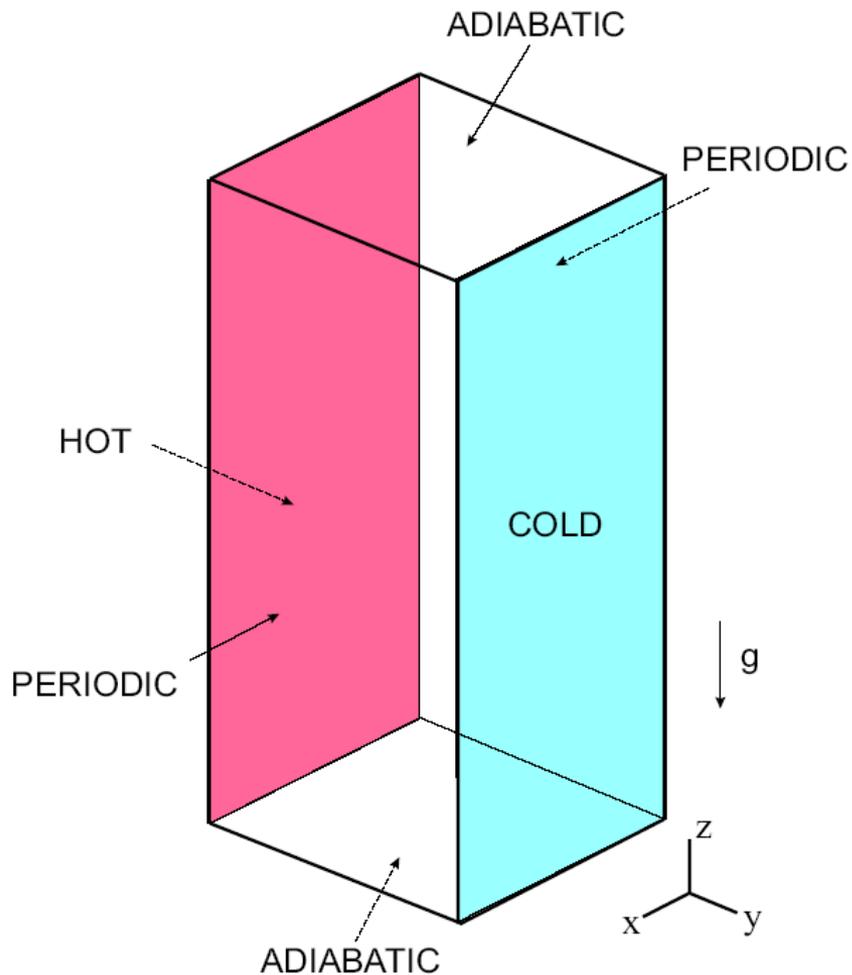
Общая процедура выполнения DNS расчета

- Верификация дискретных операторов - MMS
- Серия предварительных расчетов
- Корректность невязки решателя (запас минимум 2 порядка)
- Корректность геометрии
- Обоснование сеточного разрешения в пристенной области ($y^+ < 1$) и области турбулентного течения
- Сходимость по сетке
- Контроль качества осреднения
- Контроль момента начала осреднения
- Сравнение с экспериментальными и численными данными (если существуют) для моделируемой конфигурации или для аналогичной с меньшими числами Re (Ra)



