



РОССИЙСКИЙ ТЕХНОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ МИРЭА

ЭЛИТНАЯ
ПОДГОТОВКА



СУПЕРКОМПЬЮТЕРЫ

Занятие 05. Использование параллельного
программирования для решения проблем глобального
масштаба (задачи всепланетного уровня
GRAND CHALLENGES)

*Баканов Валерий Михайлович, д.т.н., профессор
915-053-5469, 499-122-1328,
e881e@mail.ru, http://vbakanov.ru/left_1.htm*

Кафедра КБ-5 “Аппаратное, программное и математическое обеспечение вычислительных систем”

Суперкомпьютеры в системах параллельных вычислений

Несколько определений суперкомпьютера (включая не совсем серьёзные):

Любой компьютер, весящий более тонны (1989, Гордон Белл и Дон Нельсон).

Любой компьютер, который сделал Сэймур Крэй (эпоха компьютеров фирмы Cray Research завершилась в середине 90-х г.г. со смертью основателя фирмы).

Десять (двадцать, сто, пятьсот...) наиболее мощных по производительности ЭВМ в мире или конкретной стране (идея воплощена в списках Top500.org и Top50.org).

Компьютер с производительностью более XXXYYY флопс (флопс – floating operations per second - число операций над числами с плавающей точкой в секунду; метод используется в современности при испытаниях производительности по системе LinPACK – решение систем линейных алгебраических уравнений (СЛАУ) большой размерности прямым методом (с фиксированным числом операций) Гаусса.

МЕХАНÉТИС (шутл.) – профессионально-психическое заболевание. Анамнез – пациент искренне убеждён, что решение задачи, постановку которой он (обычно) **не может даже точно сформулировать**, легко получить, если в его распоряжении появится суперкомпьютер. ОЧЕНЬ ЗАРАЗНО – передаётся воздушно-капельным путём при разговоре и чихании (особенно через InterNet); случаи передачи половым и бытовым путями не фиксированы. Прогноз: благоприятный при своевременно начатом лечении интенсивными методами показа истинной ценности пациента в данной области знаний (при этом со стороны последнего обычным является крайне агрессивное поведение с медленно затухающими эксцессами). **МЕХАНÉТИС** не возникает “на пустом месте” – известны случаи, когда при использовании суперкомпьютера наблюдалось не повышение, а резкое снижение быстродействия (причиной обычно является явление сетевой деградации вычислений)

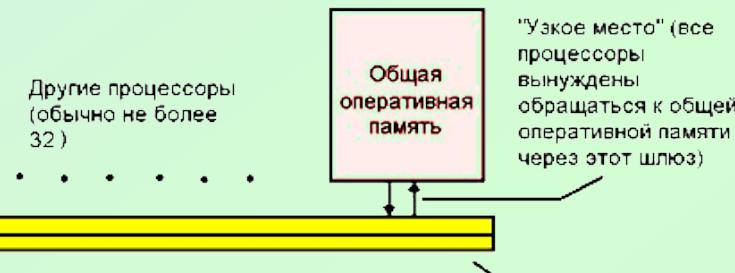
Основные архитектуры суперкомпьютерных вычислительных систем



a) SMP (Symmetric Multi Processors) - архитектура многопроцессорных вычислительных систем с общей памятью

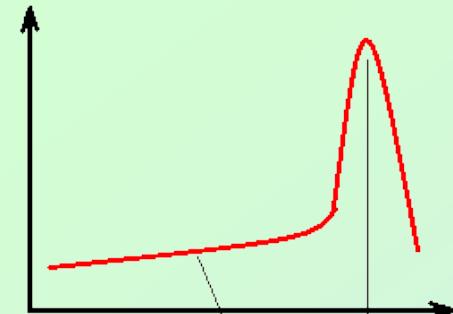


б) MPP (Massive Parallel Processing) - архитектура многокомпьютерных вычислительных систем с распределённой памятью



"Узкое место" (все
процессоры
вынуждены
обращаться к общей
оперативной памяти
через этот шлюз)

Коммуникационная
среда (обычно
быстро действующая
низколатентная
компьютерная сеть)



Траффик (интенсивность обмена
данными) коммуникационной
сети (качественная картина)



Каким образом можно сравнивать производительность компьютеров?

Производительность (быстродействие) компьютера – *число производимых им операций в единицу времени (обычно в 1 сек)*. Проблема – КАКИХ ОПЕРАЦИЙ? Обычная ЭВМ производит в единицу времени ОЧЕНЬ РАЗНОЕ количество операций – напр., арифметических операций с фиксированной запятой БОЛЬШЕ, чем с плавающей и ещё меньше “сложных” (напр., вычисление квадратного корня или тригонометрии). Часто используются т.н. **смёси Гибсона**.

Различные области применения ЭВМ отличаются преобладанием определенных операций - бухгалтерского счёта обычно достаточно операций с фиксированной точкой, для математического моделирования обязательно наличие операций с плавающей точкой и большого диапазона представления чисел.

Т.к. изначально супер-ЭВМ предназначались для использования в основном для решения сложных задач физики и математики, быстродействие принято определять в числе *арифметических операций с плавающей точкой двойной точности (double float)* в секунду (**flops, флопс**) на задачах определённого типа (решение СЛАУ большой размерности *прямым методом* Гаусса – тест LinPACK).

Почему выбран *прямой (не итерационный)* метод Гаусса? Во-первых, огромное количество научных задач (напр., *МКЭ – метод конечных элементов*) сводится к вычислительным процедурам линейной алгебры (операциям с матрицами); во-вторых – *при прямом (не итерационном) методе число операций не зависит от исходных данных*; в данном случае число операций = $2x(N^3/3+N^2) \approx O(N^3)$, где N – размérность матриц. *Применяются и другие тесты производительности.*

Джек Донгарра – разработчик теста LinPACK



В настоящее время [Джек Донгарра](#) (Jack J.Dongarra, род. July 18, 1950) - профессор [университета Теннеси](#).

В начале 1990-х годов возникла необходимость получения сравнительных характеристик и метрик суперкомпьютеров. После недостаточно показательных рейтингов 1986-1992 г.г. возникла идея начать ежегодно сравнивать суперкомпьютеры при помощи единой методики. В мае 1993 г. Джек Донгарра принял участие в этом проекте со своим тестом [LinPACK](#) (решение плотной СЛАУ большой размерности методом LU-декомпозиции - [прямой метод Гаусса](#)) разработанным им в [Аргоннской Национальной Лаборатории](#)). С июня 1993 этот тест (и его параллельная версия HPL - [High-performance Linpack](#), 64-битные “плавающие” числа по формату IEEE) используется при получения (дважды в год) данных о производительности компьютеров для списка Top500.

В 2013 году Джек Донгарра предложил новый тест для ранжирования суперкомпьютеров по вычислительным мощностям (HPCG - [High Performance Conjugate Gradient](#)). В июне 2014 был опубликован первый рейтинг на базе HPCG, первое место в котором занял китайский [Tianhe-2](#) с результатом в 0,58 HPCG PFlops (при 55 теоретических пиковых).

Критика методики основана на том, что LinPACK излишне узкоспециализирован и соответствует далеко не всем типам задач (алгоритмов), решаемых на суперкомпьютерах. Альтернативами предлагаются синтетические (состоящих из нескольких разнородных тестов) методики, более точно (не только на решении СЛАУ) определяющие производительность (напр., включающий 7 тестов для замера производительность различных подсистем компьютера, тест [HPC Challenge Benchmark](#)) или иные тесты, напр., [Graph500](#), информирующий о возможностях обработки больших массивов данных (вместо концентрации на чисто вычислительных характеристиках, как в тесте HPL), один из классов задач теста Graph500 - алгоритмы на графах.

Видео с Джеком Донгарра - [здесь](#) (Zurich, Switzerland, Sept. 7, 2012), [тут](#) (ЮУрГУ, Россия, Apr. 8, 2012) и [здесь](#) (интервью в ЮУрГУ).

Почему тест LinPACK использует “плавающие” числа двойной точности (*double float*)?

При решении задач на супер-ЭВМ общее число арифметических действий огромно (даже для ПЭВМ с производительность 1 Gflops за час работы производится $10^9 \approx 4 \cdot 10^{12}$ операций; супер-ЭВМ петафлопсного уровня за час производит $10^{15} \approx 4 \cdot 10^{18}$ операций); при каждой арифметической операции в “плавающей” арифметике **появляется погрешность**. В большинстве случаев погрешности имеют разный знак и в процессе вычислений происходит их **взаимная частичная компенсация**.

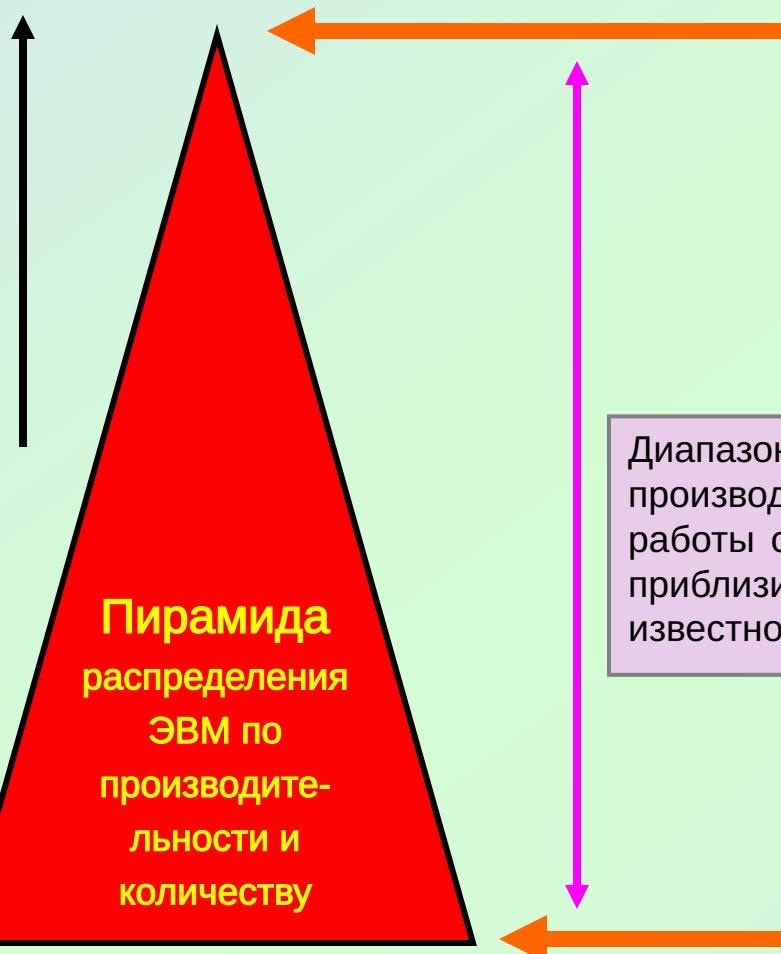
Грубую оценку вероятностной оценки реальной погрешности вычислений можно провести на основе постулирования нормального закона распределения погрешностей и использовании статистики Стьюдента; при этом для средне-квадратичного значения погрешности σ

имеем $s \gg O(b\sqrt{e})$, где e – число арифметических действий с округлением, β – значение младшего разряда числа (при представлении мантиссы числа 3-мя байтами получим $\beta=2^{-24}$).

Расчёты показывают, что за **час работы** ПЭВМ (при производительности $\sim 10^9$ опер/сек это около $4 \cdot 10^{12}$ операций) при использовании “плавающих” чисел одинарной точности (*float*, $\beta=2^{-24}$) имеем $\sigma=O(2^{-24}(2 \cdot 10^6)) \approx 0.12$, что заставляет усомниться в результате проведённых расчётов; при использовании чисел *double float* ($\beta=2^{-53}$) имеем $\sigma \approx 10^{-10}$ (это вполне удовлетворительно, конечно); для петафлопсной ЭВМ для последнего случая $\sigma \approx 10^{-7}$ (что, пожалуй, допустимо).

Каков диапазон LinPACK-производительности существующих ЭВМ ?

