



Вычислительные алгоритмы и отказоустойчивость

М.В. Якововский
lira@imatmod.ru

Ближайшие перспективы

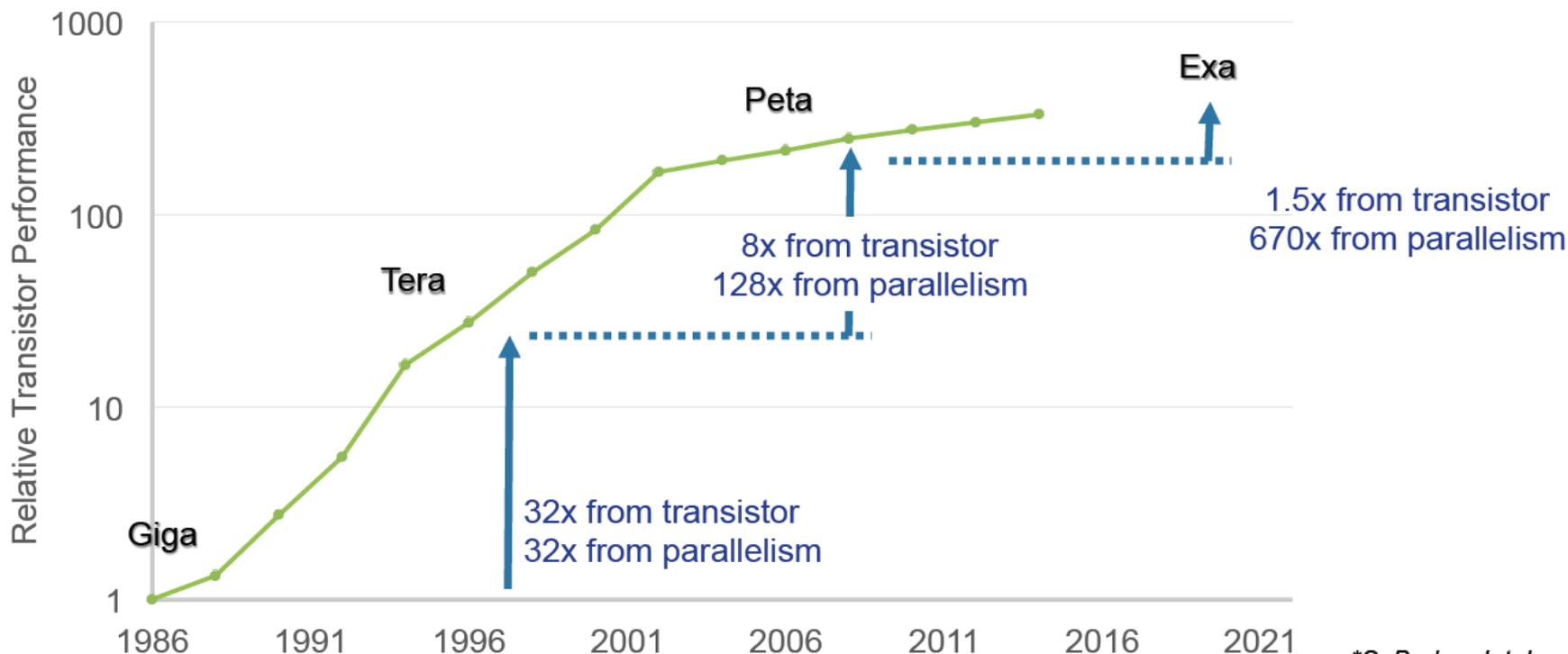
- Реальная необходимость высокопроизводительных вычислительных систем следующего поколения для решения задач:
 - нефтегазовые проблемы разведки и оптимизации добычи
 - экологические двигатели
 - ядерная энергетика и термоядерный синтез
 - фундаментальные проблемы астрофизики
- 2015 - Достаточно широкое использование PetaFLOPs вычислительных систем
- 2018...2020 - Производительность суперкомпьютеров 1 ExaFLOPs

July 31, 2016

Paul Messina, Argonne National Laboratory, ECP Director

A Path to Capable Exascale Computing

From Giga to Exa, via Tera & Peta*



*S. Borkar, Intel

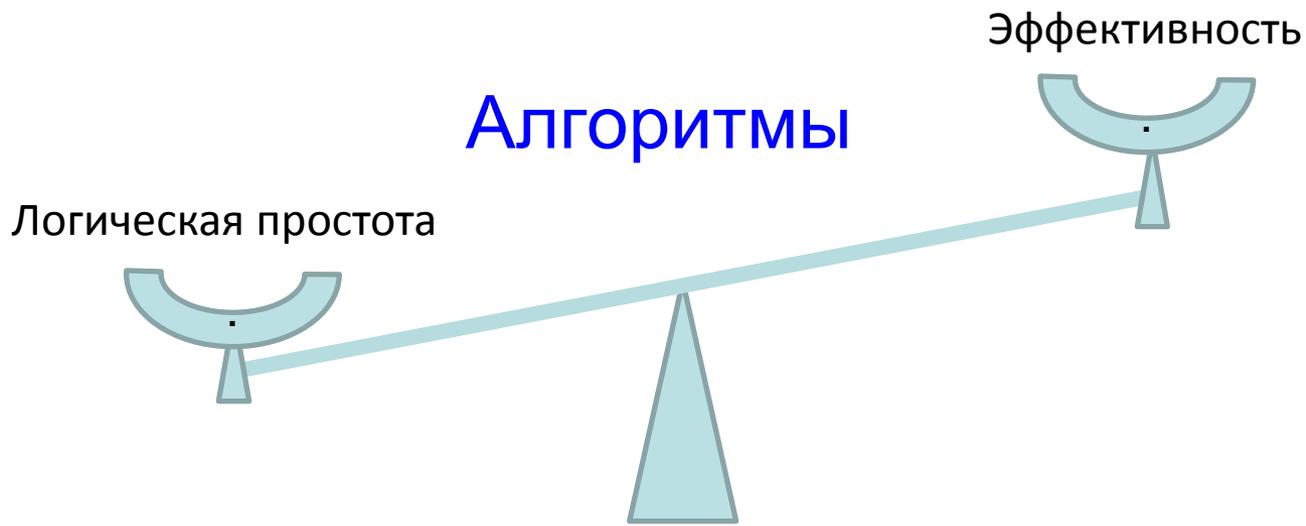
Performance from parallelism

July 31, 2016

Paul Messina, Argonne National Laboratory, ECP Director

A Path to Capable Exascale Computing

- ExaScale computer
 - Позволяет решить научную задачу в 50 раз быстрее (или сложнее), системы уровня 20 Pflops (Titan, Sequoia), или в 100 раз быстрее, чем Mira
 - Мощность порядка 20-30 МВт
 - Отказоустойчивость позволяет пользователю вмешиваться на чаще одного раза в неделю
 - Имеет стек программного обеспечения, который отвечает потребностям широкого спектра приложений и рабочих нагрузок