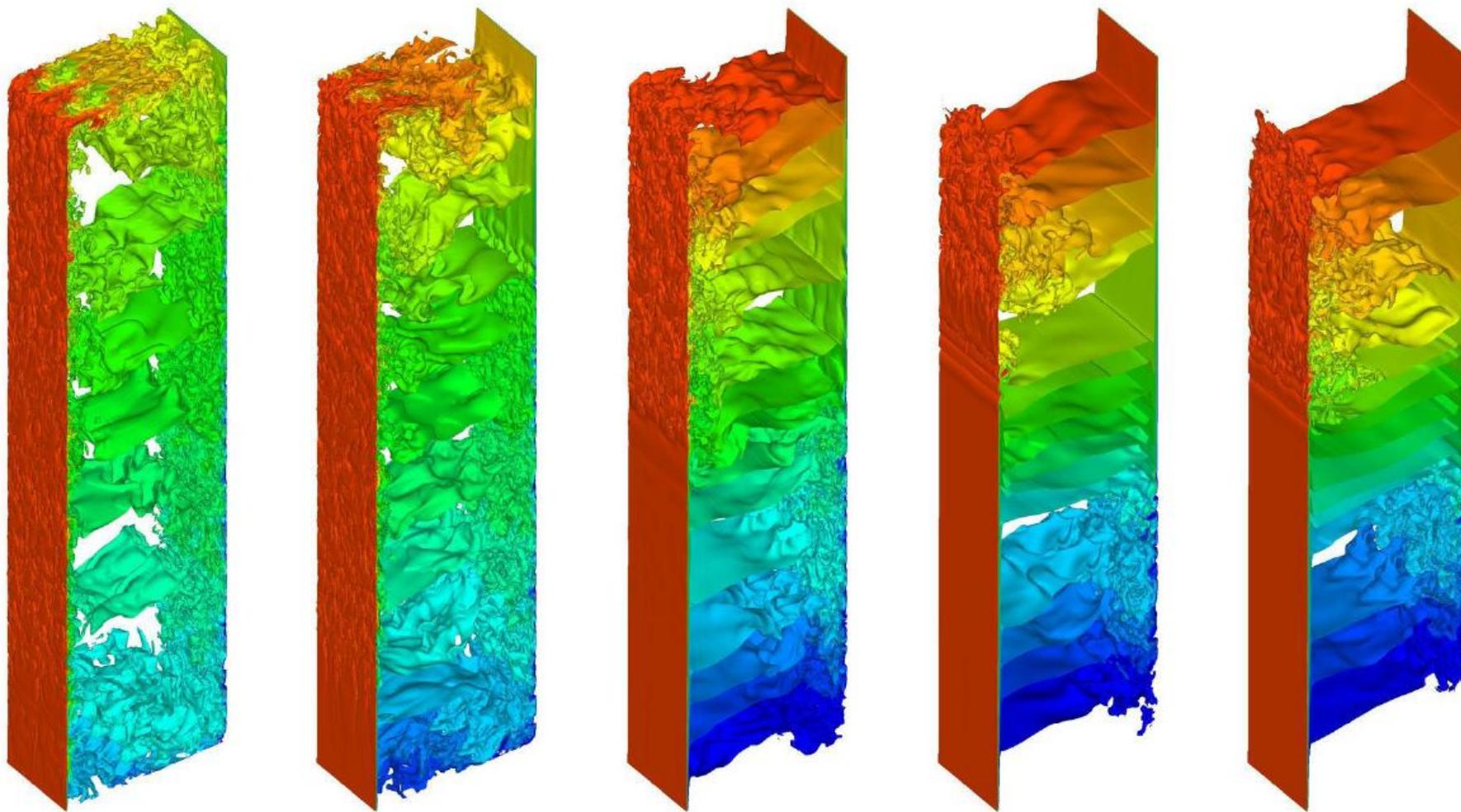
Порядок сходимости для дискретных операторов

DNS 3D каверны с разнотемпературными стенками



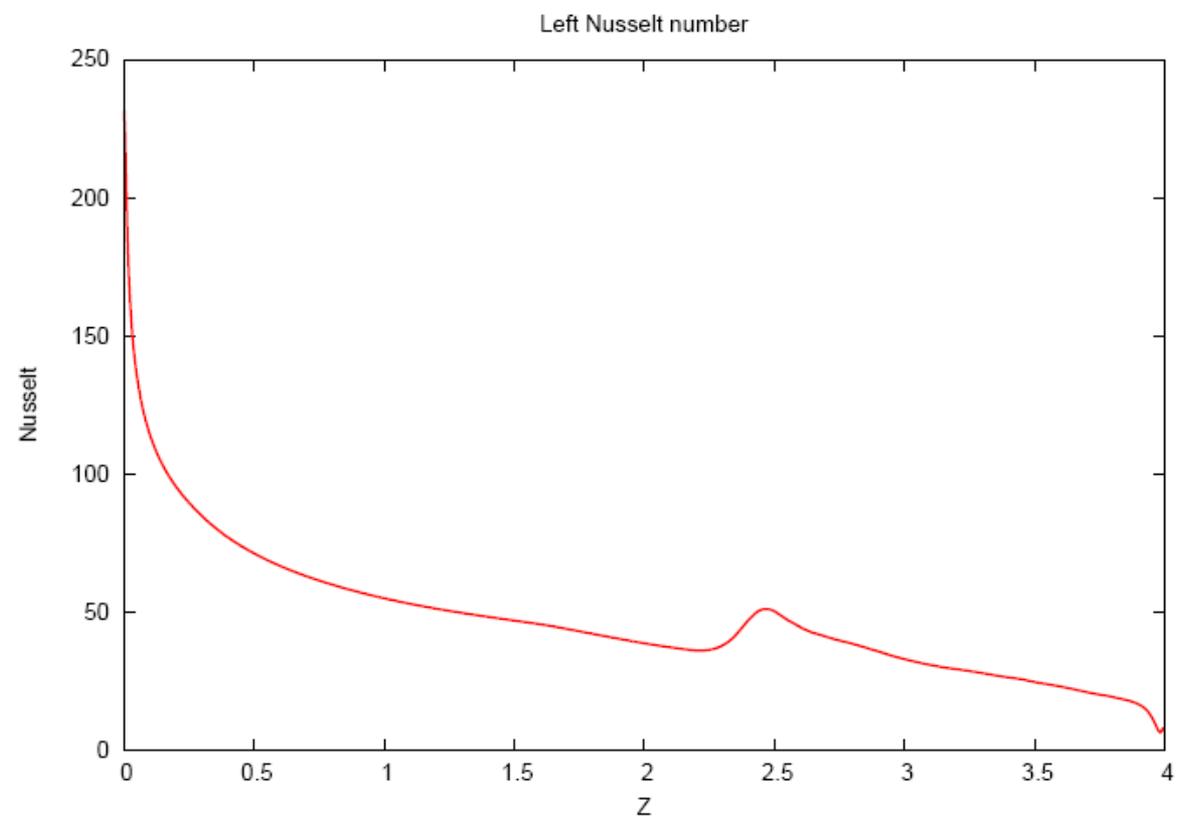
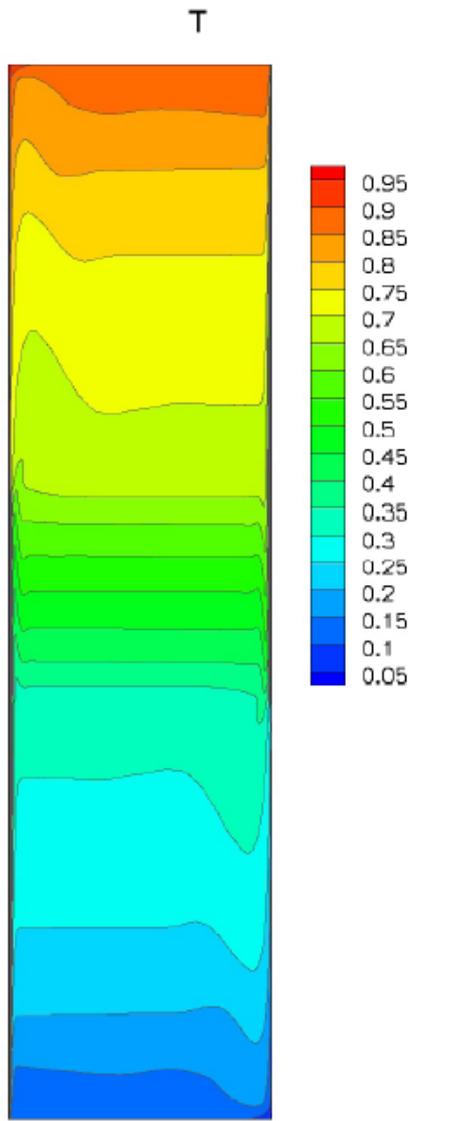
- Размер сетки 128x680x1280, 111×10^8 узлов
- DNS выполнено на 512 процессорах суперкомпьютера Marenostrum
- схема 4-го порядка аппроксимации
- $Ra = 10^{11}$, $Pr = 0.71$ (воздух)
- Соотношение высоты и ширины – 4 к 1

