



# ОБЗОР ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Если создать систему, которой сможет пользоваться даже дурак, то только дурак захочет ею пользоваться.

*А. Блох*

# Содержание

2

- Пути достижения параллелизма вычислений
- Определение суперкомпьютера
- Примеры суперкомпьютеров

# Параллельные вычисления

3

- *Параллельные вычисления* – процессы решения задач, в которых в один и тот же момент времени могут выполняться одновременно несколько вычислительных операций.

# Пути достижения параллелизма

4

- Независимое функционирование отдельных устройств компьютера (устройства ввода-вывода, обрабатывающие процессоры и устройства памяти).
- Избыточность элементов вычислительной системы
  - использование специализированных устройств
    - отдельные процессоры для целочисленной и вещественной арифметики, устройства многоуровневой памяти;
  - дублирование устройств
    - использование нескольких однотипных обрабатывающих процессоров или нескольких устройств оперативной памяти
- Конвейерная реализация обрабатывающих устройств.

# Пути достижения параллелизма

5

- *Многозадачный (псевдопараллельный) режим*
  - для выполнения нескольких процессов используется единственный процессор
- *Параллельный режим*
  - в один и тот же момент времени может выполняться несколько команд обработки данных (обеспечивается при наличии нескольких процессоров или при помощи конвейерных и векторных обрабатывающих устройств)
- *Режим распределенных вычислений*
  - для параллельной обработки данных используется несколько удаленных друг от друга обрабатывающих устройств, а передача данных по линиям связи приводит к существенным временным задержкам.

# Суперкомпьютер

6

- *Суперкомпьютер* – вычислительная система, обладающая предельными характеристиками по производительности среди компьютерных систем, имеющихся в данное время.
- Другие определения ☺
  - Суперкомпьютер – любой компьютер, весящий более одной тонны.
  - Суперкомпьютер – любой компьютер, который создал Сеймур Крей.
- Рейтинги-листы суперкомпьютеров
  - Мировой: TOP500 ([top500.org](http://top500.org))
  - Российский: TOP50 ([supercomputers.ru](http://supercomputers.ru))



Cray-2, самый быстрый компьютер 1985-89 гг.



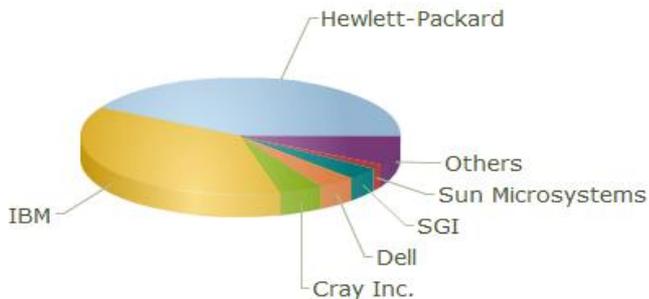
# Суперкомпьютеры TOP500

7

## TOP 10 Systems - 11/2008

- 1 Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Valtaire Infiniband
- 2 Jaguar - Cray XT5 QC 2.3 GHz
- 3 Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0/2.8 GHz
- 4 BlueGene/L - eServer Blue Gene Solution
- 5 Blue Gene/P Solution
- 6 Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband
- 7 Franklin - Cray XT4 QuadCore 2.3 GHz
- 8 Jaguar - Cray XT4 QuadCore 2.1 GHz
- 9 Red Storm - Sandia/ Cray Red Storm, XT3/4, 2.4/2.2 GHz dual/quad core
- 10 Dawning 5000A - Dawning 5000A, QC Opteron 1.9 Ghz, Infiniband, Windows HPC 2008

Производители / системы  
(11/2008)



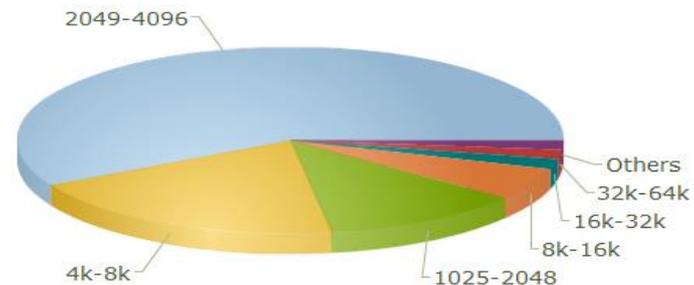
Производители / производительность  
(11/2008)