- Явные схемы позволяют создавать логически простые алгоритмы, но имеют строгие ограничения на дискретизацию по времени из условий устойчивости:

- для уравнений гиперболического типа условие устойчивости

$$\Delta t \leq h,$$

где Δt - шаг дискретизации по времени, h - шаг дискретизации по пространству

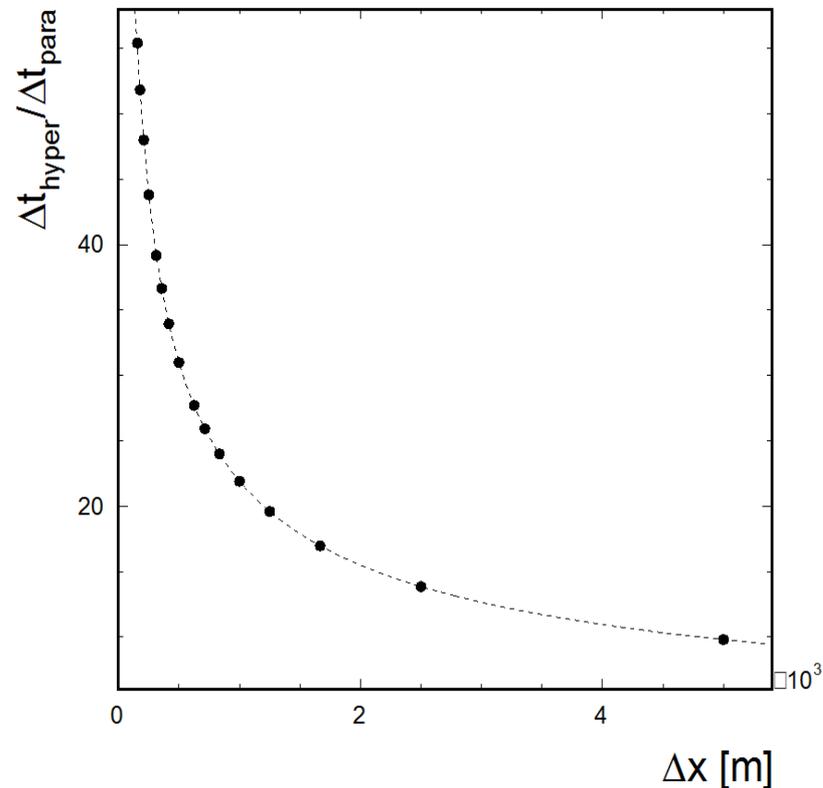
- для параболического типа уравнений условие устойчивости

$$\Delta t \leq h^2$$

Это условие практически исключает возможность использования высокого разрешения по пространству

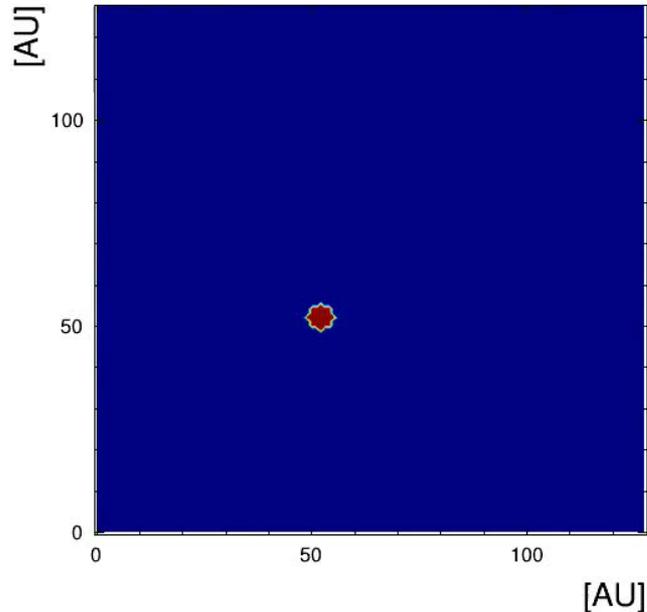
Уравнение Пуассона (гравитационный потенциал)

Отношение временной дискретизации параболического и гиперболического метода как функция дискретизации по пространству



Аккреция облака межзвездного газа на компактном астрономическом объекте

низкое разрешение



высокое разрешение



- Межзвездное облако 5 AU
- Плотность 0.8×10^{-11} кг/м³
- Скорость облака 300 м/с
- Импакт параметер 4-10 AU
- Компактный объект:
 - масса 10^{30} Кг
 - радиус 0.5 AU
- Температура пространства $T = 20$ К

Blue Waters system

Cray HLRS – Germany, Stuttgart

- Каждые 4.2 часа фиксируется отказ, требующий восстановления части системы
- Полный отказ системы каждые 160 часов