Производительность
растёт, количество
уменьшается

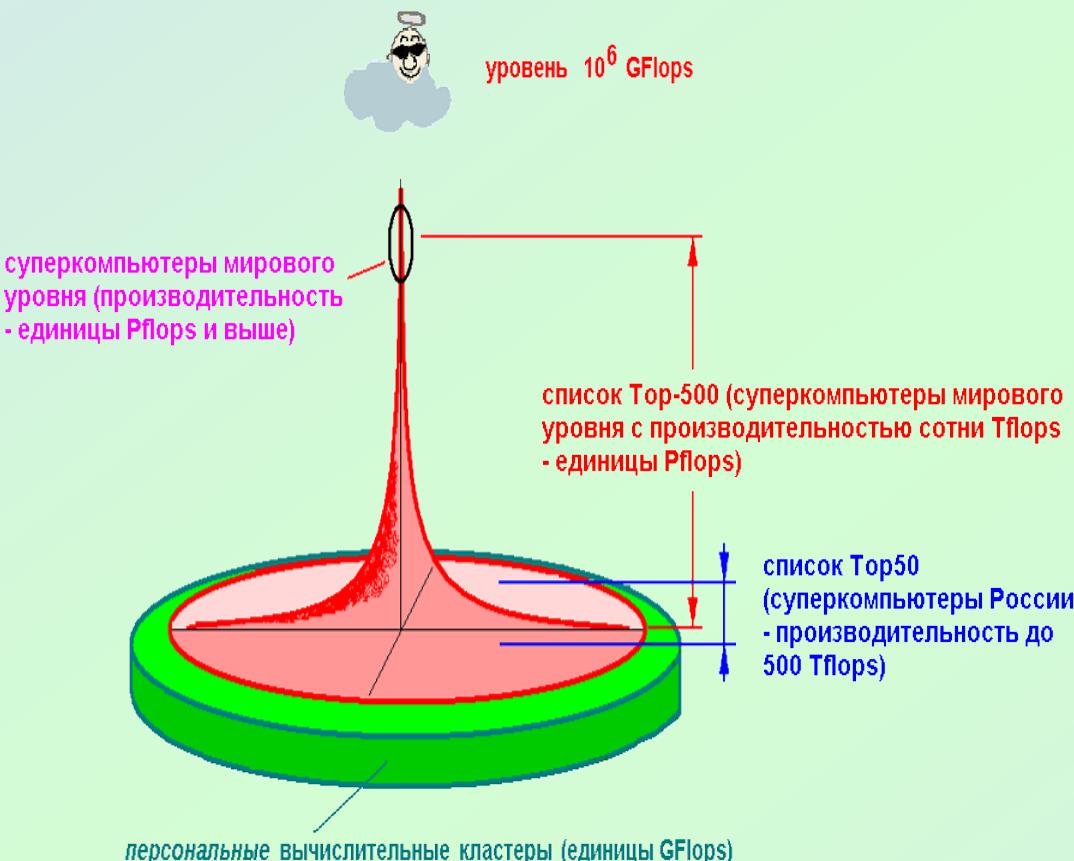


Уровень наиболее мощных супер-ЭВМ: производительность порядка сотен ПетаФлопс ($1 \text{ PFlops} = 10^{15} \text{ Flops}$); количество – единицы (сложность и стоимость столь высоки, что разработку могут позволить себе лишь наиболее технически развитые страны).

Диапазон по производительности между наиболее производительной и общедоступными ЭВМ равен $\approx 10^8$ (1 сек работы суперкомпьютера по числу выполненных операций равна приблизительно 10 годам работы ПЭВМ средней мощности... известно, что в году $\sim 3,6 \cdot 10^6$ сек).

Уровень персональных ЭВМ (ПЭВМ): производительность превышает 10^9 flops ($1 \text{ GFlops} = 1 \text{ ГигаФлопс}$); количество – сотни миллионов.

Распределение вычислительных мощностей 500 суперкомпьютеров в мире по данным июня 2021 г. списка Топ-500 (<http://top500.org>)



1-е место: Суперкомпьютер **Fugaku** - Суперкомпьютер Fugaku, A64FX 48C 2,2 ГГц, межсоединение Tofu D, Fujitsu RIKEN Center for Computational Science Japan. 7'630'848 ядер, 442 / 537 PFlops, 29'899 kW

2-е место: **Summit** - IBM Power System AC922, IBM POWER9 22C 3,07 ГГц, NVIDIA Volta GV100, Dual-Rail Mellanox EDR Infiniband, IBM DOE / SC / Национальная лаборатория Ок-Ридж, США. 2'414'592 ядер, 148 / 200 PFlops, 10'096 kW

3-е место: **Sierra** - IBM Power System AC922, IBM POWER9 22C 3,1 ГГц, NVIDIA Volta GV100, Dual-Rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE / NNSA/LLN США. 1'572'480 ядер, 94 / 125 PFlops, 7'438 kW

Россия:

61-е место: **Christofari** - NVIDIA DGX-2, Xeon Platinum 8168 24C 2.7GHz, Mellanox InfiniBand EDR, NVIDIA Tesla V100, Nvidia. SberCloud, Russia. 99'600 ядер, 7'7 / 8,7 PFlops

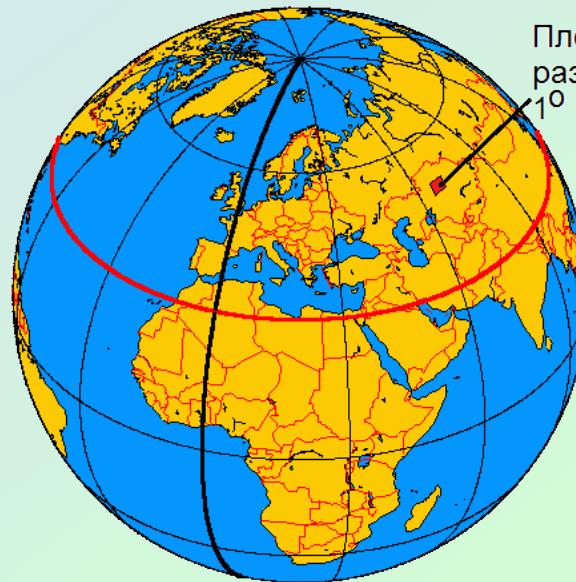
Современные задачи науки и техники, требующие для решения суперкомпьютерных мощностей

Области человеческой деятельности, где возникают задачи подобного рода, часто именуемые проблемами класса **GRAND CHALLENGES** или **ЗАДАЧАМИ БОЛЬШОГО ВЫЗОВА** (имеется в виду вызов окружающему миру) - фундаментальные научные или инженерные задачи с широкой областью применения, эффективное решение которых возможно исключительно с использованием мощных (суперкомпьютерных, неизбежно использующих параллельные вычисления) вычислительных ресурсов:

- Предсказания погоды, климата и глобальных изменений в атмосфере
- Генетика человека
- Моделирование и прогнозирование экономики
- Управляемый термоядерный синтез
- Науки о материалах
- Построение полупроводниковых приборов
- Сверхпроводимость
- Структурная биология
- Разработка фармацевтических препаратов
- Квантовая хромодинамика
- Астрономия
- Транспортные задачи большой размерности
- Гидро- и газодинамика
- Эффективность систем сгорания топлива
- Разведка нефти и газа
- Вычислительные задачи наук о мировом океане
- Распознавание и синтез речи, распознавание изображений

Нетрадиционное использование суперкомпьютеров в биржевой деятельности – концепция “**быстрых брокеров**” (суперпроизводительность позволяет удовлетворительно предсказывать короткосрочные тренды и осуществлять в единицу времени огромное число мелких транзакций; многие биржи на основании этого запрещают излишне частые запросы от одного и того же брокера).

Каково число арифметических действий необходимо для решения конкретной задачи уровня *GRAND CHALLENGES* – моделирование климата Земли?

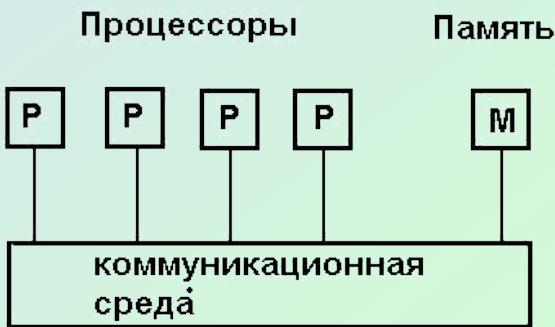


Климатическая система включает атмосферу, океан, сушу, криосферу и биоту. Климат – ансамбль состояний, через которые проходит климатическая система за достаточно большой промежуток времени. Климатическая модель – математическая модель, описывающая климатическую систему (в основе этой модели лежат уравнения динамики сплошной среды, уравнения равновесной термодинамики; в целом модель представляет систему трёхмерных нелинейных дифф. уравнений в частных производных). Обычно при численном решении таких моделей используются алгоритмы, основанные на методе дискретизации. Для данной задачи логично разбить область определения на элементы, размер основания которых определяется сеткой по поверхности Земли в виде квадрата в 1 градус (всего $360 \cdot 360 = 130 \cdot 10^3$ элементов).

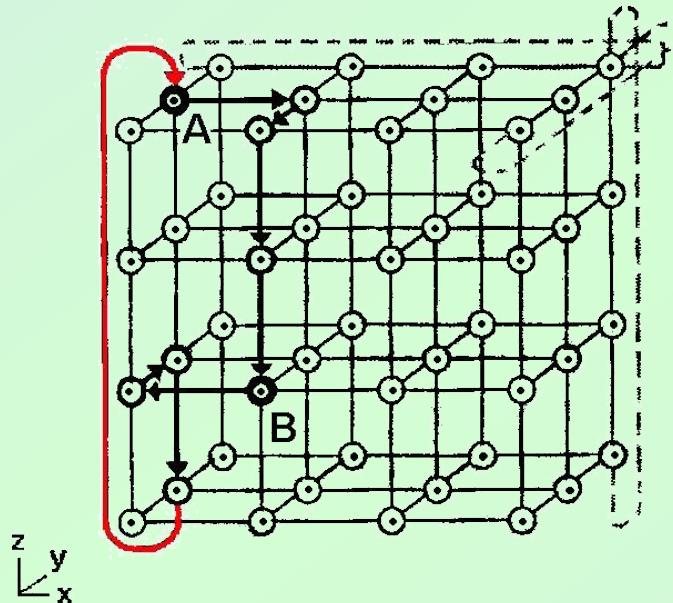
Т.к. необходима 3D-модель, используем 20 слоёв по высоте (от уровня Земли вверх в атмосферу), получаем $2,610^6$ объёмных элементов. Состояние каждого элемента описывается в этой модели примерно 10 компонентами (температура, плотность, компоненты скорости ветра etc); т.о. в любой фиксированный момент времени состояние атмосферы земли характеризуется ансамблем из $2,610^7$ (вещественных) чисел. Для достаточной точности моделирования необходимо пересчитывать ансамбли не реже 10 мин, для периода моделирования климата в 100 лет требуется рассчитать $1,310^{14}$ ансамблей (проверьте!). Реальность такова, что для вычисления каждого промежуточного результата необходимо выполнить $10^2 \div 10^3$ арифметических действий; т.о. всего в процессе одного численного эксперимента по моделированию климата требуется совершить порядка $10^{16} \div 10^{17}$ арифметических операций с плавающей запятой (обычно двойной точности).

При использовании ПЭВМ с производительностью 1 Gflops (10^9) время расчёта 0,3–3 года (что нереально); при использовании ЭВМ с производительностью 1 Tflops (10^{12}) время расчёта 3–30 час. Следует помнить, что для полного анализа проблемы необходимо проводить сотни/тысячи таких расчётов... при этом практически необходим вычислитель с производительностью порядка 1 Pflops (10^{15} операций в секунду) !