F. X. Trias, A. Gorobets, M. Soria, A. Oliva, Direct numerical simulation of a differentially heated cavity of aspect ratio 4 with Ra-number up to 10^{11} - Part I: Numerical methods and time-averaged flow, International Journal of Heat and Mass, 53 (2010), 665-673

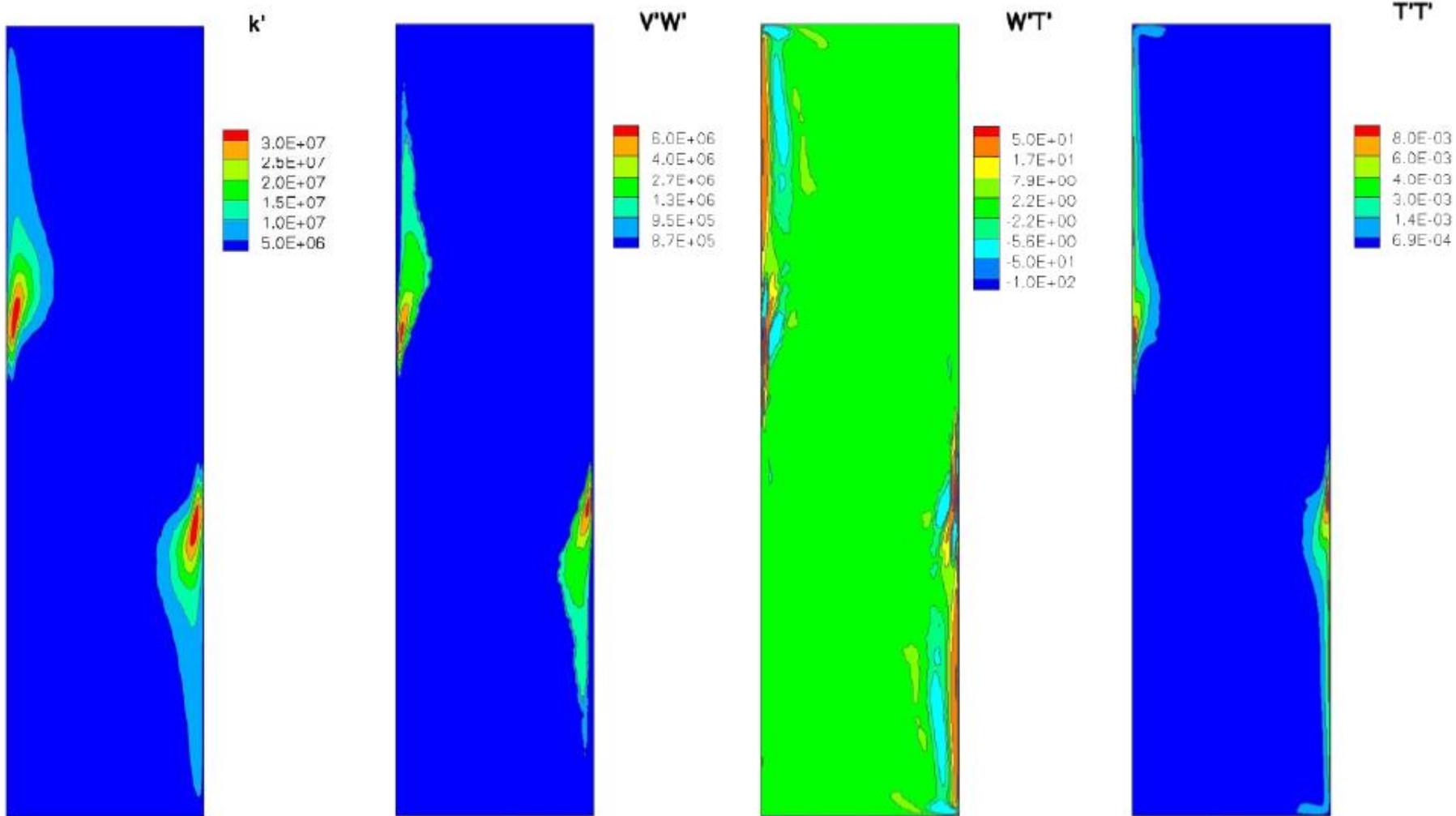


Эволюция картины течения со временем
(изоповерхности температуры)