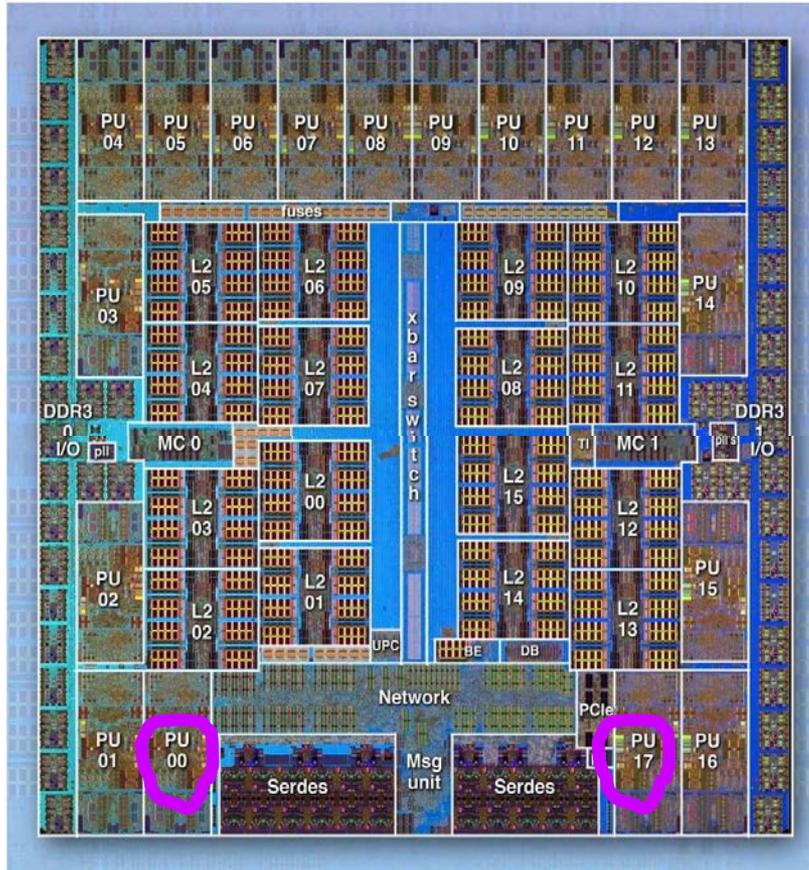


Di Martino, Catello, Zbigniew Kalbarczyk, Ravishankar K. Iyer, Fabio Baccanico, Joshi Fullop, and William Kramer. "Lessons learned from the analysis of system failures at petascale: The case of blue waters." *In Dependable Systems and Networks (DSN), 2014 44th Annual IEEE/IFIP International Conference on*, pp. 610-621. IEEE, 2014.

Статистика отказов

Системы	Количество ядер	Надежность
ASCI Q	8192	MTBF: 6.5 часов Источники аппаратные сбоев: устройство хранения, CPU, память
ASCI White	8192	MTBF: 5 часов (2001 год) и 40 часов (2003 год) Источники аппаратные сбоев: устройство хранения, CPU, внешние устройства
PSC Lemieux	3016	MTBI: 9.7 часов
Google	15000	20 перезагрузок/день 2-3% компьютеров заменяется ежегодно Источники аппаратных сбоев: устройства хранения, память

IBM PowerPC® A2 1.6 GHz, 16 ядер на процессор



Robert W. Wisniewski.

BlueGene/Q: Architecture,

CoDesign; Path to Exascale / Blue

Gene Supercomputer Research,

January 25, 2012

Однако ядер 18, а не 16

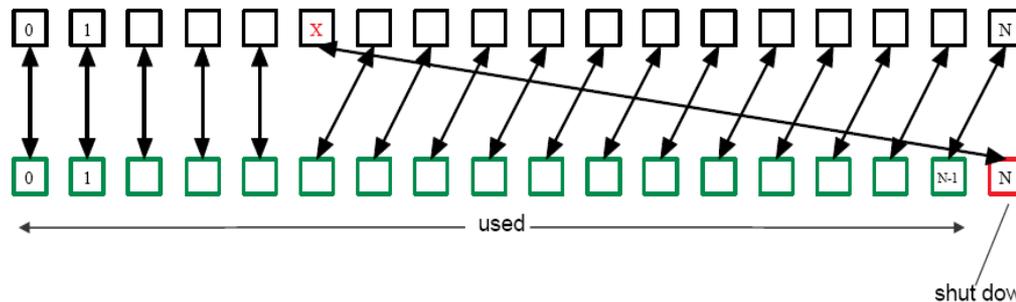
Одно – сервисное

Одно – запасное

В последних процессорах поддерживается горячая замена ядра

Physical Processor core IDs

Logical Processor core IDs



Время между отказами на экзафлопсных системах ~ 30 минут

Marc Snir, et al. Addressing failures in exascale computing. International Journal of High Performance Computing Applications, 28(2):129–173, May 2014

Частота аппаратных отказов будет возрастать

- Уменьшение размера транзистора делает его менее устойчивым к космической радиации
- Ёмкости меньшего размера содержат меньший заряд, - его проще изменить

Программное обеспечение становится сложнее и содержит больше ошибок

- Оборудование становится сложнее (неоднородные ядра, многоуровневая иерархия памяти, сложная топология объединения узлов) существенно усложняет программное обеспечение
- *Мультифизичность и многомасштабность* решаемых задач приводит к объединению большого числа программных модулей.
- Сокращение обменов, использование асинхронных взаимодействий, обеспечение защищённости от отказов оборудования приводит к созданию сложных прикладных кодов

Время создания контрольной точки ~ 30 минут

System from TOP 500	Max performance	Checkpoint time (minutes)
LLNL Zeus Lawrence Livermore National Laboratory	11 TeraFLOPS	26
LLNL BlueGene/L	500 TeraFLOPS	20
Argonne BlueGene/P	500 TeraFLOPS	30
LANL RoadRunner Los Alamos National Labs	1 PetaFLOPS	~ 20

Cappello F. 2009. Fault Tolerance in Petascale/ Exascale Systems: Current Knowledge, Challenges and Research Opportunities. International Journal of High Performance Computing Applications 23, 3, 212–226.



Уровни управления контрольными точками

- Системный уровень
 - Простота использования
- Уровень пользователя
 - Радикальное сокращение объёмов контрольных точек
 - Вместо рестарта всей системы - замена вычислительного узла
 - Хранение данных не только на локальных дисках HDDs но и в оперативной памяти



Доступные инструменты

Automatic (based on BLCR) system level checkpoint :

- MPICH, MVAPICH, OpenMPI

Semi-automatic, user level checkpoint :

- C³ - Cornell Checkpoint pre-Compiler, (Greg Bronevetsky, Daniel Marques, ...)
- ULFM(FT-MPI)

Egwutuoha, I.P. A survey of fault tolerance mechanisms and checkpoint/restart implementations for high performance computing systems. / I.P. Egwutuoha, D. Levy, B. Selic, S. Chen // The Journal of Supercomputing. — 2013. — Vol. 65, No.3. —P. 1302-1326.