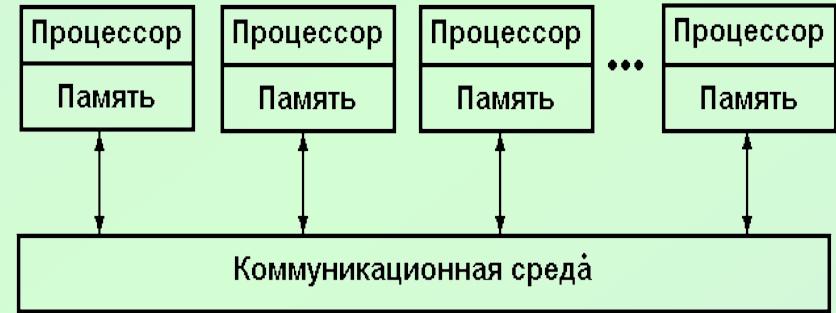
Основные архитектуры многопроцессорных (многокомпьютерных) вычислительных систем (МВС)



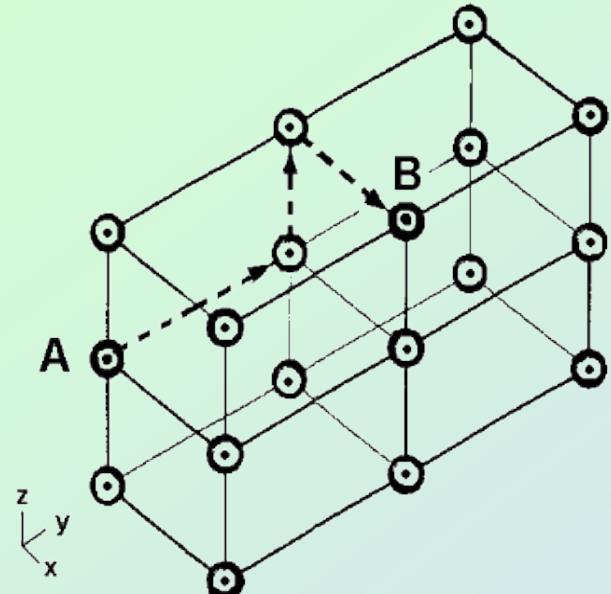
Архитектура SMP (*Symmetric Multi Processors*)



Объединение вычислительных узлов MPP-вычислителя с помощью 3D-решётки (красным цветом показан дополнительный маршрут, превращающий 3D-решётку в 3D-тор)

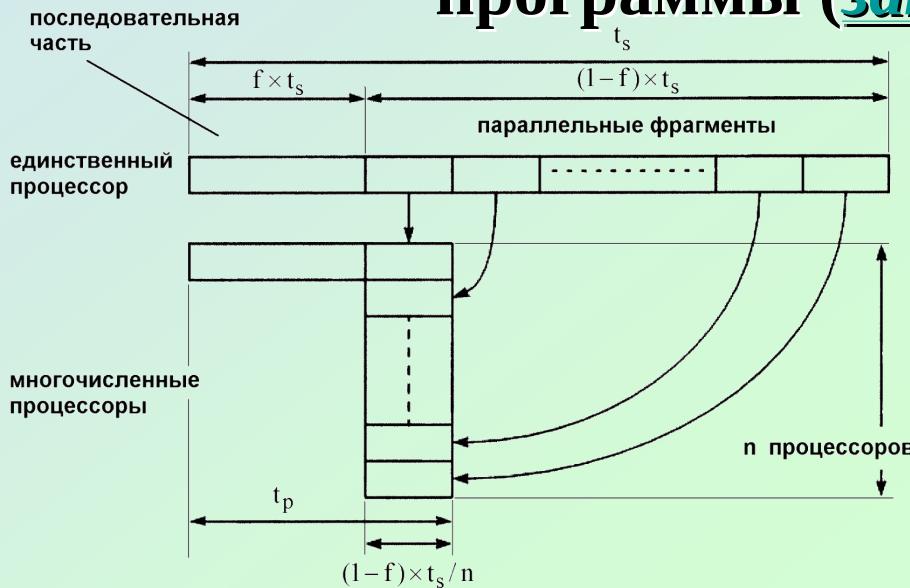


Архитектура MPP (*Massive Parallel Processing*)



Нахождение минимального пути между узлами в 3D-решётке

Ускорение вычислений на МВС от количества процессоров и степени распараллеливания программы (закон Амдаля)



Если в программе доля последовательных операций f (тогда часть удачно распараллеленных операций $1-f$), то во сколько раз ускорится выполнения такой программы на p процессорах (по отношению к последовательному выполнению на одном процессоре)?

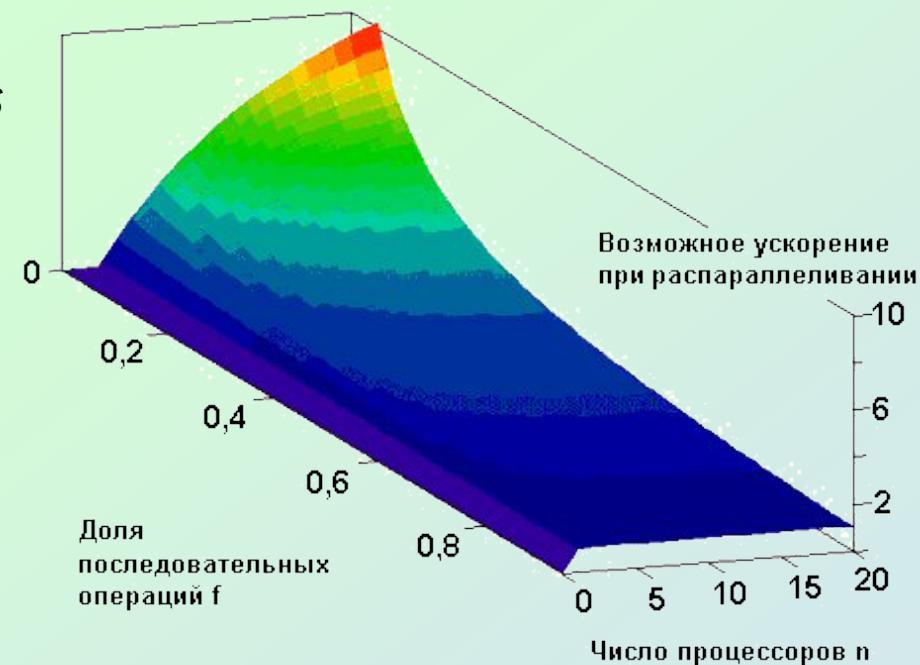
Закон Амдаля (Gene Amdahl, 1967) утверждает, что ускорение (убыстрение) вычислений S ограничено сверху:

$$S \leq \frac{1}{f + \frac{\alpha(1-f)}{p}}$$

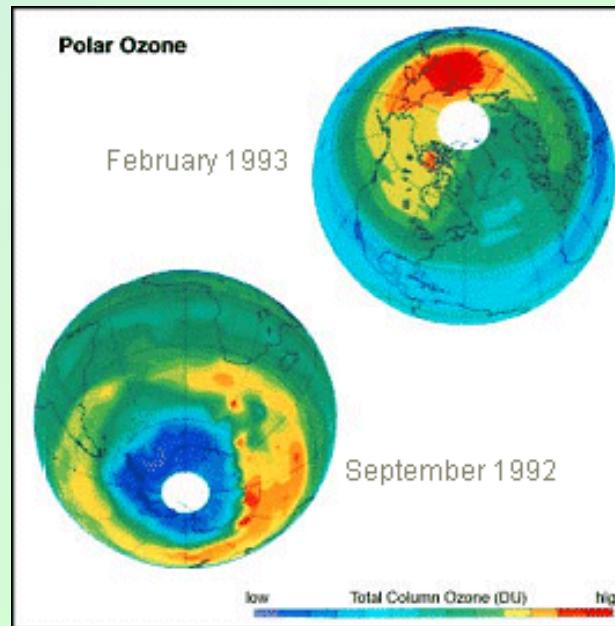
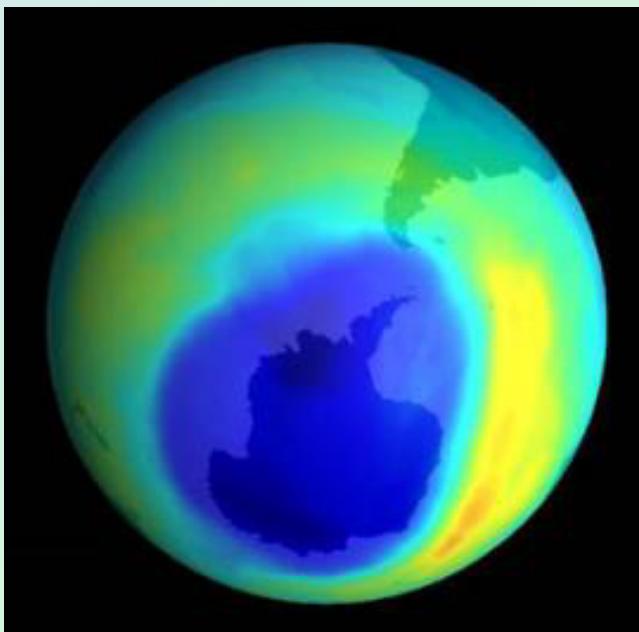
при $f=0,5$ при любом p ускорение не превышает 2, при $f=0,1$ при любых p имеем $S \leq 10$, т.к.

$$\lim_{p \rightarrow \infty} \frac{1}{f + \frac{\alpha(1-f)}{p}} = \frac{1}{f}$$

Закон Амдаля не учитывает задержек при передаче данных между вычислительными узлами по компьютерной сети; это учтено в сетевом законе Амдаля. К закону Амдаля близок закон Густавсона-Барсиса.

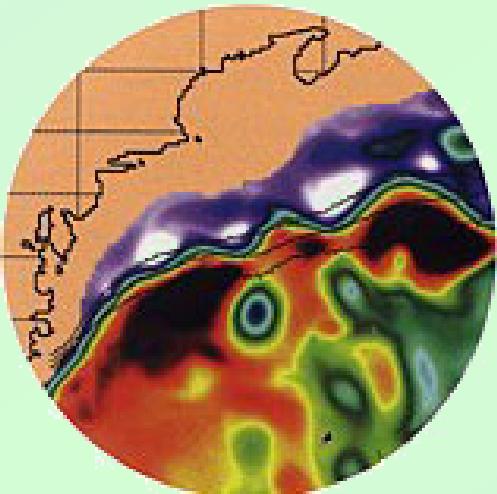


Примеры использования суперкомпьютеров для решения задач класса **GRAND CHALLENGES**

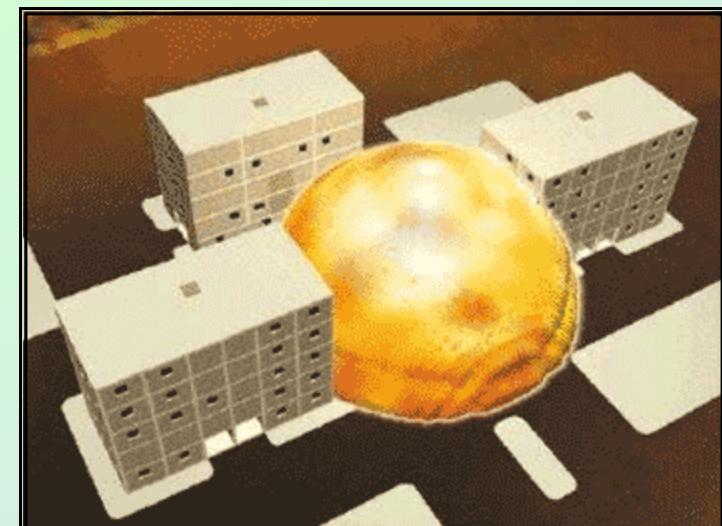


Моделирование
появления и
распространения
'озоновой дыры'
над Планетой

Математическое
моделирование
ядерных взрывов
вместо проведения
натурных
испытаний



Моделирование
поведения
Мирового океана
вблизи берегов
(приливы/отливы,
сейши, цунами)