DNS 3D каверны с разнотемпературными стенками

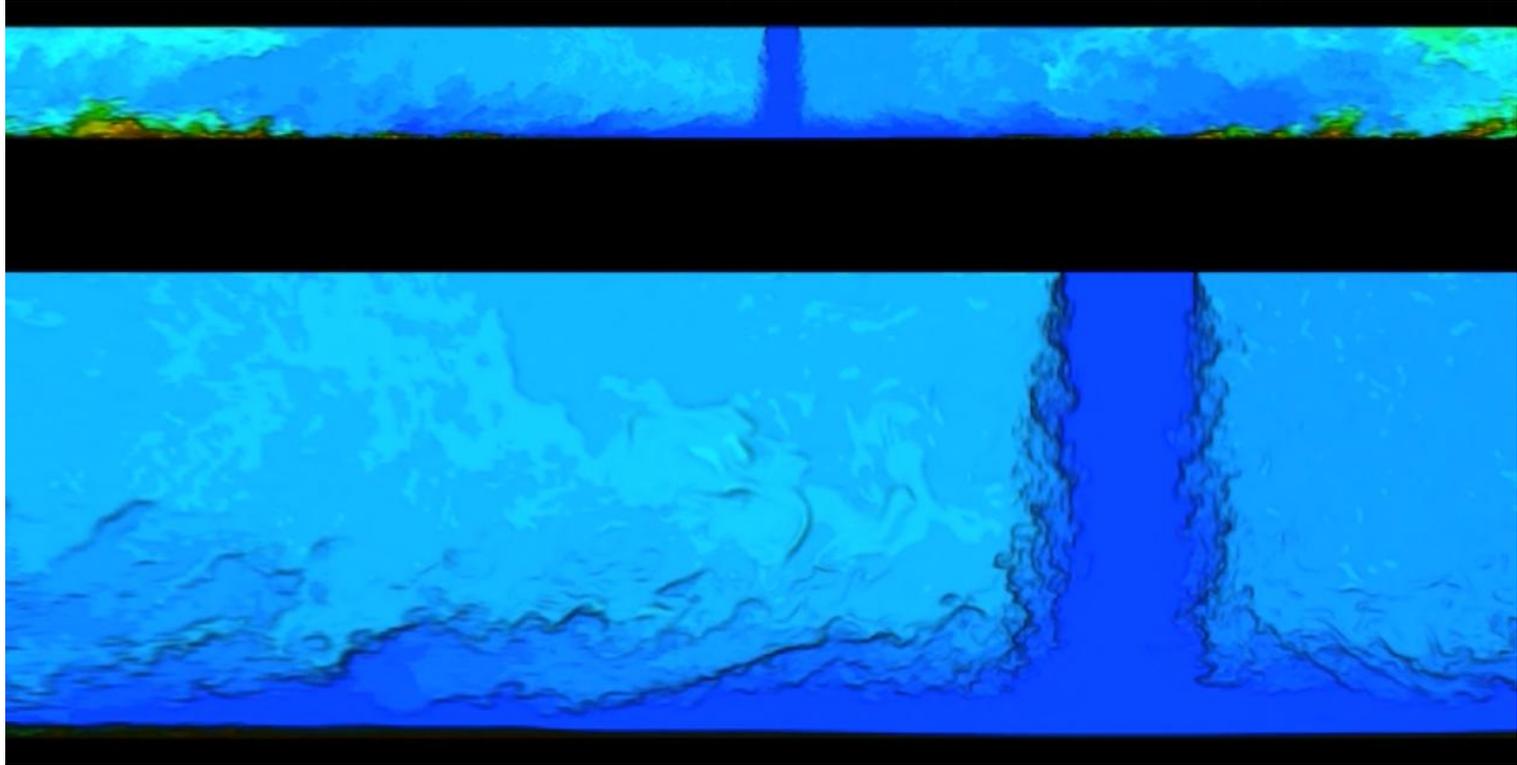


DNS 3D каверны с разнотемпературными стенками



DNS отбойной струи

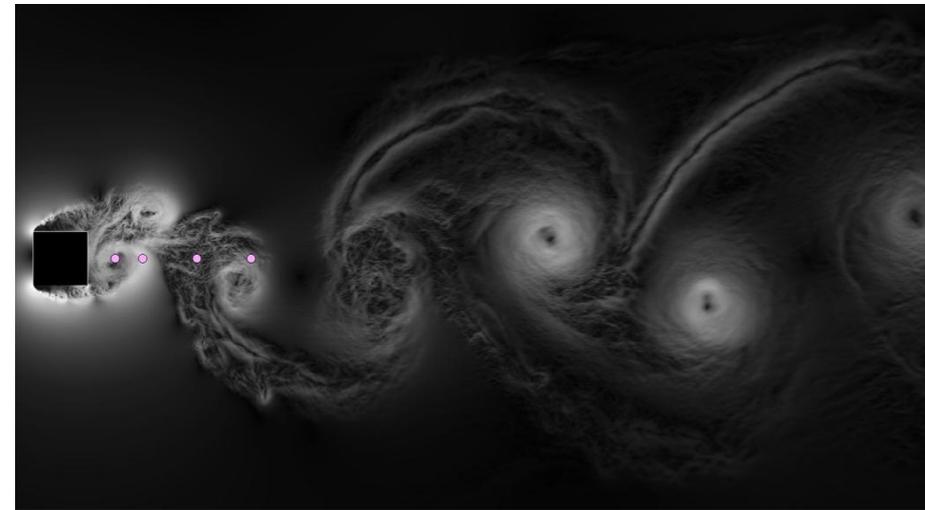
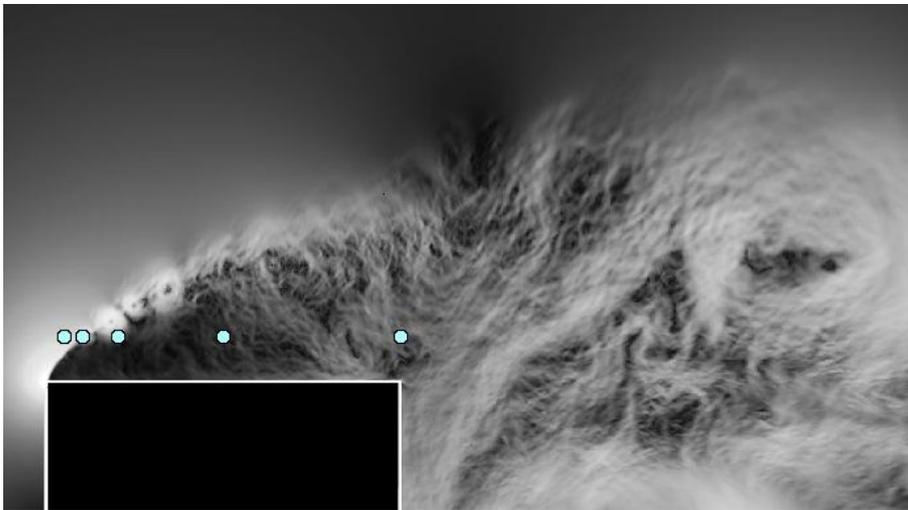
- $Re_{\tau} = 20000$
- Высота канала = 4 ширины щели
- Сетка 100M узлов
- 320 CPU на MareNostrum



J. E. Jaramillo, F. X. Trias, A. Gorobets, C. D. Perez-Segarra, A. Oliva, DNS and RANS modelling of a Turbulent Plane Impinging Jet, International Journal of Heat and Mass Transfer, Volume 55, Issue 4, 31 January 2012, Pages 789–801