Cappello, F. Fault tolerance in petascale/exascale systems: Current knowledge, challenges and research opportunities // International Journal of High Performance Computing Applications. — 2009. — Vol. 23, No. 3. — P. 212–226

ULFM –

User-Level Failure Mitigation

Версия MPI 3.1 не имеет механизмов управления и штатной работы при отказах

Стандарт ULFM предложен в качестве минимального, но достаточного интерфейса обеспечивающего восстановление MPI приложений, даже при отказах

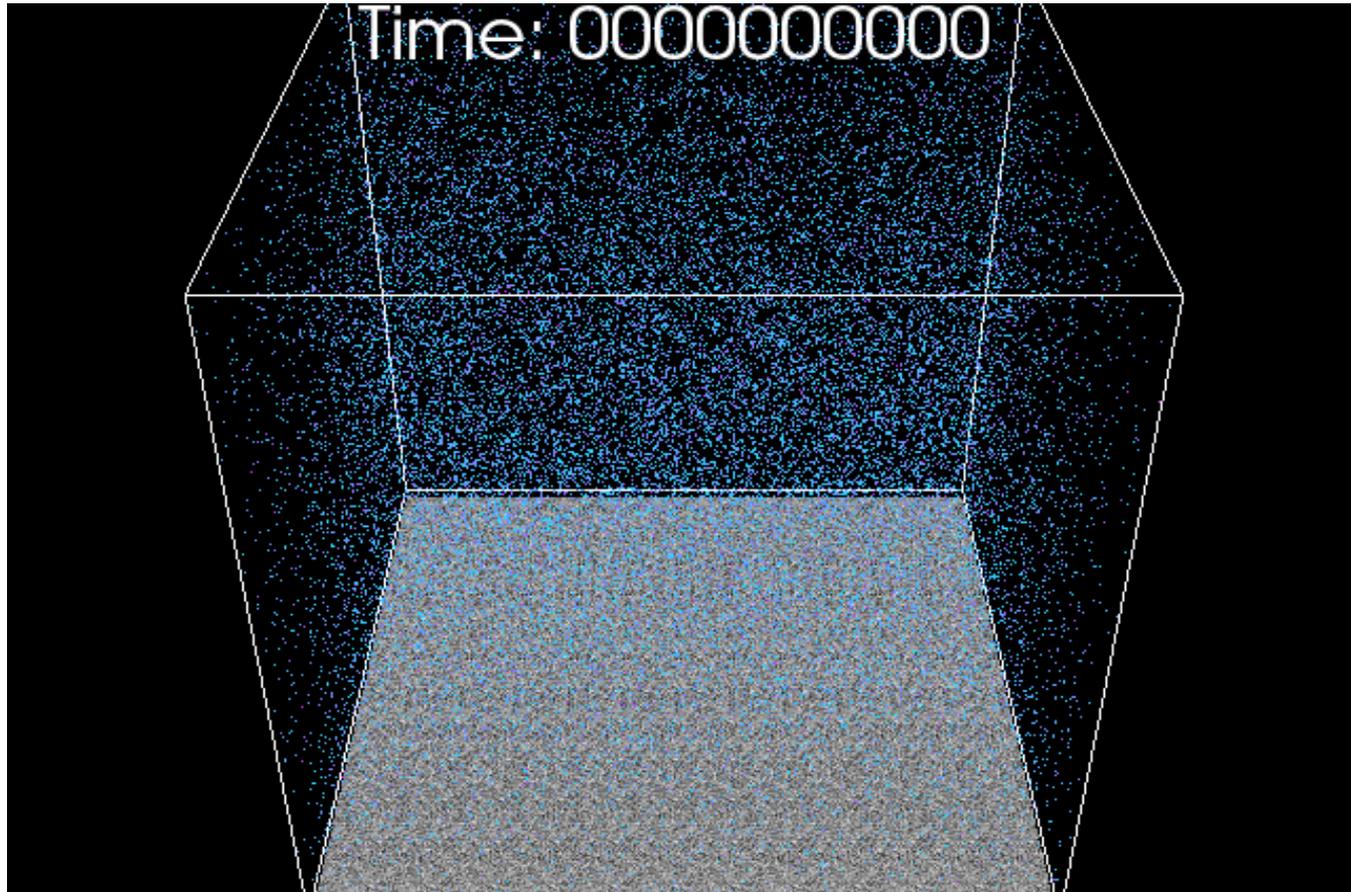
- **MPI_COMM_REVOKE**
- **MPI_COMM_SHRINK**
- **MPI_COMM_FAILURE_GET_ACKED**
- **MPI_COMM_FAILURE_ACK**
- **MPI_COMM_AGRE**

<http://fault-tolerance.org/>

ULMF is a part of new version of MPI (MPI 4.1)

Моделирования взаимодействия Ni-N₂

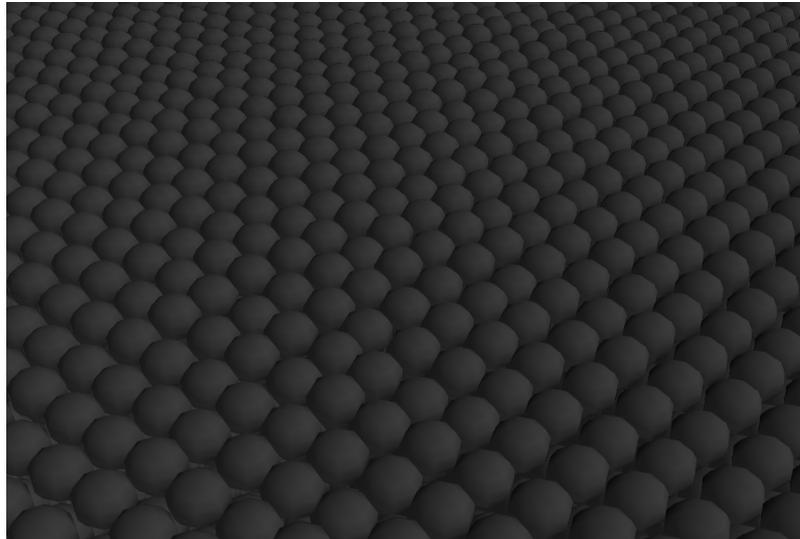
Size: 8 128 512 + 423 840 = 8 552 352 particles,
Temperature $T_{Ni} = 273.15$ K, $T_{N_2} = 273.15$ K



The problem is split into gas dynamics and molecular dynamics:
Flow and Particles

Пример расчета

- KIAM_JOB_CONTROL позволяет поддерживать долговременные вычисления с использованием различных вычислительных ресурсов
В качестве примера представлена задача МД моделирования и визуализации МД данных.