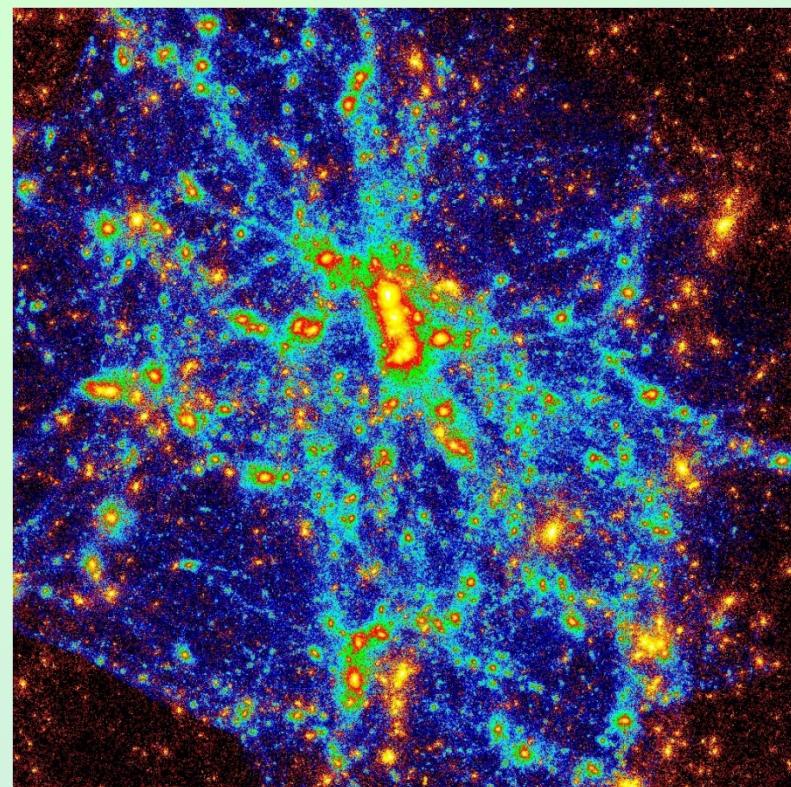


Моделирование движения N тел в гравитационном поле (процесс образования галактик)

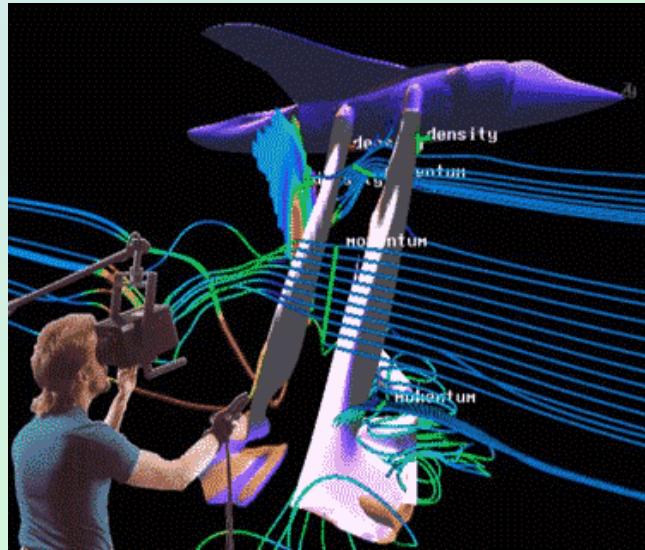


Один из начальных моментов “стягивания” протозвезд в скопления (будущие галактики) при суперкомпьютерном численном моделировании и числе $N \approx 10^7$ (подробнее об используемой аппаратуре см. слайд 28).

Процесс взаимодействия двух галактик (наблюдение оптическим телескопом)

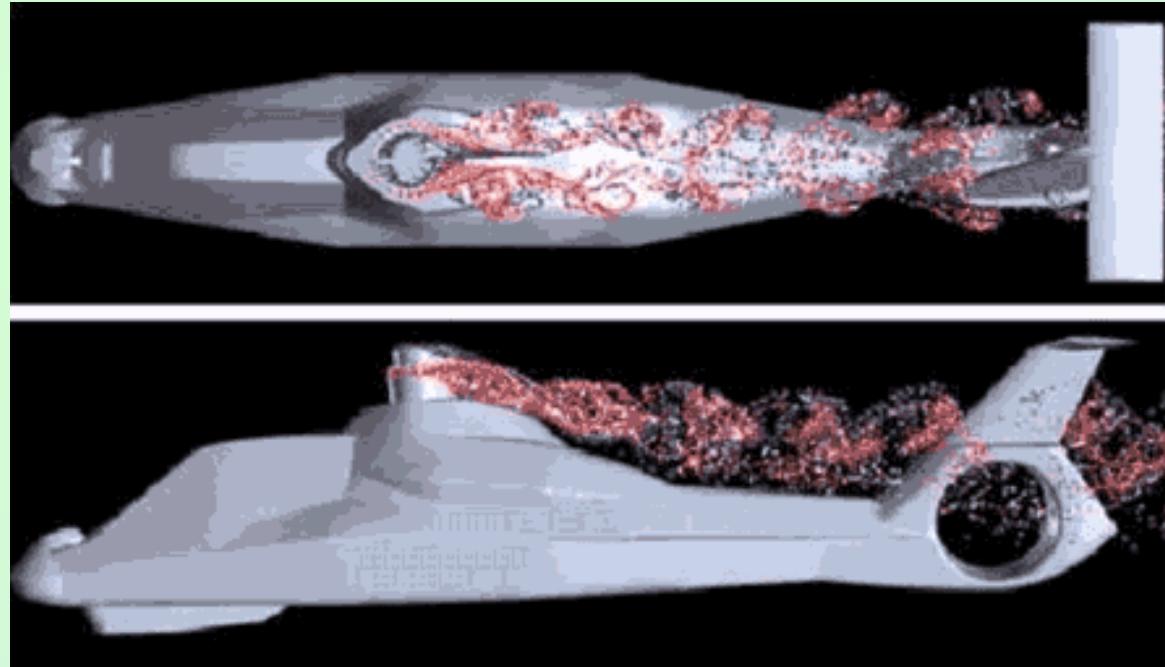


Примеры использования суперкомпьютеров для решения научно-технических задач (аэродинамика)



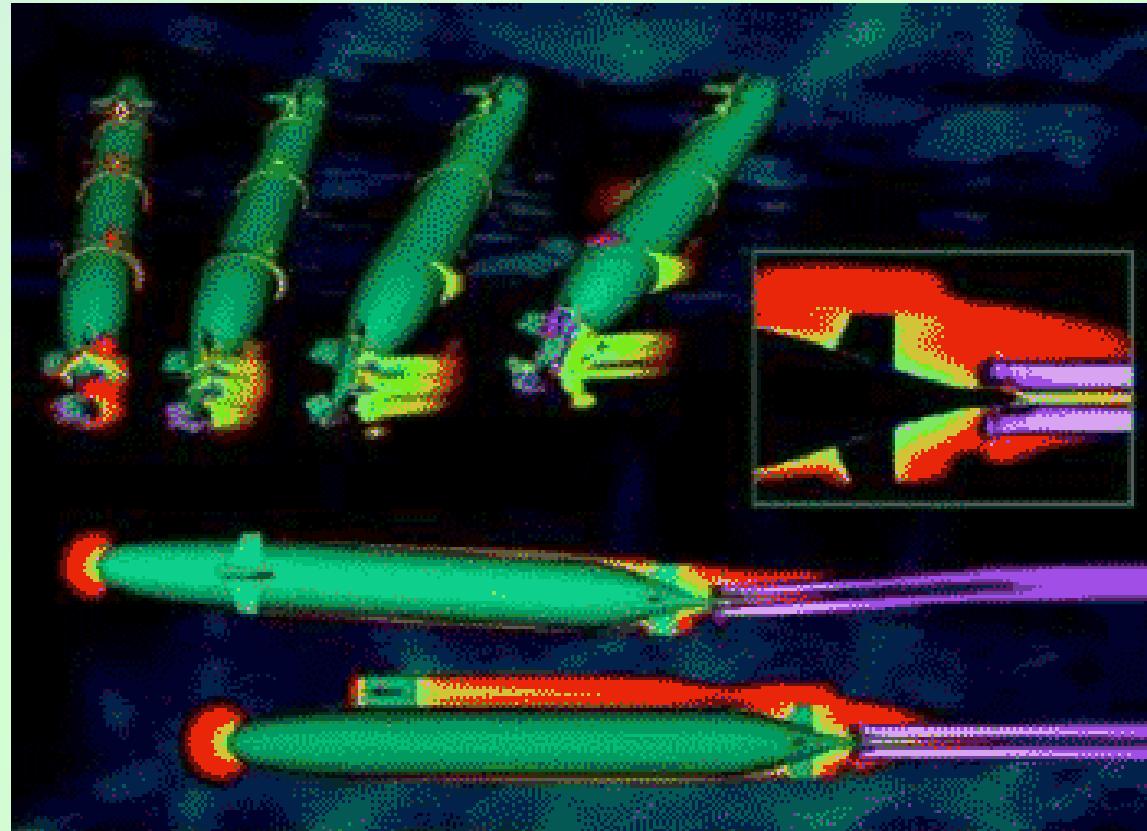
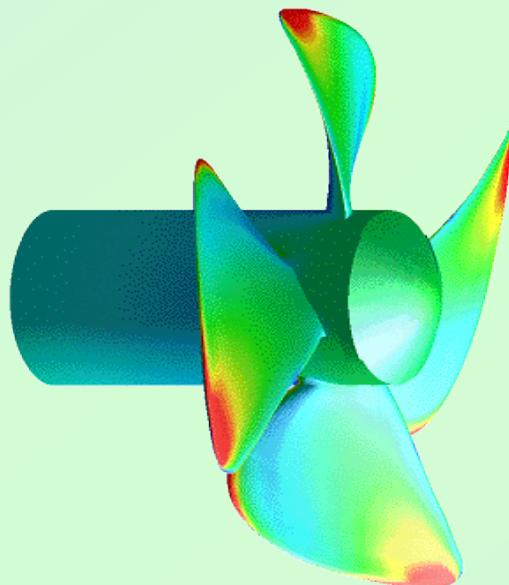
Результаты математического моделирования турбулентности при особых режимах движения вертолета

Интерактивное моделирование процесса обтекания воздухом вертикальных рулей скоростного летательного аппарата



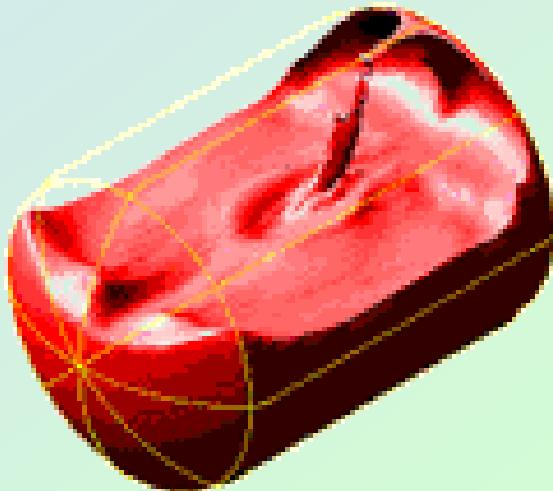
Примеры использования суперкомпьютеров для решения научно-технических задач (гидродинамика)

Результаты математического моделирования сложного пространственного перемещения подводной лодки

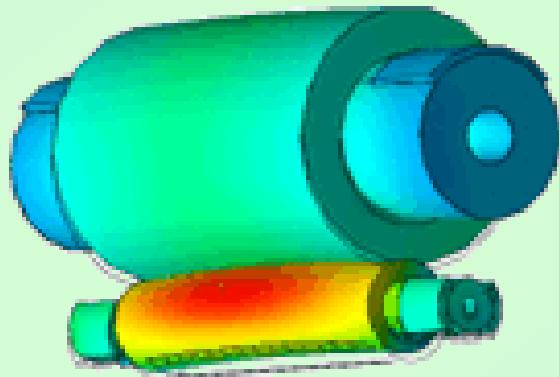


Моделирование возникновения кавитации при обтекании жидкостью винта плавсредства

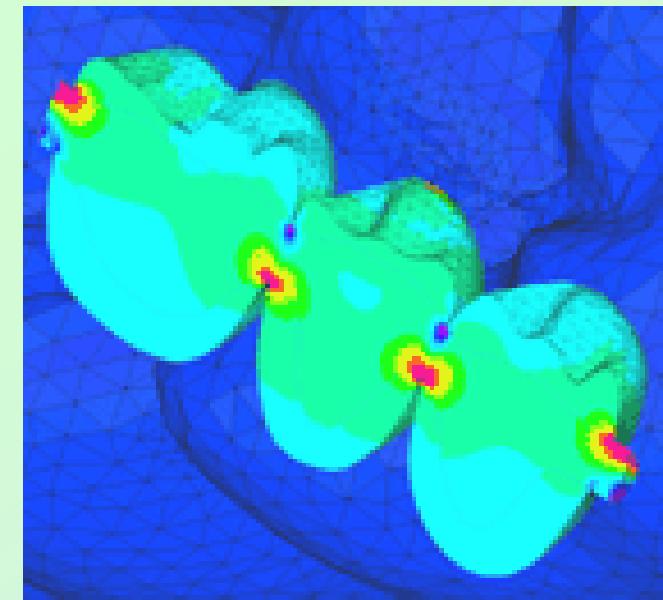
Примеры использования суперкомпьютеров для решения сложных технических задач