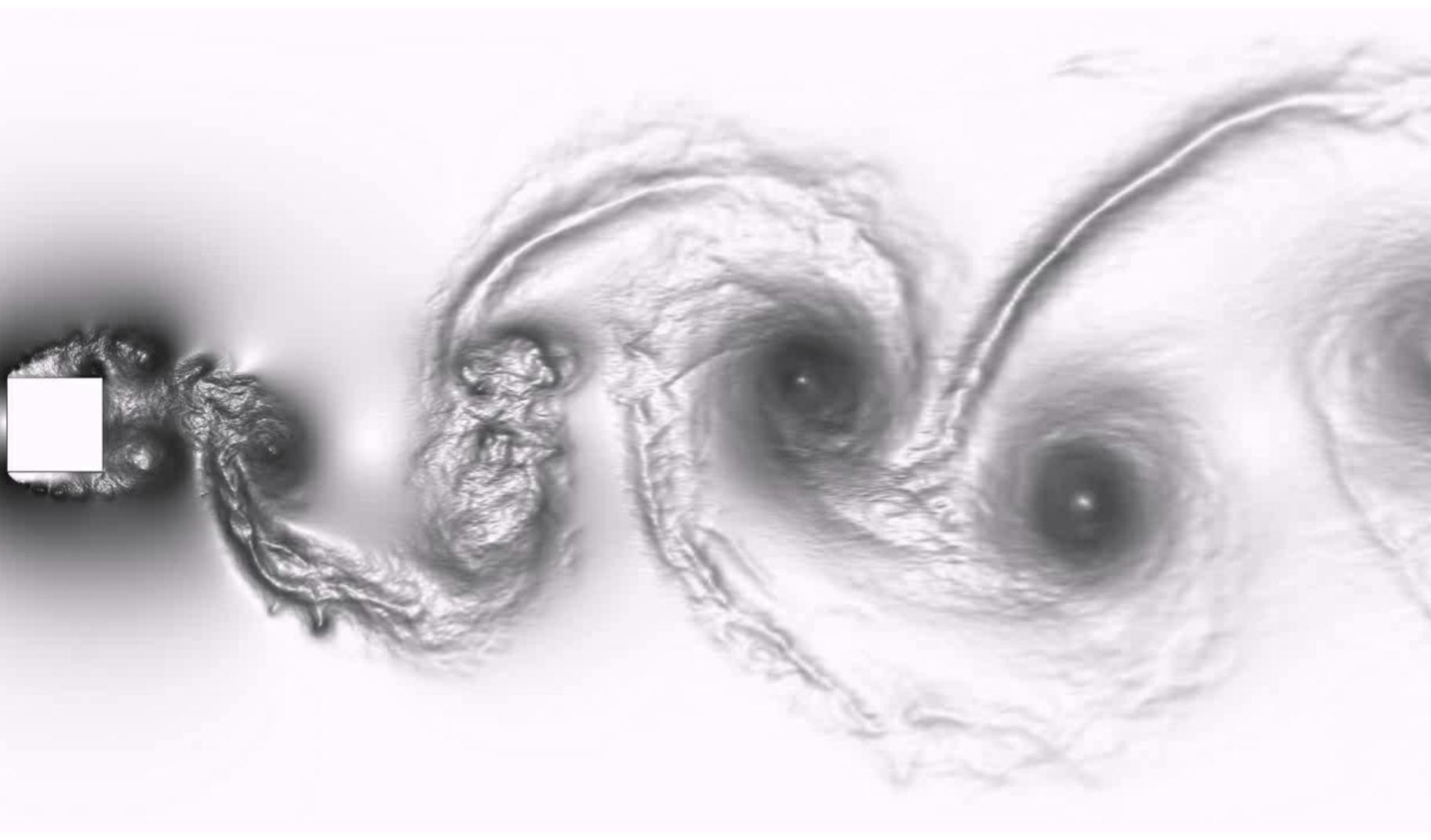
DNS течения вокруг квадратного цилиндра

- $Re = 22000$
- $Pr = 0.71$
- Сетка $1174 \times 1174 \times 216 \approx 300M$ узлов
- 784 CPU на суперкомпьютере MareNostrum
- Вычислительная стоимость $\sim 270K$ CPUh