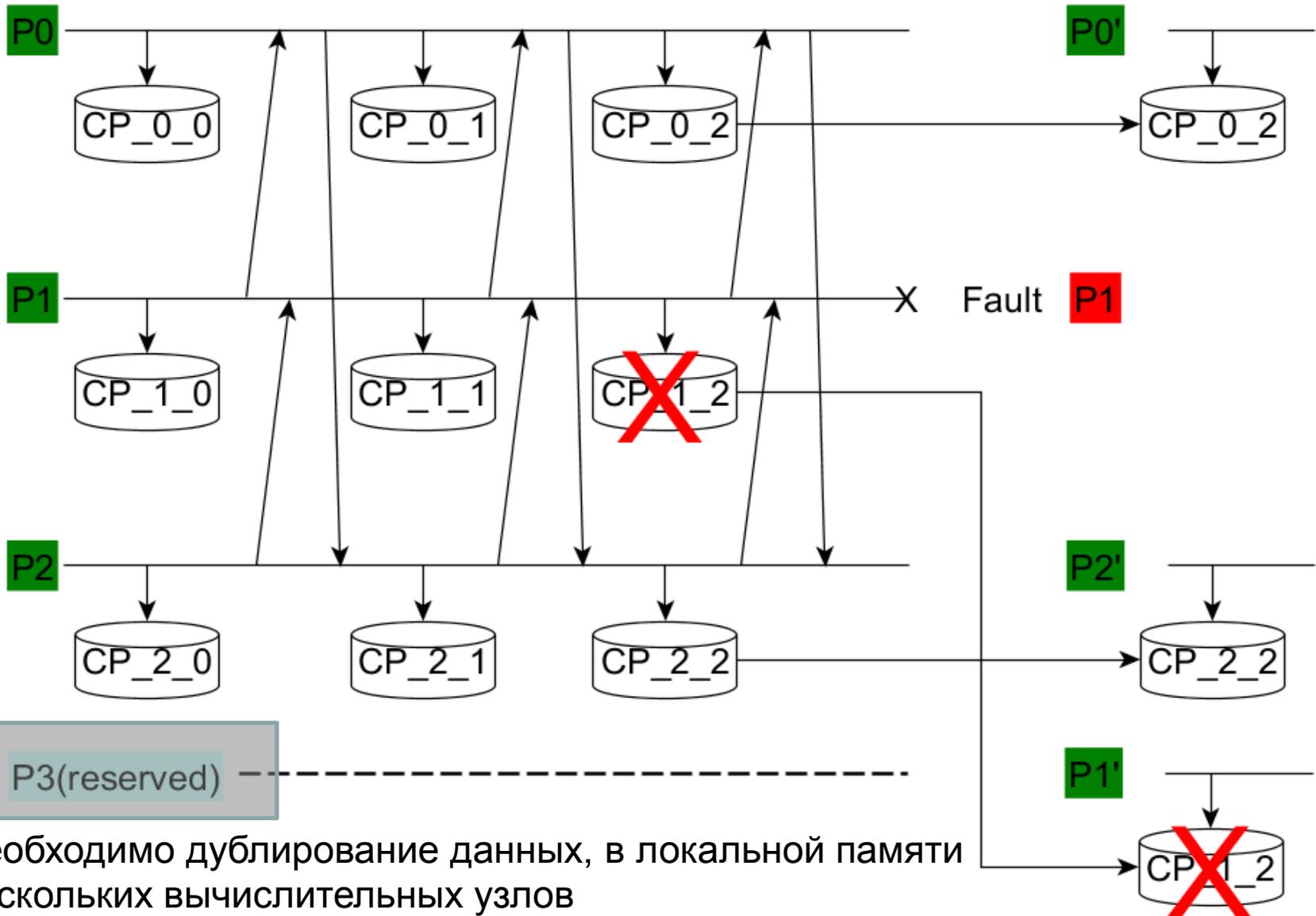




НРС ВЫЗОВ

- Разработка принципов управления контрольными точками, при которых время накладных расходов меньше чем MTBF
- Разработка алгоритмов, дающих возможность продолжать расчет даже при регулярных отказах части процессов

Обеспечение отказоустойчивости

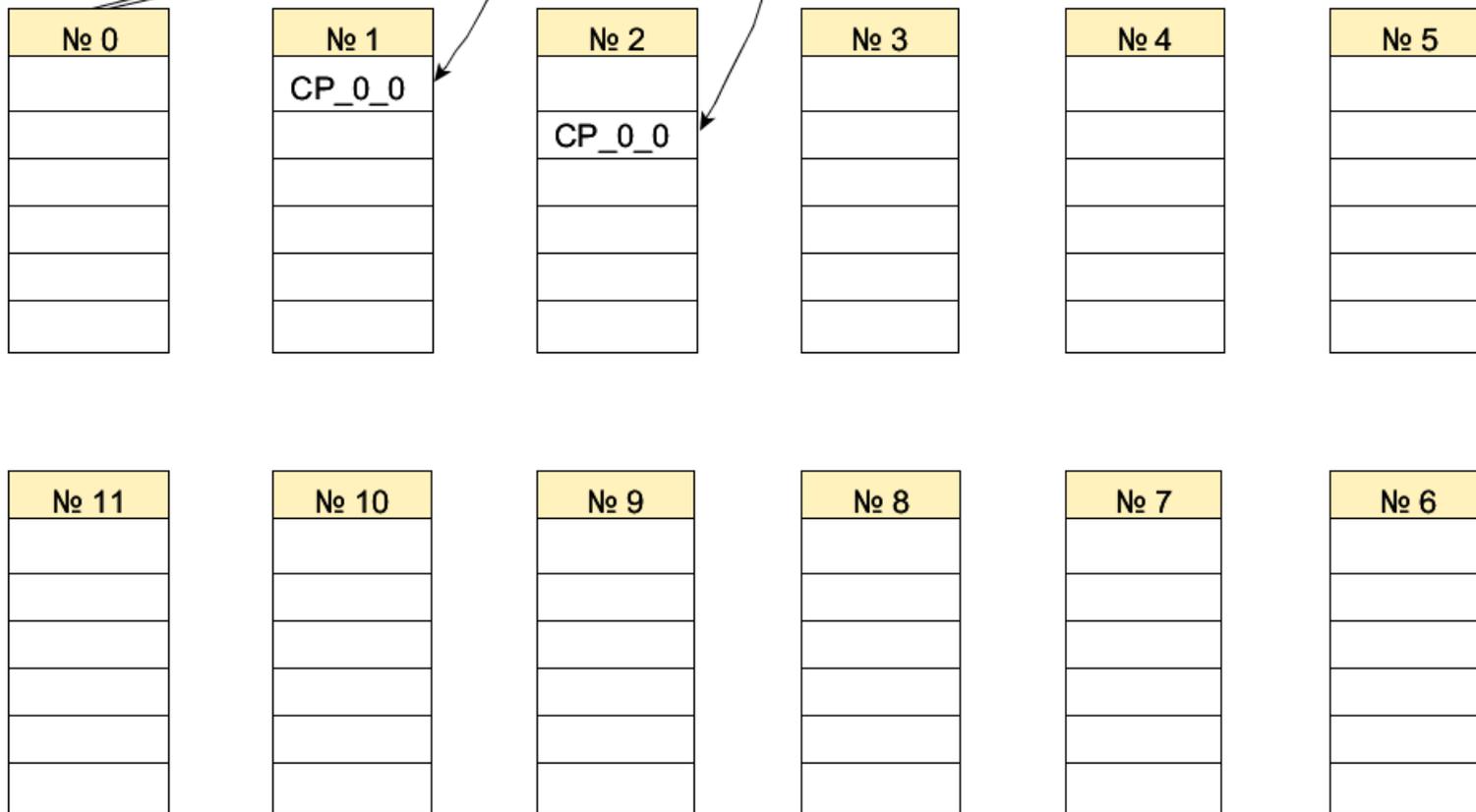


Необходимо дублирование данных, в локальной памяти нескольких вычислительных узлов