Моделирование колебаний жидкости в цистерне при перевозке



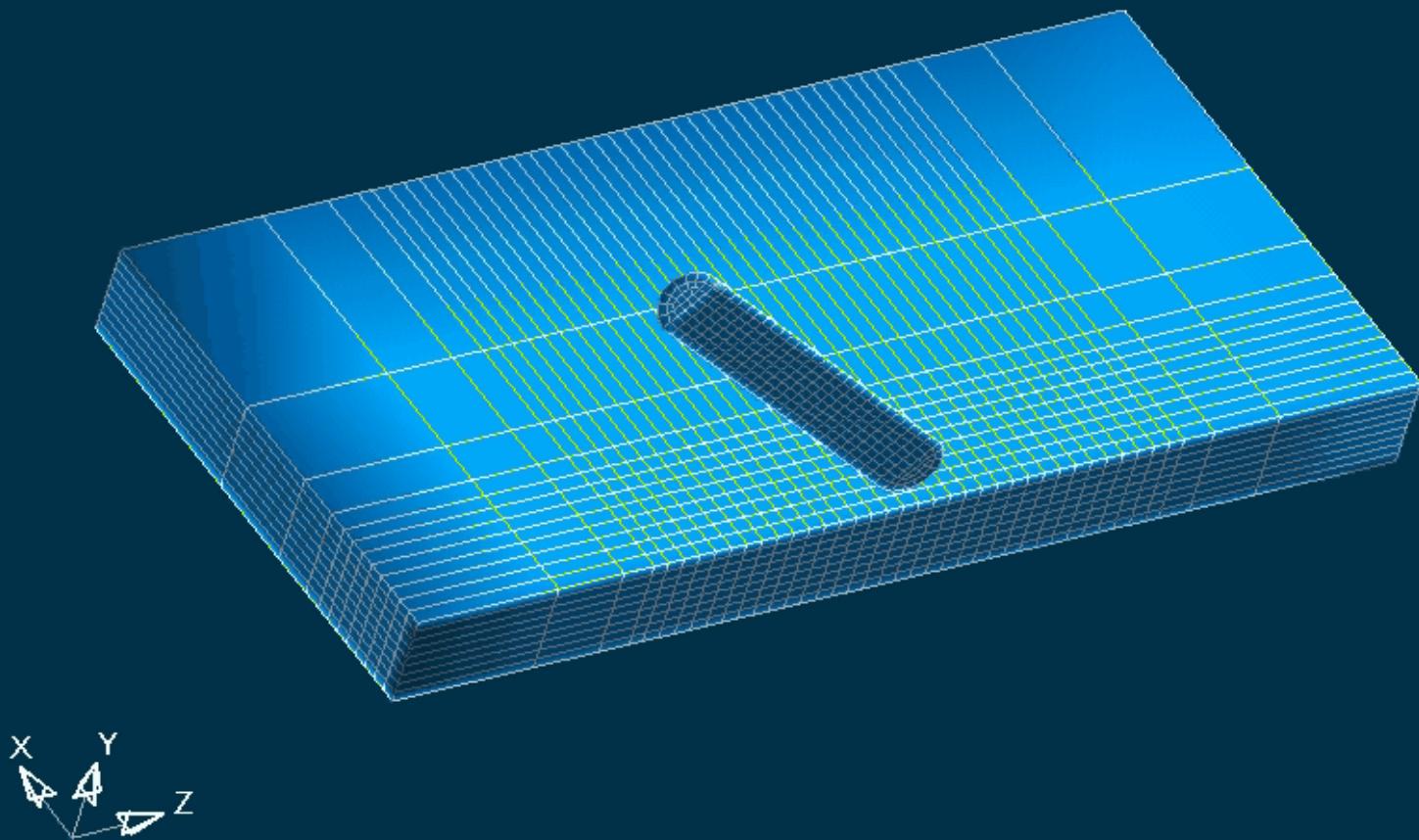
Оптимизация системы закрепления
искусственных зубов при имплантации



Визуализация напряжённо-деформированного состояния (НДС) металлодеформирующего оборудования (процесс плоской листовой прокатки)

Визуализация численного решения задачи моделирования проникновения твердого тела через сплошную пластину

Stress Von Mises



0.014111	*
0.012700	
0.011289	
0.009878	
0.008466	
0.007055	
0.005644	
0.004233	
0.002822	
0.001411	
0.000000	

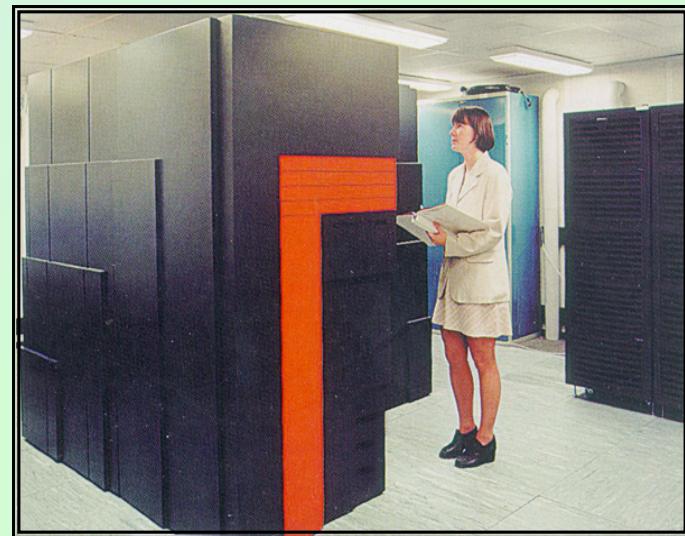
Суперкомпьютеры серии Cray (фирма *Cray Research*) (1)



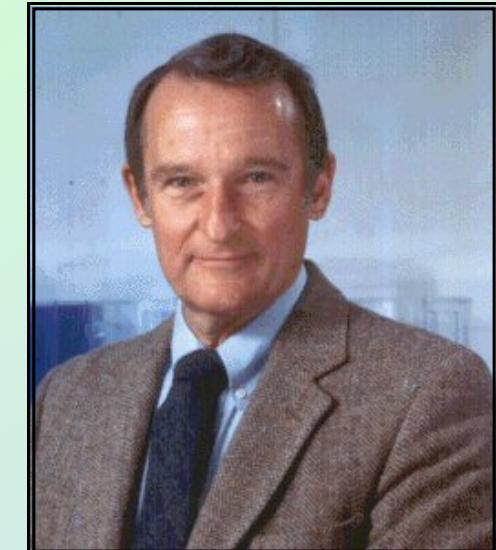
Векторный суперкомпьютер Cray 1 (1976, до 160×10^6 операций/сек)



Суперкомпьютер Cray 2



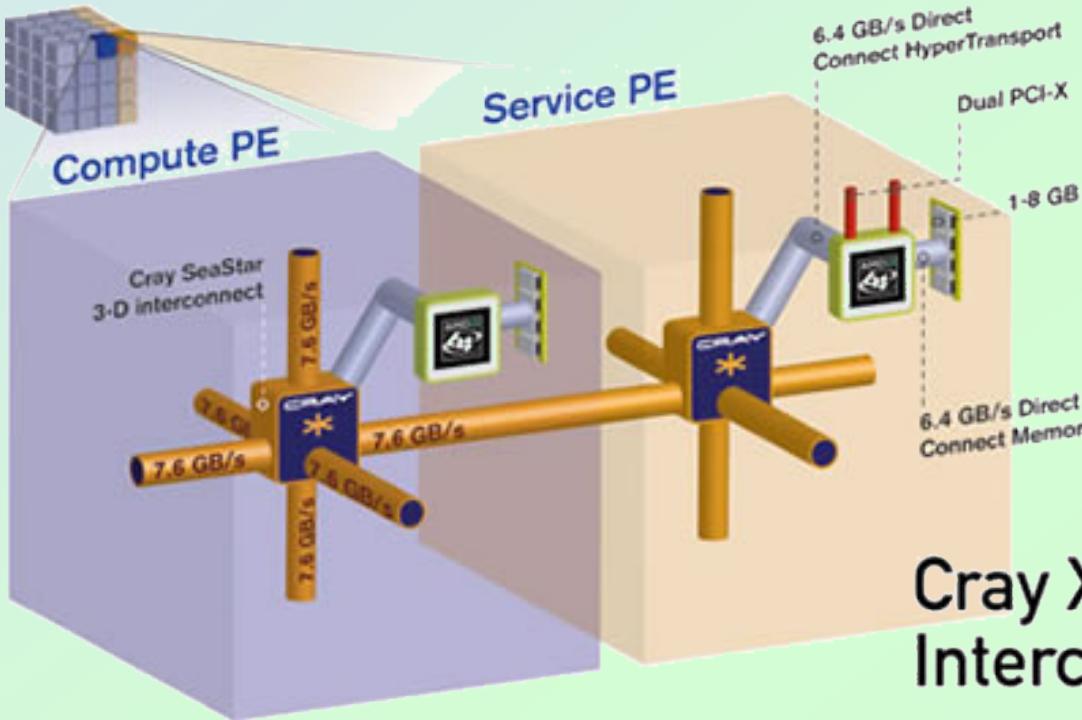
Массово-параллельный суперкомпьютер Cray T3E (1993, до 2'048 процессоров DEC Alpha, общая производительность до 2.5×10^{12} "плавающих" операций/сек)



[Сеймур Крэй](#) (29.IX.1925÷05.X.1996)

Суперкомпьютеры серии Cray (фирма *Cray Research*) (2)