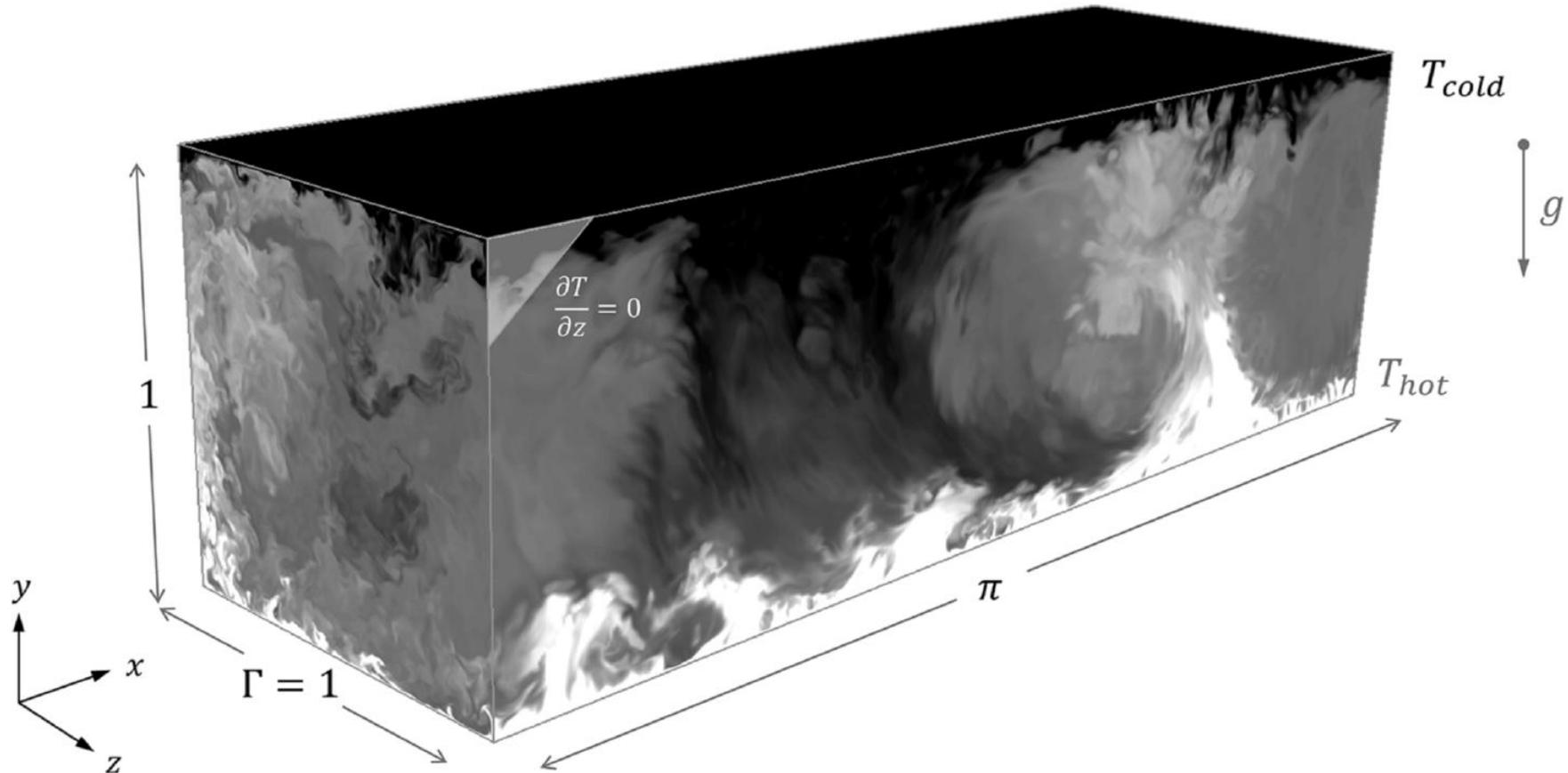
F.X.Trias, A. Gorobets, and A. Oliva. Turbulent flow around a square cylinder at Reynolds number 22000: a DNS study, Computers & Fluids, 2015, Volume 123, Pages 87–98

DNS течения вокруг квадратного цилиндра



DNS конвекции Рэля — Бенара

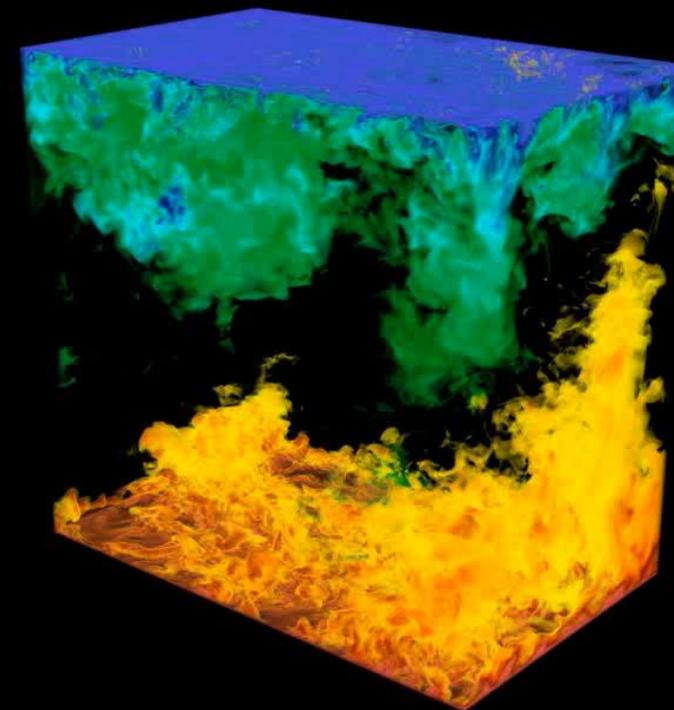
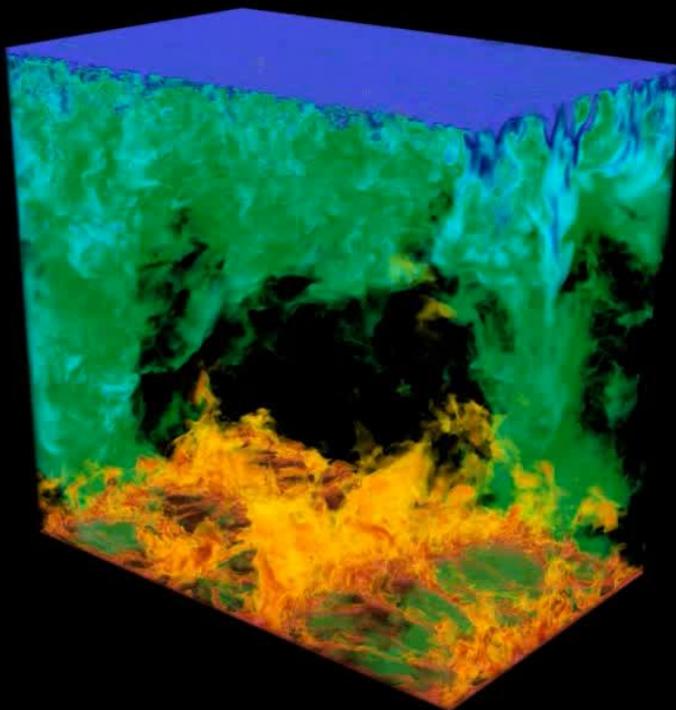
$Ra = 10^{10}$, 600M узлов,
суперкомпьютер MareNostrum