Пример стратегии дублирования

$N = 12$, $DF = 2$, $SD = 3$

CP_0_0

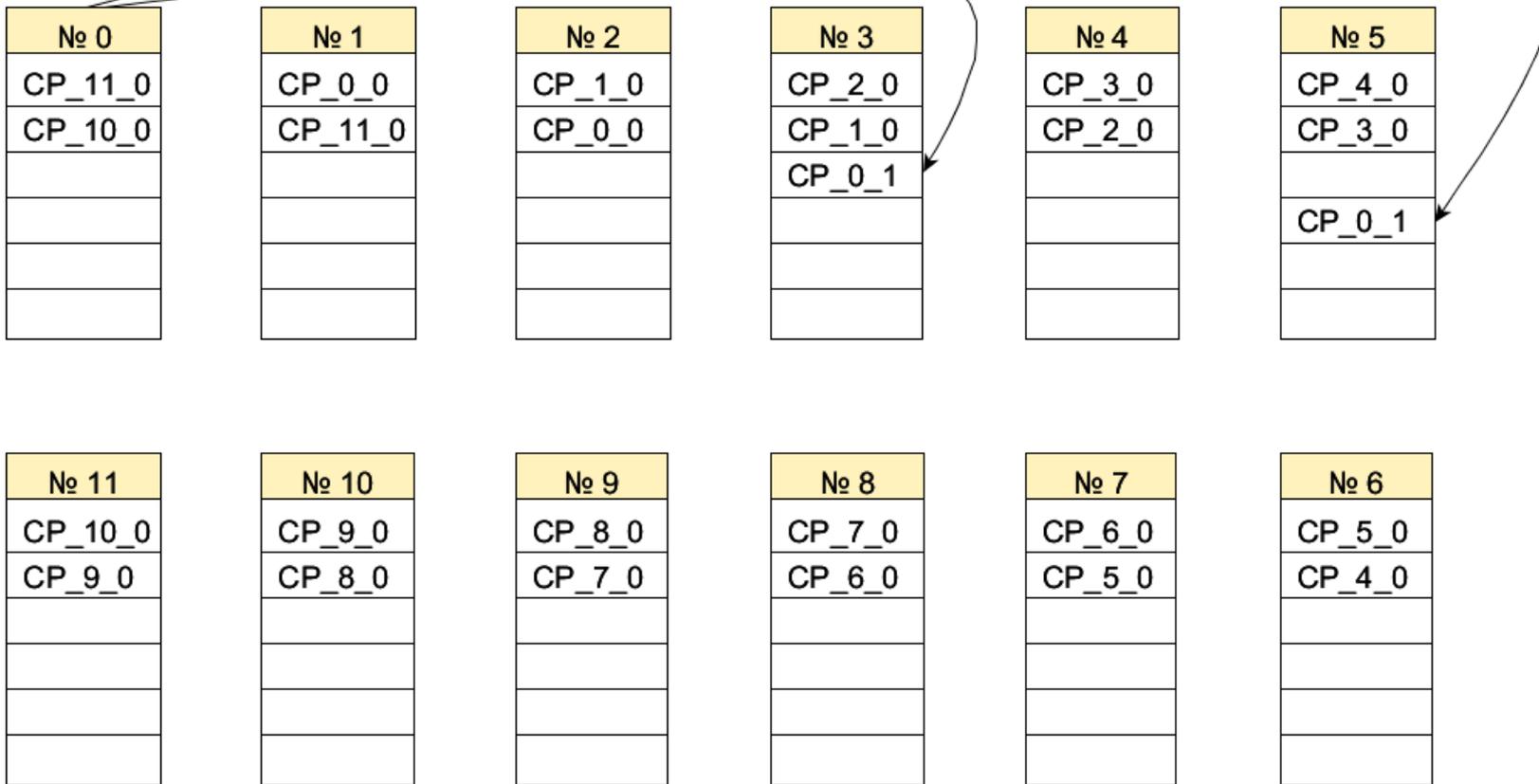


Пример стратегии дублирования

$$N = 12, DF = 2, SD = 3$$

CP_0_1

CP_0_0



Пример стратегии дублирования

N – 12 DF – 2 SD – 3

№ 0
CP_11_0
CP_10_0
CP_9_1
CP_7_1
CP_6_2
CP_2_2

№ 1
CP_0_0
CP_11_0
CP_10_1
CP_8_1
CP_7_2
CP_3_2

№ 2
CP_1_0
CP_0_0
CP_11_1
CP_9_1
CP_8_2
CP_4_2

№ 3
CP_2_0
CP_1_0
CP_0_1
CP_10_1
CP_9_2
CP_5_2

№ 4
CP_3_0
CP_2_0
CP_1_1
CP_11_1
CP_10_2
CP_6_2

№ 5
CP_4_0
CP_3_0
CP_2_1
CP_0_1
CP_11_2
CP_7_2

№ 11
CP_10_0
CP_9_0
CP_8_1
CP_6_1
CP_5_2
CP_1_2

№ 10
CP_9_0
CP_8_0
CP_7_1
CP_5_1
CP_4_2
CP_0_2

№ 9
CP_8_0
CP_7_0
CP_6_1
CP_4_1
CP_3_2
CP_11_2

№ 8
CP_7_0
CP_6_0
CP_5_1
CP_3_1
CP_2_2
CP_10_2

№ 7
CP_6_0
CP_5_0
CP_4_1
CP_2_1
CP_1_2
CP_9_2

№ 6
CP_5_0
CP_4_0
CP_3_1
CP_1_1
CP_0_2
CP_8_2

Одномерное гиперболическое уравнение

$$\frac{\partial^2 \Phi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} = F(x, t)$$