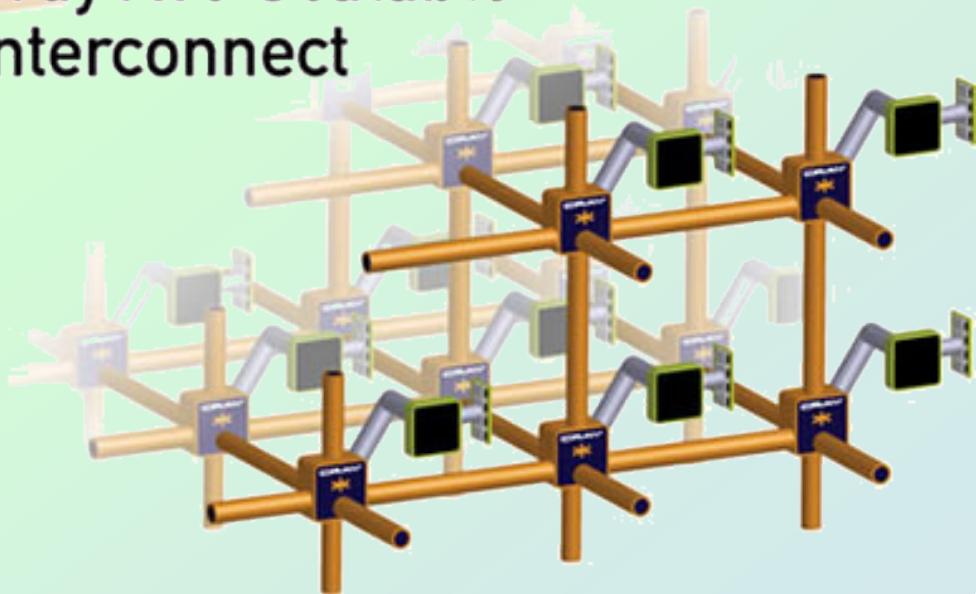
Cray XT3 Scalable Architecture



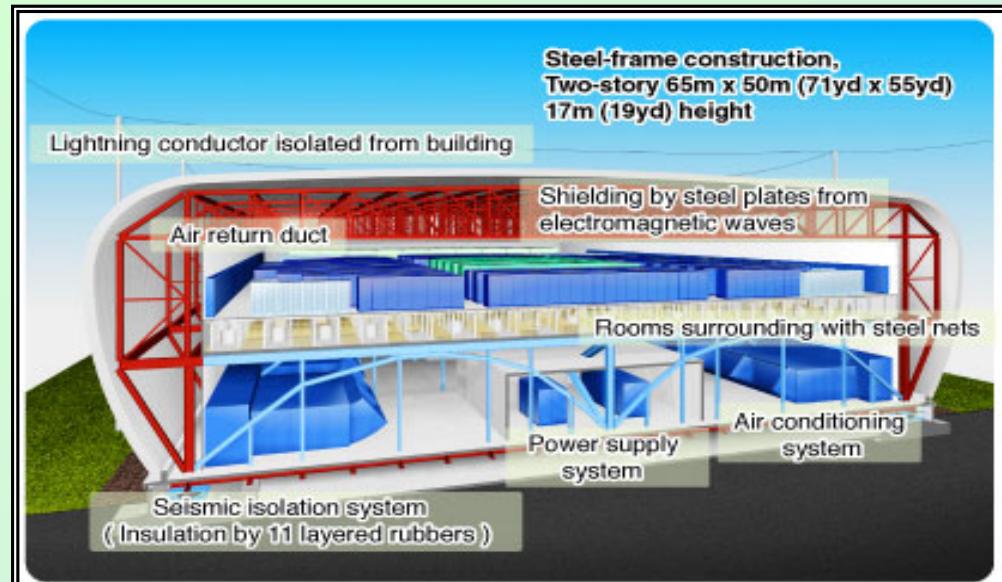
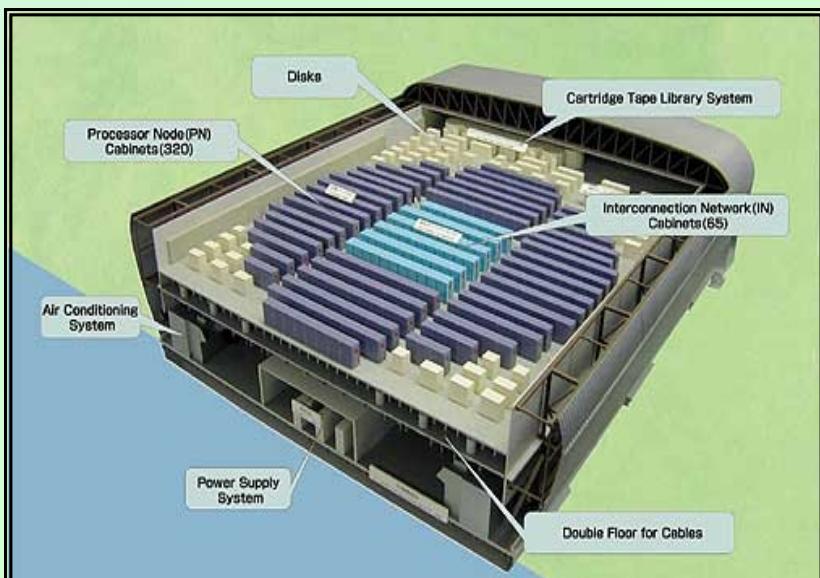
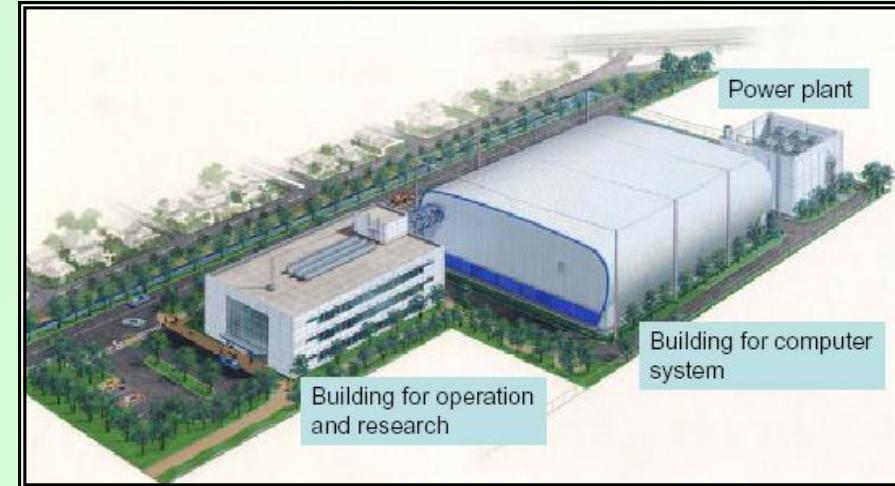
Масштабируемая архитектура объединения вычислительных узлов суперкомпьютера Cray XT3 (конфигурации - от 548 процессоров, RAM 4,3 Tb, 2,6 Tflops до 30'508 процессоров, RAM 239 Tb, 147 Tflops)

Масштабируемая архитектура суперкомпьютера Cray XT3 – показано объединение двух вычислительных узлов: вычислительного (ВУ) и сервисного). Каждый ВУ – процессор AMD Opteron 2,4 GHz, RAM 1÷8 Gb и коммутационный канал Hyper Transport

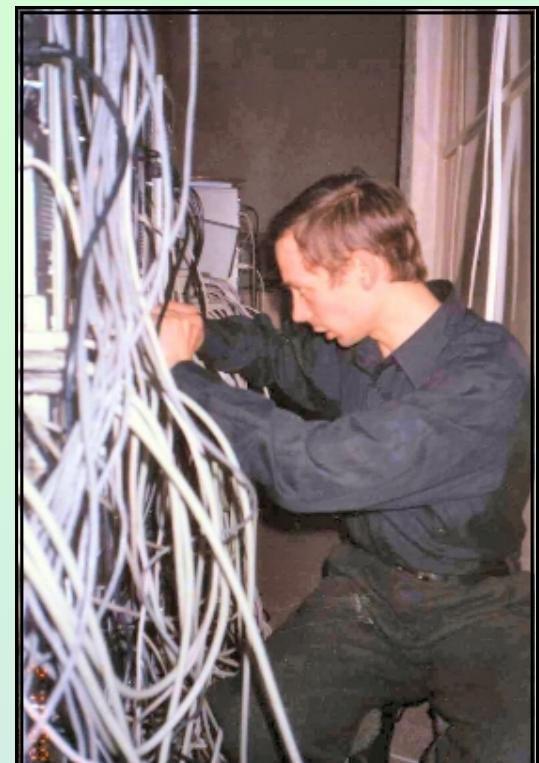
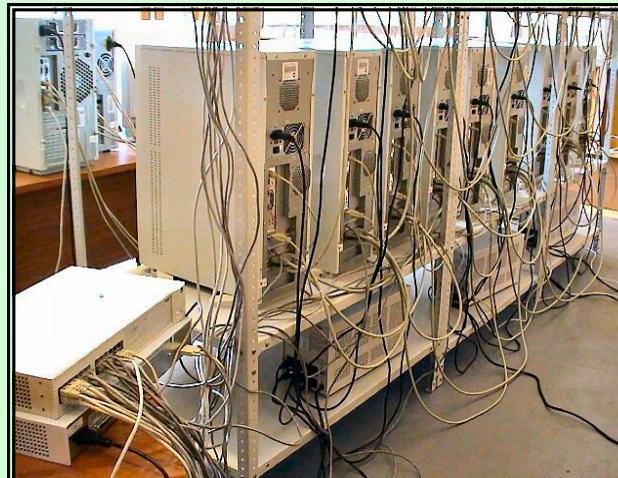
Cray XT3 Scalable Interconnect



Суперкомпьютер 'Earth Simulator' (Япония, фирма NEC Vector, производительность 40×10^{12} флопс, включает 5'120 процессоров SX-6, занимает площадь 65×50 м)



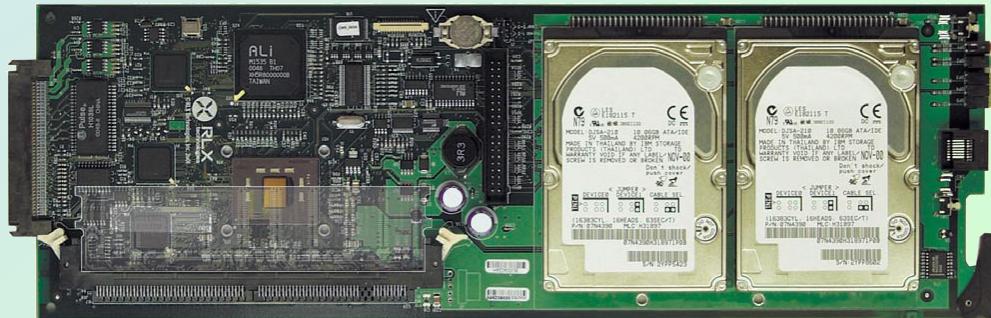
Проект BEOWULF (<http://www.beowulf.org>) – объединение десятков/сотен/тысяч ПЭВМ в единый вычислительный Linux-клáстер путём применения быстродействующих низколатентных (малоинерционных) сетей (Fast/Gigabit Ethernet, Myrinet, SCI etc)



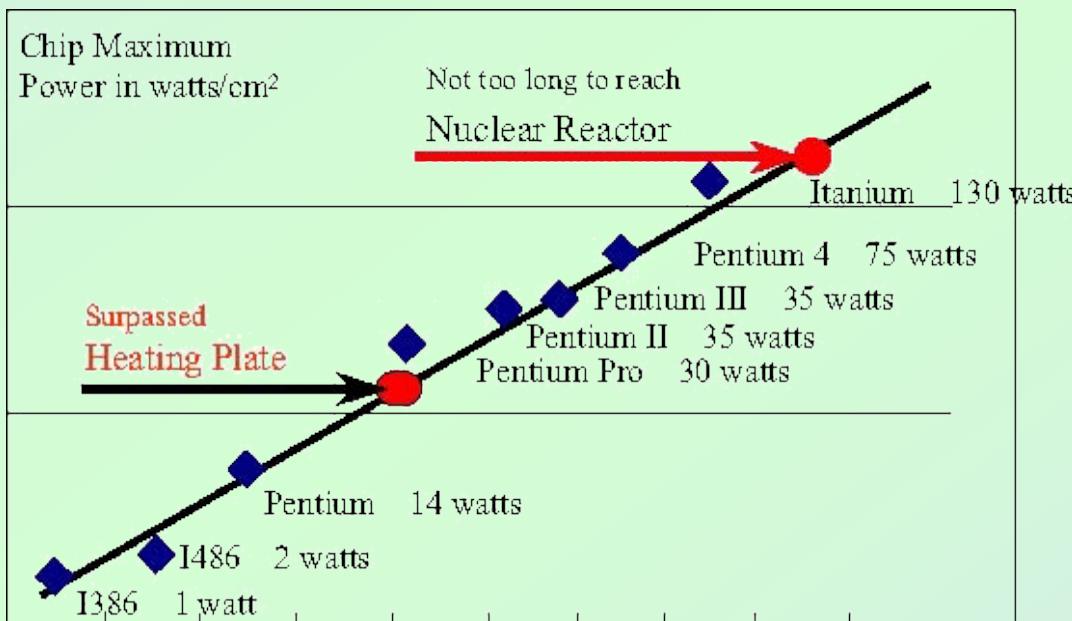
Некоторые известные суперкомпьютерные установки, созданные в учебных и академических учреждениях по традиционной схеме вычислительных Linux-кластеров МРР-архитектуры



Аппаратура для моделирование движения N тел в гравитационном поле (процесс образование галактик); Los Alamos National Laboratory



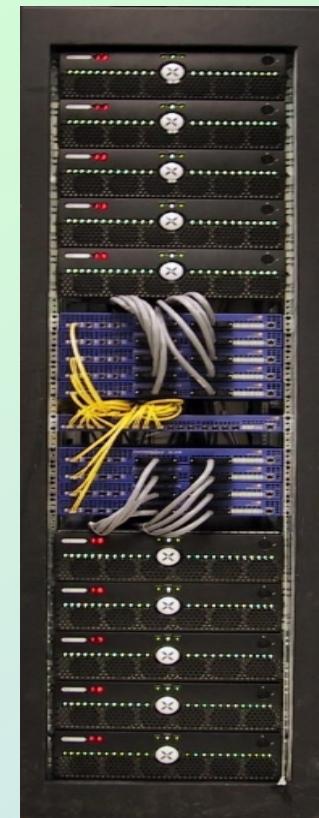
“Корзина” (24 платы) и монтажная стойка системы Bladed Beowulf (10 “корзин” - 240 вычислительных узлов) →



Лезвийная (“bladed”) плата (процессор Transmeta Crusoe TM5600 633 Mhz, RAM 256 Mb, HDD 10 Gb, 3x100 Mb Fast Ethernet) суперпроизводительного вычислительного кластера RLX System 324 (*Bladed Beowulf*)



Потребляемая современными микропроцессорами мощность (для Transmeta Crusoe всего 6 Watt)



Посещение студентами кла́стеров учебно-научного центра по высокопроизводительным вычислениям НИИВЦ МГУ



Суперклластер **ЛОМОНОСОВ**
(Россия, НИИВЦ МГУ, 2011, Т-
Платформы, 33'072 ядра Xeon
X5570/X5670 2.93 GHz, Nvidia 2070
GPU, 2.8×10^{15} flops); 18-е место в
списке Top500.org на ноябрь 2011 г.