F.Dabbagh, F.X.Trias, A.Gorobets, and A.Oliva. A priori study of subgrid-scale features in turbulent Rayleigh-Benard convection. *Physics of Fluids*, 29, 105103 (2017); <https://doi.org/10.1063/1.5005842>

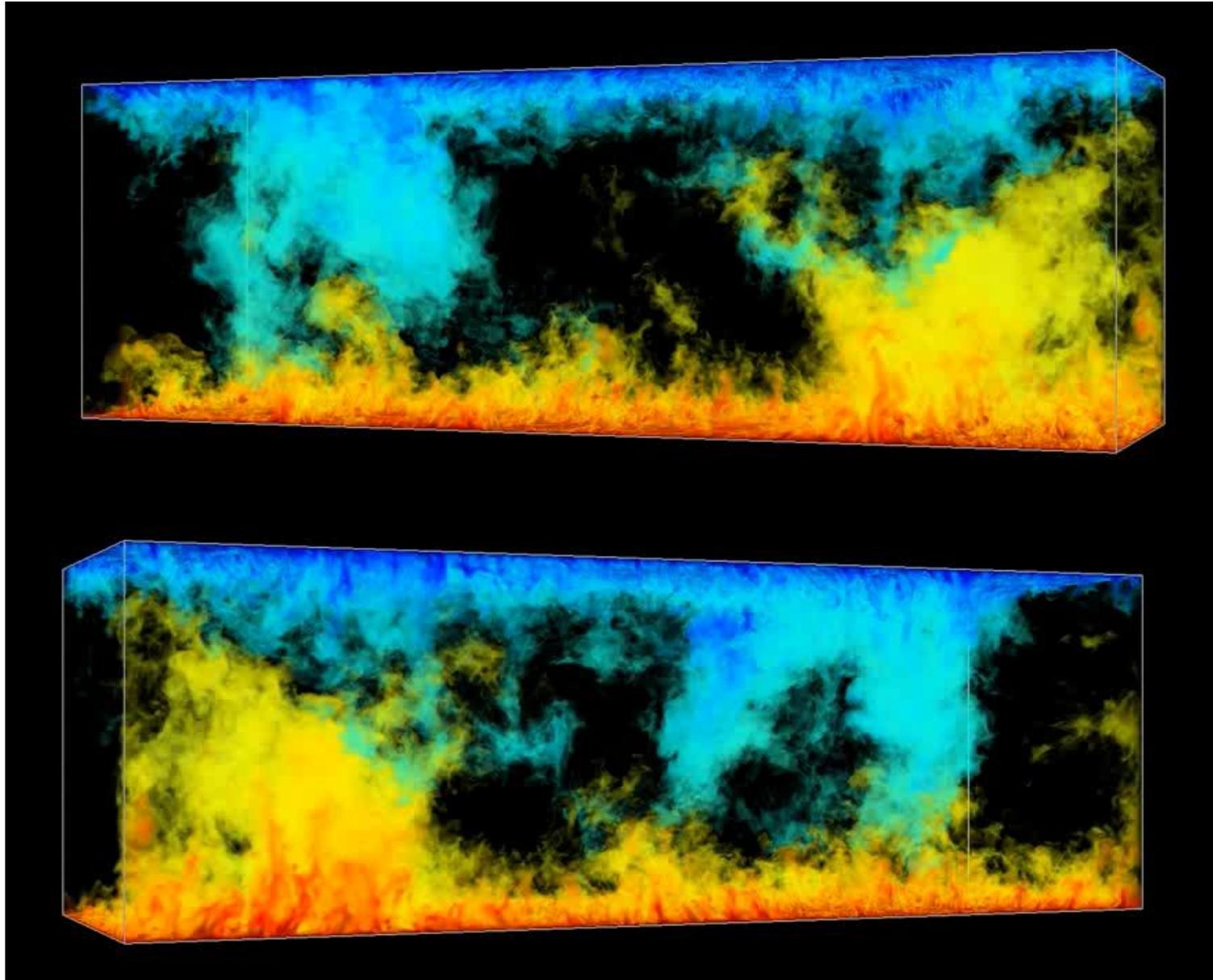
DNS конвекции Рэля — Бенара

$Ra = 10^{10}$, 600M узлов,
суперкомпьютер MareNostrum



F.Dabbagh, F.X.Trias, A.Gorobets, and A.Oliva. A priori study of subgrid-scale features in turbulent Rayleigh-Benard convection. *Physics of Fluids*, 29, 105103 (2017); <https://doi.org/10.1063/1.5005842>

DNS конвекции Рэля — Бенара



DNS конвекции Рэля — Бенара