Архитектура / системы  
(11/2008)



Количество процессоров / системы  
(11/2008)



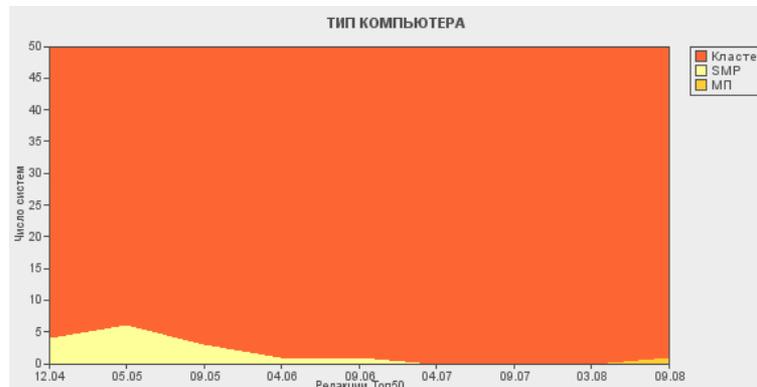
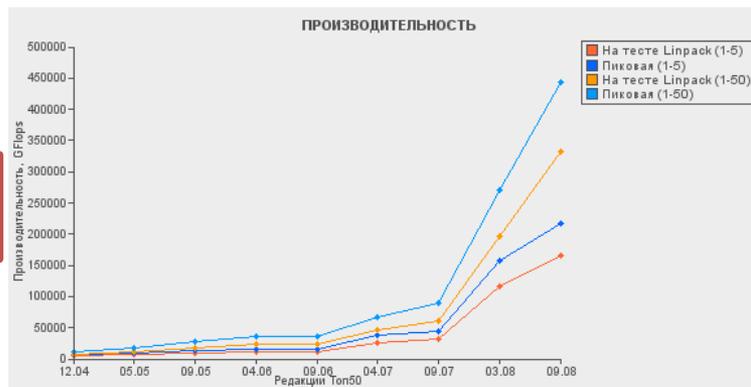
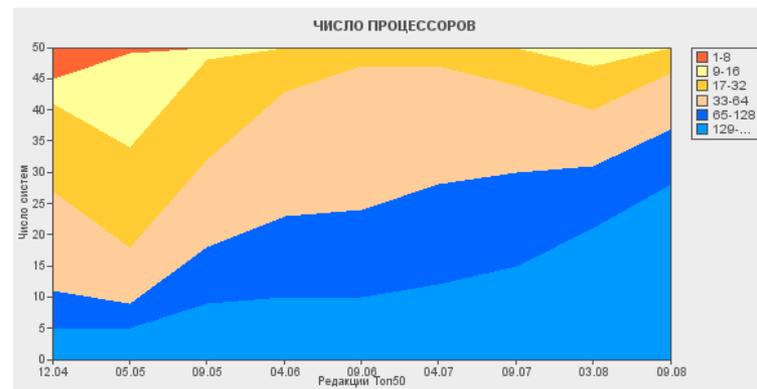
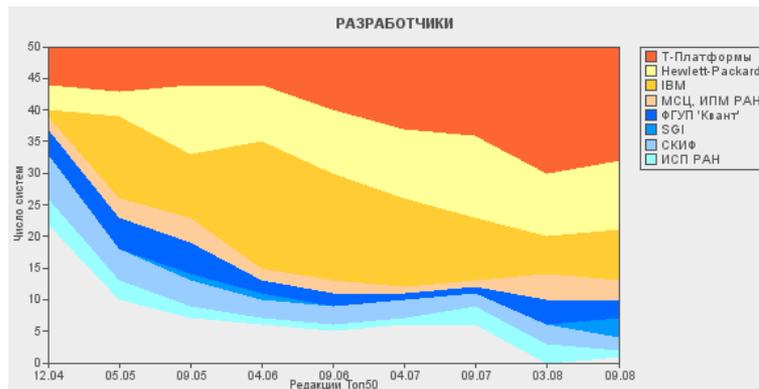
# Суперкомпьютеры TOP50