Студенты в зале посетителей суперкластера ЛОМОНОСОВ НИИВЦ МГУ в апреле 2009 г.



Почему далеко не всегда суммарная производительность МВС пропорциональна числу вычислительных узлов (метафора [Воеводина Вл.В.](#), руководителя [НИИВЦ МГУ](#))

Формулировка вопроса: *Если 1 землекоп может выкопать яму заданного размера за 1 час, то за какое время выкопают такую же яму 10 землекопов?*



На самом деле в вопросе не хватает указания на **форму ямы** – чем она уже и глубже, тем сильнее землекопы при работе мешают друг другу и процесс идёт всё медленнее...

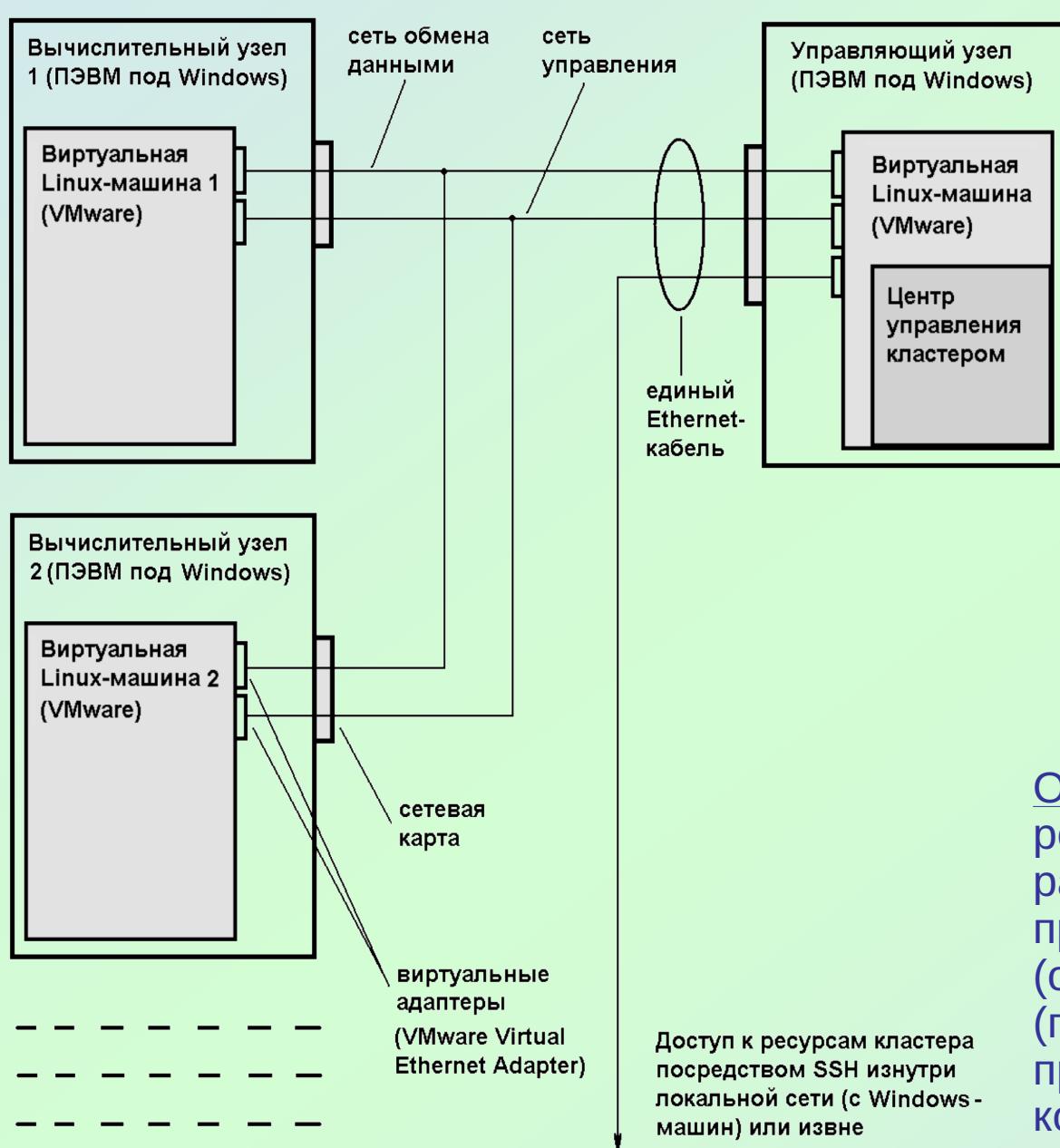


Аналогии метафоры:

- 1) землекопы отдельные вычислительные узлы (ВУ) МВС,
- 2) процесс выкапывания ямы выполнение программы,
- 3) явления взаимодействия (*“если землекопов много, они при работе мешают друг другу”*) возрастание трафика в объединяющей ВУ компьютерной сети и, как следствие, задержки в доставке данных для обработки ВУ (они вынуждены ожидать, не производя при этом вычислений); при этом общее быстродействие МВС может катастрофически падать. **Что делать?**

Аппаратный путь – применять специальные сетевые технологии с минимальными задержками при обмене данными (весьма дорого).

Интеллектуальный путь – рационально (из условия минимума пересылок блоков данных) распределять по ВУ общие пулы данных (аппаратная часть МВС не изменяется, но требуется глубокий анализ последовательности выполнения конкретного алгоритма; [см., напр., здесь](#)).



Принципиальная схема виртуального кластера MVS-900 разработки ИПМ им. М.В.Келдыша РАН (испытательная лаборатория проекта MBC, Алексей Лацис, lacis@a5.kiam.ru)

Основные преимущества – реализован режим совместной работы Windows-машин при проведении занятий и (одновременной) отладке (прогонке) параллельных программ на едином комплексе оборудования.

Количественные характеристики коммуникационной среды кластерных систем с распределённой памятью

$T = X/S + L$, где T – затрачиваемое на передачу сообщения время (сек),
 S – пропускная способность канала “точка-точка” (Мбайт/сек),
 X – длина сообщения (Мбайт),
 L – латентность (время “запуска” обмена), сек;
 L не зависит от X , при $X \rightarrow 0$ имеем $T \rightarrow L$

$P = LS$, где P – “цена обмена” (латентность, приведённая к максимальной скорости обмена), Мбайт

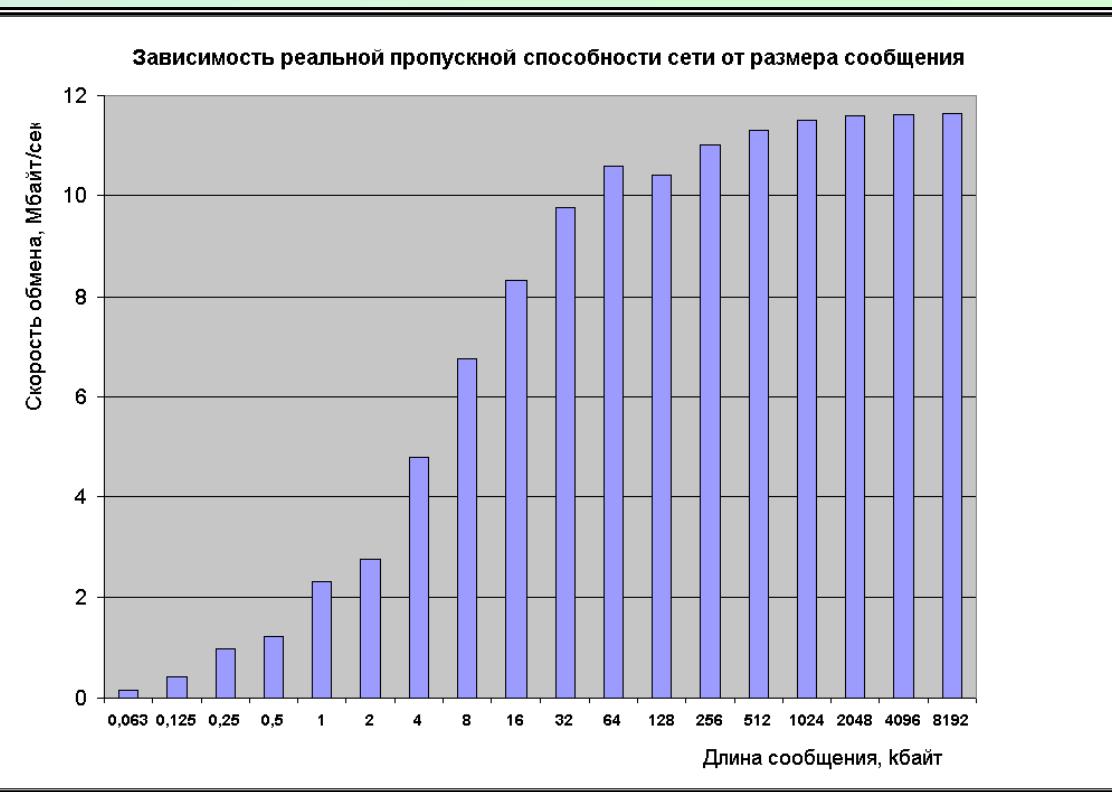
Для обычных недорогих Ethernet (Fast/Gigabit Ethernet) – сетей латентность L может достигать 200-400 мксек (теоретический максимум – 50-70 мксек). Специальное (дорогостоящее) сетевое оборудование (Myrinet, SCI, Giganet, Quadrics) позволяет снизить значение L до единиц мксек!



Низколатентные сетевые адAPTERы технологии SCI фирмы *Dolphin Interconnect Solutions, Inc.*

Опыты по определению пропускной способности коммуникационной сети клáстера

Опыты проводились в лаборатории кафедры ИТ-4 МГУПИ; время 16-19 час. Пропускная способность сети замерялась на операциях “точка – точка” программой PROGMPI.C для MVS-900 на [Slackware Linux](#). HOST и 2 NODE’ы на Athlon 1800 MHz, ОП 256 Mb, HDD 80 Gb, сетевой адаптер NVIDIA nForce встроенный, сетевой коммутатор (switch) COMPEX PS2208D 100 Мбит/сек. На NODE’ах сторонние задачи (кроме VMware Workstation 3.2.0 build-2230, ОП: заказано 64 Mb, реально выделено 88 Mb) не выполнялись, нагрузка (согласно “Диспетчеру задач Windows”) не выше 45-55%