Две характеристики $x - ct$ и $x + ct$,
определяющими область, влияющую на
решение $\Phi(x, t)$ в точке (x, t)

Геометрические размеры области на
момент $(t - \Delta t)$, определяющие $\Phi(x, t)$,

**Область ускоренного расчета при
парировании потери данных, вызванной
выходом из строя одного процессора**

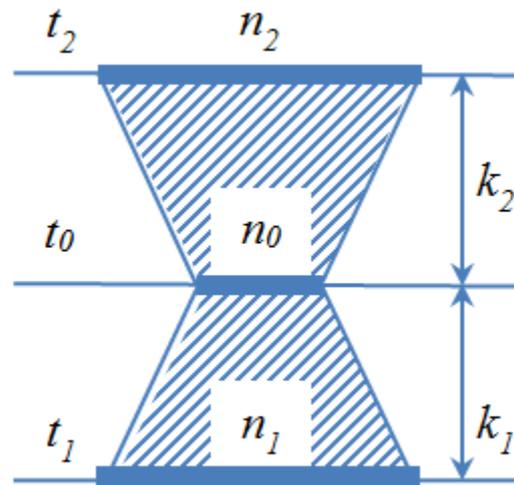
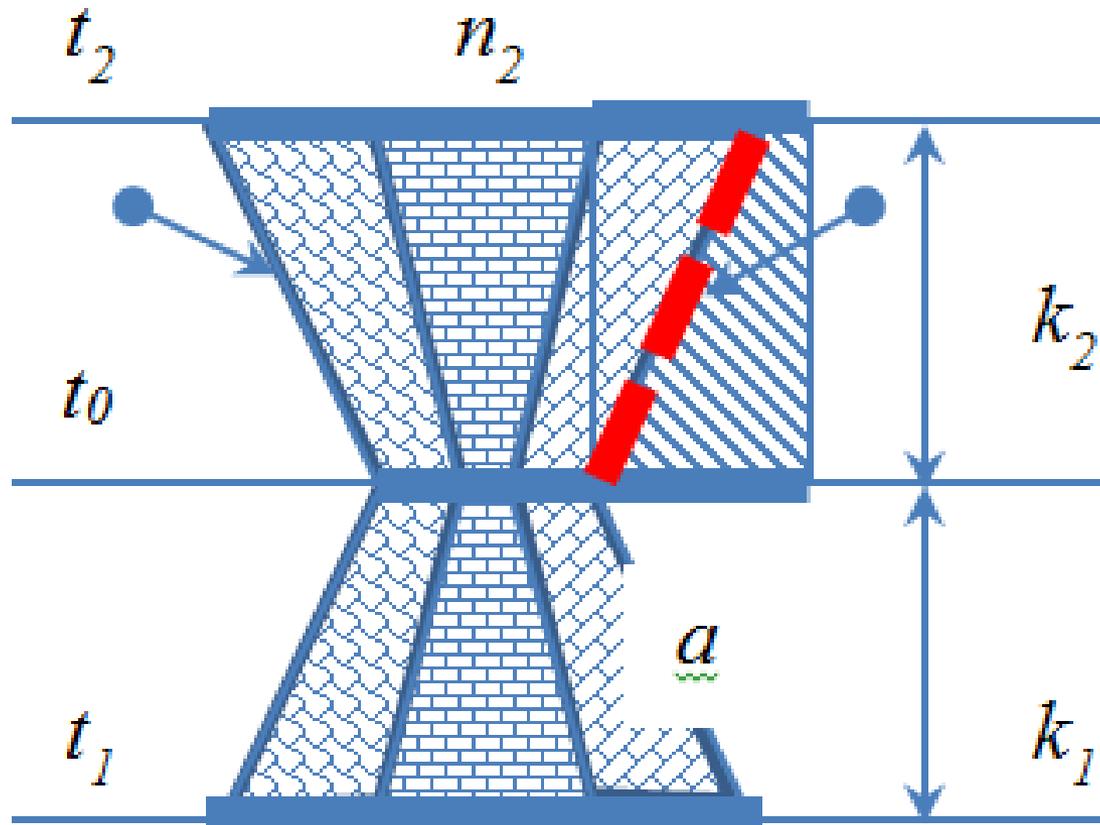
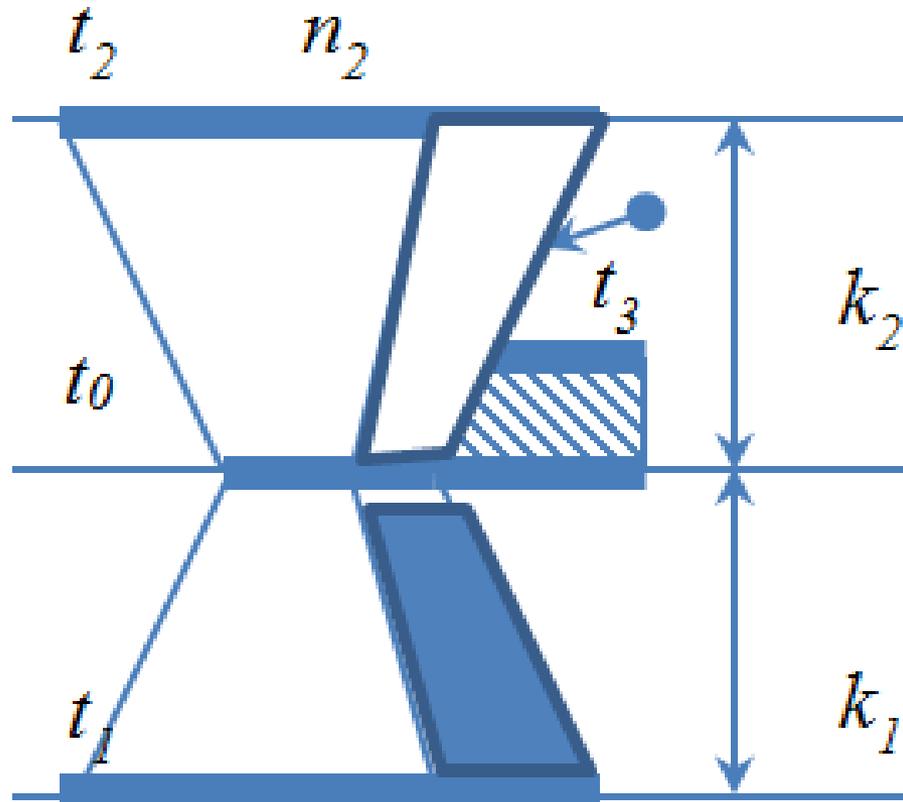