8

## Текущий рейтинг

9-ая редакция от 23.09.2008

N	Место	Кол-во CPU/ядер
1	Москва <a href="#">Межведомственный суперкомпьютерный центр Российской академии наук 2008 г.</a>	1564/6256
2	Москва <a href="#">Московский государственный университет имени М.В. Ломоносова 2008 г.</a>	1250/5000
3	Москва <a href="#">РНИЦ Курчатовский институт 2008 г.</a>	864/3456
4	Москва <a href="#">Факультет Вычислительной математики и кибернетики Московский государственный университет имени М.В. Ломоносова 2008 г.</a>	2048/8192
5	Уфа <a href="#">УГАТУ 2007 г.</a>	532/2128
6	Москва <a href="#">Росгидромет 2008 г.</a>	354/1416
7	Красноярск <a href="#">СФУ 2007 г.</a>	452/1808
8	Челябинск <a href="#">Южно-Уральский государственный университет 2008 г.</a>	332/1328
9	Рыбинск <a href="#">НПО Сатурн 2008 г.</a>	336/1344
10	Москва <a href="#">Росгидромет 2008 г.</a>	832/1664



# Производительность суперкомпьютеров

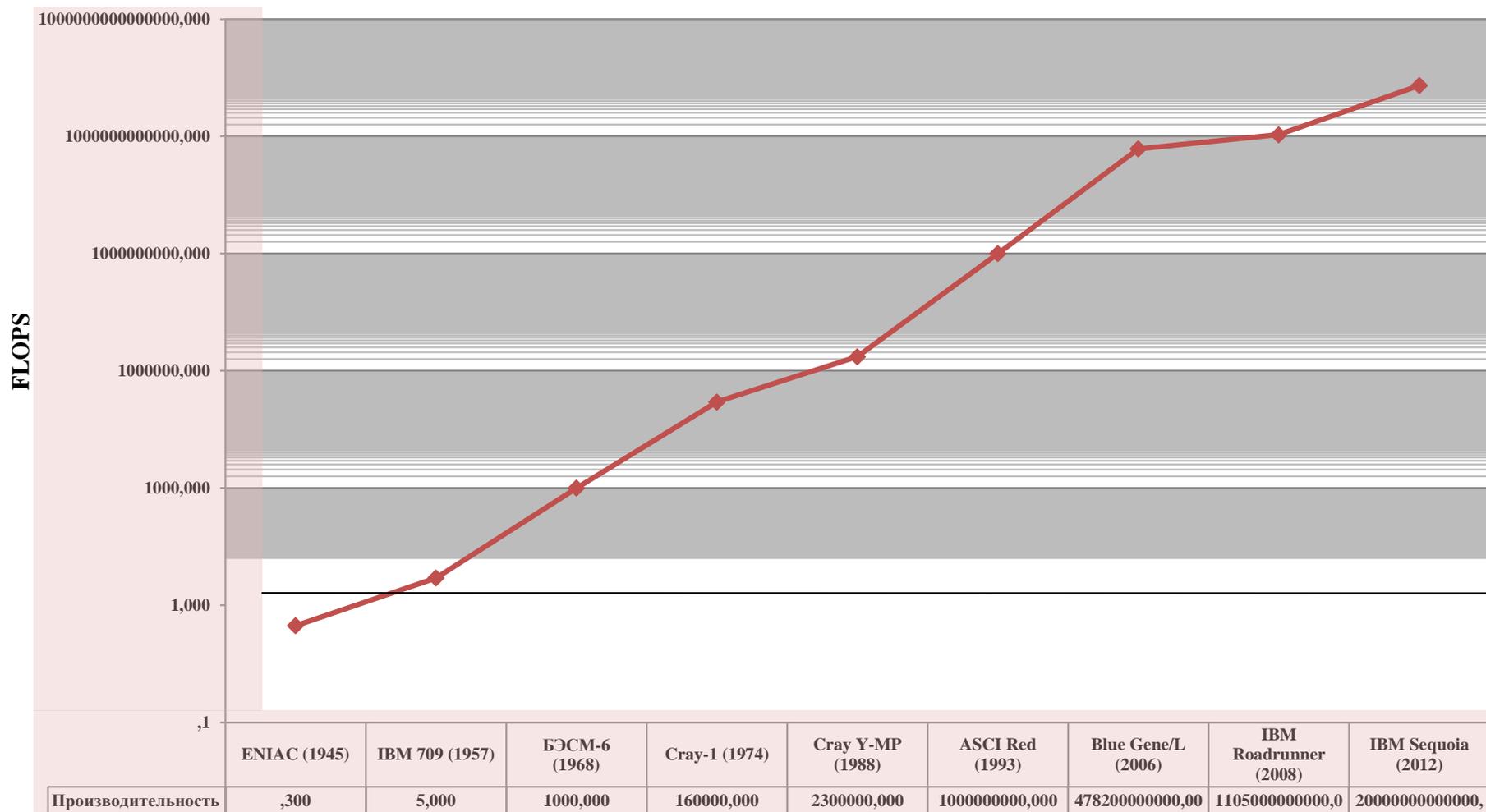
9

- *FLOPS (Floating point Operations Per Second, флорпс)* – мера производительности, показывающая, сколько операций с плавающей точкой в секунду выполняет данная вычислительная система.

Название	год	FLOPS
флорпс	1941	10
килофлорпс	1949	10 <sup>3</sup>
мегафлорпс	1964	10 <sup>6</sup>
гигафлорпс	1987	10 <sup>9</sup>
терафлорпс	1997	10 <sup>12</sup>
петафлорпс	2008	10 <sup>15</sup>
эксафлорпс	2021	10 <sup>18</sup>
зеттафлорпс	2037	10 <sup>21</sup>
йоттафлорпс	2066	10 <sup>24</sup>
ксерофлорпс	2082	10 <sup>27</sup>

# Производительность суперкомпьютеров

10



# Границы применимости FLOPS

11

- Флопс не всегда адекватная мера производительности, поскольку неоднозначным является уже само его определение.
  - Под "операцией с плавающей запятой" может скрываться масса разных понятий.
  - Существенную роль в данных вычислениях играет разрядность операндов, которая также нигде не оговаривается.
  - Величина флопс подвержена влиянию очень многих факторов, напрямую не связанных с производительностью вычислительного модуля:
    - пропускная способность каналов связи с окружением процессора
    - производительность основной памяти
    - синхронность работы кэш-памяти разных уровней.
- Результаты, полученные на одном и том же компьютере при помощи разных программ, могут существенным образом отличаться. Более того, с каждым новым испытанием разные результаты можно получить при использовании одного алгоритма.