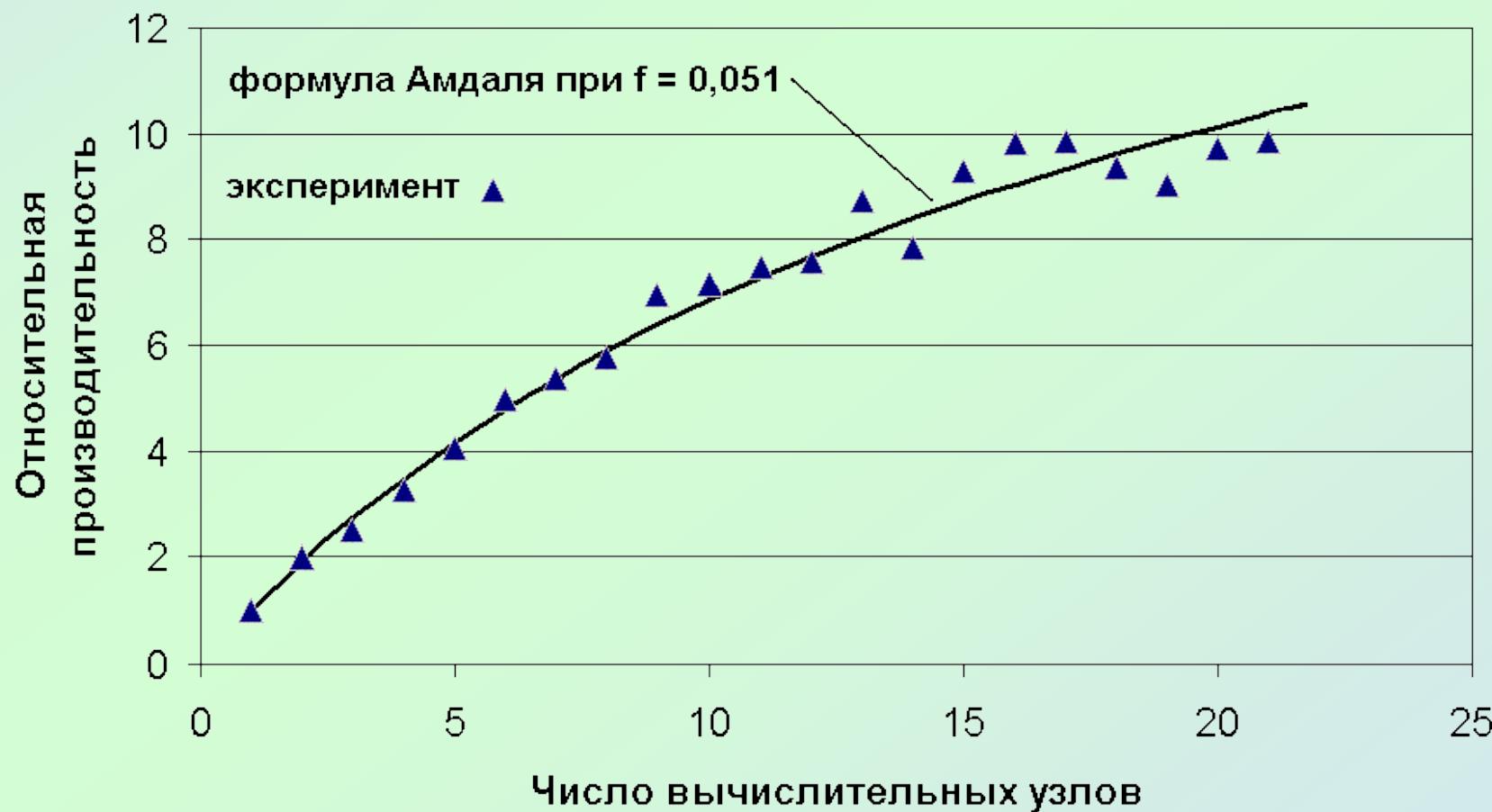


Количественная (на основе данных эксперимента) оценка латентности сетевого оборудования дает значение около 330 мксек (~0,3 мсек)

Экспериментальное определение производительности вычислительного кластера

Задача: экспериментально оценить зависимость производительности вычислительного кластера от числа вычислительных узлов (процессоров).

Тестовый пример: умножение матриц размерностью 1000x1000 double-чисел ([программа MM_MPI_1.C](#)); кластер SCI-MAIN НИВЦ МГУ (опыты 19.III.2005, удаленный доступ по протоколу SSH).



Проект “Одомашненный BEOWUF” (направление исследований: введение технологий параллельных вычислений в быт – *персонализация кластерных технологий*)

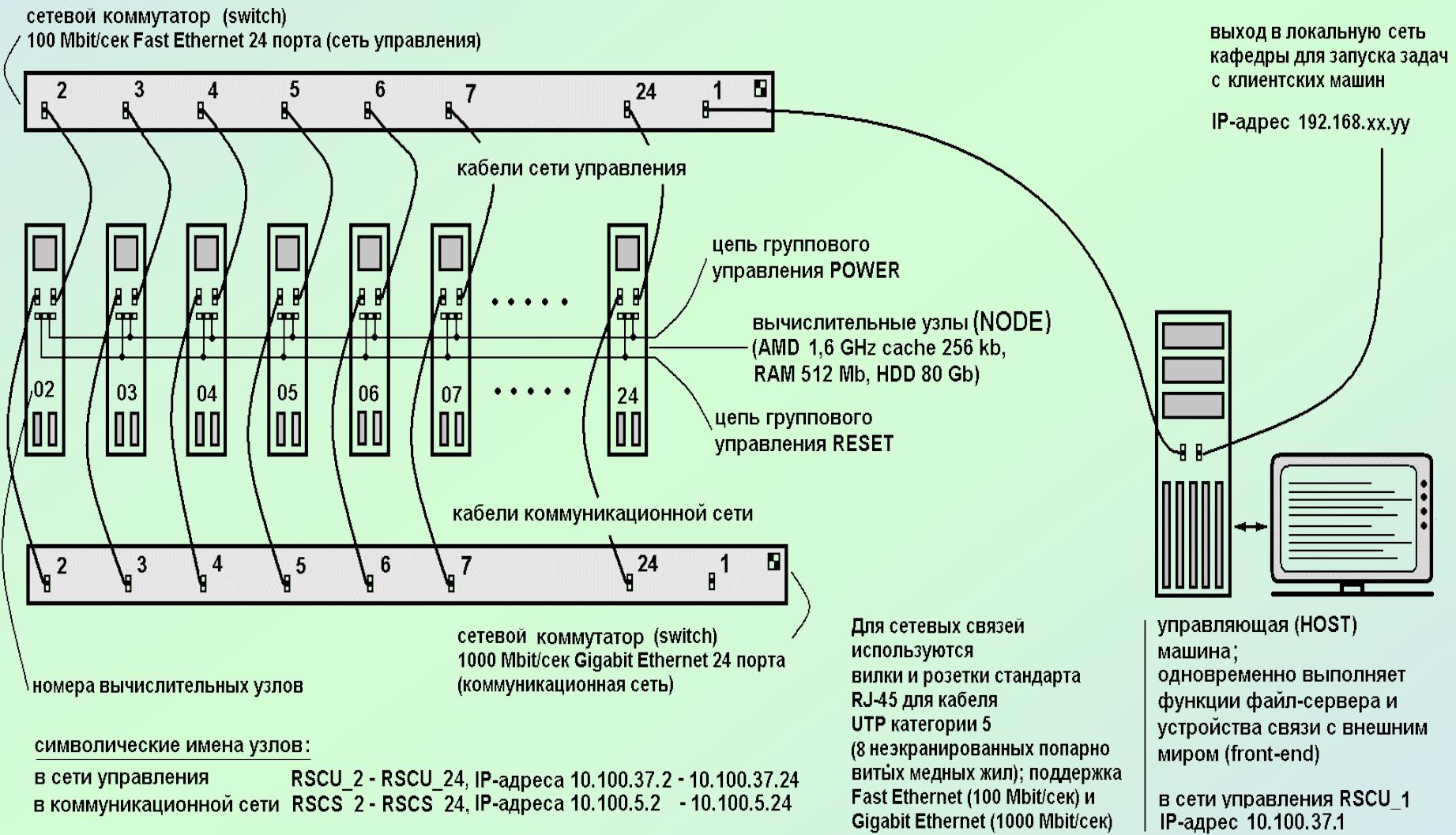
Цель проекта – создание типовой модели дешевого (стоимость не более 200 \$US на узел) общедоступного **персонального вычислительного кластера** для освоения принципов системного администрирования и различных технологий параллельного программирования).

Созданное оборудование включает вычислительный блок (размер 760x250x280 мм, вес ~17 кг) из 5 стандартных системных плат формата ПАТХ с процессорами Celeron 2,0 GHz, HDD 20 Gb, RAM 256 Mb, L2-кэш 256 kb и управляющую ПЭВМ (обычный персональный компьютер), операционная система – Slackware 9.1.

Программное обеспечение – библиотека MPI (в версии LAM), пакеты AZTEC и ScaLAPACK; устанавливается система DVM.



Схема вычислительного кластера уровня кафедры (реализованный вариант)



Внешний вид МВС (вычислительный кластер) кафедрального уровня (реализованный вариант)



Основные технические данные:

1. Всего вычислительных узлов – до 24 (размещение – по 6 узлов на каждой полке).
2. Монитор технического обслуживания (дисплей 10" и клавиатура) располагаются на второй полке сверху.
3. Максимальное энергопотребление – до 2 kW (ток 10 А при напряжении 220 вольт); тепловыделение до 7 Мега Джоулей/час).
4. В настоящее время МВС как вычислительный ресурс включена в локальную сеть кафедры КБ-5 МГУПИ / МИРЭА под IP-адресом **192.168.47.61**.



28 июля 2009
года; 14:00

МОСКВА, КРЕМЛЬ. Дмитрий Медведев провёл совещание с членами Совета Безопасности по вопросам создания и применения суперкомпьютеров.

Президент сформулировал пять основных задач для работы в этой сфере:

- 1. Определение приоритетных направлений использования суперкомпьютерных и грид-технологий (использование множества серверов как единого целого) в области обеспечения национальной безопасности и социально-экономического развития страны.**
- 2. Разработка мер повышения уровня отечественной электронной компонентной базы до потребностей производства суперкомпьютеров.**
- 3. Формирование нормативно-правовой базы применения суперкомпьютеров.**
- 4. Создание условий для построения грид-сетей, разработка специализированного программного обеспечения.**
- 5. Организация специальной системы подготовки специалистов в ведущих ВУЗах страны.**

InterNet - ресурсы для проведения Исследований:

Математическая (компьютерная) модель вычислителя потоковой архитектуры для исследования произвольных алгоритмов на наличие внутреннего (скрытого) параллелизма и параметров его практического использования (Data-Flow) -

Описание здесь: <http://vbakanov.ru/dataflow/>

Инсталляция здесь: http://vbakanov.ru/dataflow/content/install_df.exe

Программная система для исследования и выбора рациональных методов построения планов (расписаний) выполнения программ на заданном поле параллельных вычислителей (SPF@home) -

Описание здесь: <http://vbakanov.ru/spf@home/>

Инсталляция здесь: http://vbakanov.ru/spf@home/content/install_spf.exe

О вычислительном кластере кафедры КБ-5 МИРЭА -

http://vbakanov.ru/hist_clu/clusters.htm

Тематические статьи на Habr'e -

<https://habr.com/ru/post/530078/> <https://habr.com/ru/post/534722/>

<https://habr.com/ru/post/535926/> <https://habr.com/ru/post/540122/>

<https://habr.com/ru/post/545498/> <https://habr.com/ru/post/551688/>

“Стихотворные” произведения о ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЯХ ([“околонаучный стёб”](#)) -

[Размышления о необходимости параллельных вычислений](#)

[О пользе потоковых \(DATA-FLOW\) вычислительных архитектур](#)

Информация для самостоятельной работы

Ключевые слова для поиска в сети InterNet: суперкомпьютеры; задачи уровня **GRAND CHALLENGES**; задачи БОЛЬШОГО ВЫЗОВА; журнал “Суперкомпьютеры”; проблема петафлопса – ссылка 1, ссылка 2, ссылка 3; список Top 500; Top 50; BEOWULF; Томас Стерлинг о ЗЕТАФЛОПСЕ; многопроцессорные (многомашинные) вычислительные системы (МВС); ILLIAK-IV; SMP и MPP-архитектура; смесь Гибсона; flops; флопс; коммуникационная среда; латентность; закон Амдаля; закон Густавсона-Барсиса; суперкомпьютер ЭЛЬБРУС для системы ПРО А-135 (ещё); быстродействующие низколатентные сетевые технологии; Ethernet; Mirinet; SCI; Infiniband; Cray Research; Сеймур Крэй - дополнительно; LinPACK; HPL; ScalapACK; прямой метод Гаусса решения систем линейных алгебраических уравнений (СЛАУ); Graph500; алгебра логики; дизъюнктивная нормальная форма; потеря точности вследствие округления при вычислениях; лаборатория Параллельных Информационных Технологий НИИВЦ МГУ, <http://parallel.ru>; Межведомственный Суперкомпьютерный Центр (МСЦ РАН), <http://www.jsc.ru>); Институт Прикладной Математики (ИПМ им.М.В.Келдыша РАН, <http://www.kiam.ru>)

О выдающихся ПЕРСОНАЛИЯХ в области информационных технологий см. также здесь ...ещё больше поисковых систем

Некоторые известные (и популярные!) системы поиска информации в Сети: (больше – здесь)



Яндекс



YAHOO!
по-русски

Баканов Валерий Михайлович

E-Mail: e881e@mail.ru

WEB: <http://vbakanov.ru>