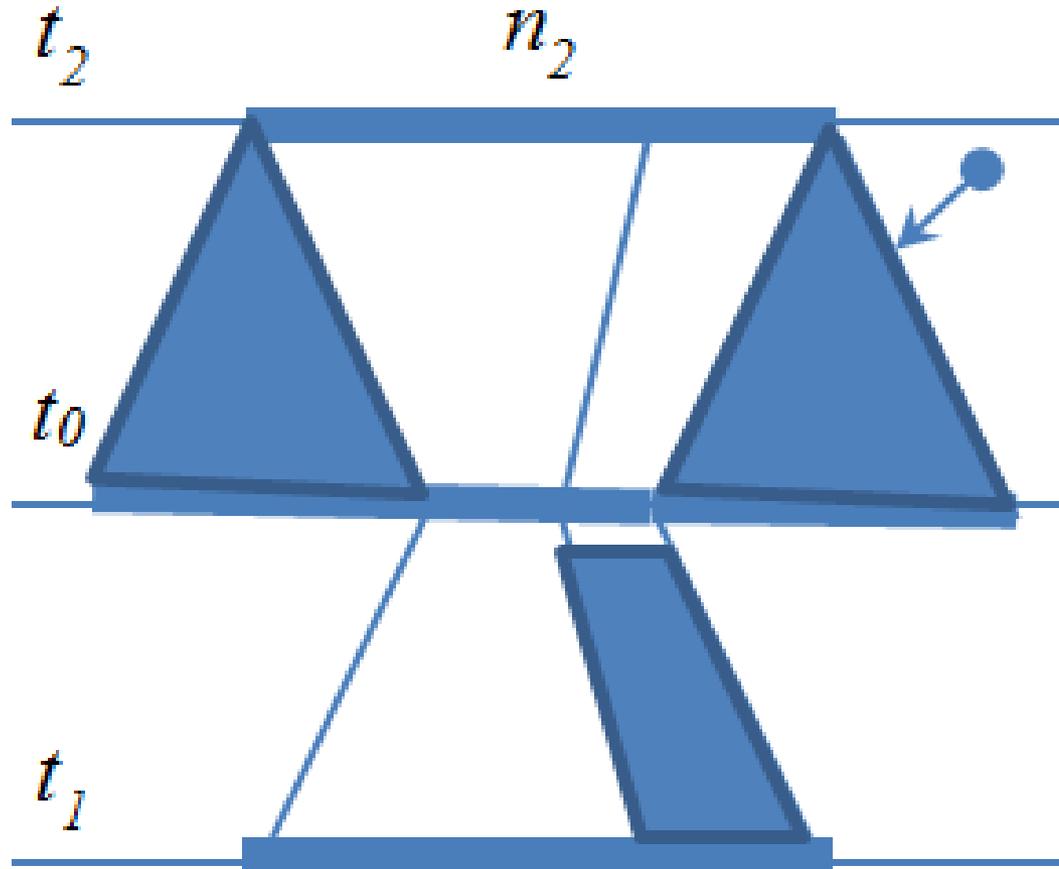
Схема замены одного неисправного процессора тремя резервными



Одновременный расчет двумя процессорами разных фрагментов области



Одновременный расчет двумя процессорами разных фрагментов области



Оценки числа дополнительных процессоров

$$p_{1D} \geq 2\alpha$$

$$p_{2D} \geq \frac{2}{3}(1 + \alpha + \alpha^2)$$

$$p_{3D} \geq \frac{1}{2}(1 + \alpha + \alpha^2 + \alpha^3)$$

$$\gamma = \frac{c\Delta t}{h}$$

$$\alpha = 1 + 2\gamma \frac{k_1}{n_0}$$

Заключение

- Необходима поддержка разработки алгоритмов и прикладного математического обеспечения, адаптируемого к архитектуре систем с экстремально параллелизмом и объемом данных и их использования для решения пилотных задач:
 - наиболее сложных в решении
 - обладающих особой важностью для развития страны
- Их создание обеспечит возможность эффективного роста практически во всех областях:
 - фундаментальная наука
 - промышленность
 - государственное и корпоративное управление
- Инструментом поддержки должна стать общегосударственная программа фундаментальных исследований, создания алгоритмов и математического обеспечения для систем сверхвысокой производительности, работы с большими данными и решения пилотных задач.

Литература

- *Четверушкин Б. Н., Якобовский М. В.* Вычислительные алгоритмы и отказоустойчивость гиперэкзафлопсных вычислительных систем // *Доклады Академии наук.* — 2017. — Т. 472, № 1. — С. 1–5. [[DOI](#)]
- *Бондаренко А. А., Якобовский М. В.* Моделирование отказов в высокопроизводительных вычислительных системах в рамках стандарта mpi и его расширения ulfm, Вестник Южно-Уральского государственного университета // *Вестник Южно-Уральского государственного университета. Серия "Вычислительная математика и информатика".* — 2015. — Т. 4, № 3. — С. 5–12.
- *Бондаренко А. А., Якобовский М. В.* Обеспечение отказоустойчивости высокопроизводительных вычислений с помощью локальных контрольных точек // *Вестник Южно-Уральского государственного университета. Серия "Вычислительная математика и информатика".* — 2014. — Т. 3, № 3. — С. 20–36.
- Cappello, F. Toward Exascale Resilience: 2014 update / F. Cappello, A. Geist, W. Gropp, S. Kale, B. Kramer, M. Snir // *Supercomputing frontiers and innovations.* — 2014. — Vol. 1, No. 1. — P. 1–28. DOI: 10.14529/jsfi140101.
- Bland, W. Post-failure recovery of MPI communication capability: Design and rationale / W. Bland, A. Bouteiller, T. Hérault, G. Bosilca, J. Dongarra // *International Journal of High Performance Computing Applications.* — 2013. — Vol. 27, No. 3. — P. 244–254. DOI: 10.1177/1094342013488238.
- ICL Fault Tolerance URL: <http://fault-tolerance.org/ulfm/ulfm-specification> (дата обращения: 01.03.2015).

Контакты

Якобовский М.В.

чл.-корр. РАН, д.ф.-м.н., проф.,

зав. сектором

«Программного обеспечения многопроцессорных систем и
вычислительных сетей»

Института прикладной математики им. М.В.Келдыша Российской
академии наук

[mail: lira@imamod.ru](mailto:lira@imamod.ru)

web: <http://lira.imamod.ru>