# Границы применимости FLOPS

12

- Суперкомпьютер MDGrape-3 (Исследовательский институт RIKEN, Япония), пиковая производительность 1 Пфлопс.
  - Данный компьютер не является компьютером общего назначения и приспособлен для решения задач моделирования сворачивания белков, и стандартный тест LINPACK на нем выполнить невозможно в силу особенностей его архитектуры.
- Игровая приставка Xbox 360 (Microsoft), пиковая производительность 1 Тфлопс, приставка PlayStation 3 (SONY) – 2 Тфлопс (суперкомпьютеры начального уровня!).
  - Операции с 3d графикой, которые они в основном выполняют, очень хорошо поддаются распараллеливанию, что с успехом используется в графических процессорах.
  - Однако эти процессоры не в состоянии выполнять большинство задач общего назначения, и их производительность не поддаётся оценке теста LINPACK и сравнению с другими системами.

# Пиковая и Linpack производительность

13

- *Пиковая производительность* – суммарная производительность процессоров (ядер), из которых состоит суперкомпьютер
- *Linpack-производительность* определяется путём запуска на суперкомпьютер теста LINPACK (решение систем линейных алгебраических уравнений), в ходе которого решается задача с известным количеством операций и подсчитывается время, за которое она была решена.

# Популярность Linpack

14

- Флопс является абсолютной величиной.
- Многие задачи инженерной и научной практики в конечном итоге сводятся к решению СЛАУ, а тест LINPACK как раз и базируется на измерении скорости решения таких систем.
- Подавляющее большинство компьютеров (включая суперкомпьютеры) построены по классической архитектуре с использованием стандартных процессоров, что позволяет использовать общепринятые тесты с большой достоверностью.

# Программа ASCI (США)

15

- *ASCI (Accelerated Strategic Computing Initiative)* – программа правительства США по развитию американских суперкомпьютерных технологий.
  - 1996 – ASCI Red (Intel), 1 Тфлопс
  - 1999 – ASCI Blue Pacific (IBM), ASCI Blue Mountain (SGI), 3 Тфлопс
  - 2000 – ASCI White, пиковая производительность свыше 12 Тфлопс, Linpack-производительность 4938 GFlops.



IBM Blue Pacific

# ASCI White

16

- 512 симметричных мультипроцессорных (SMP) узлов, каждый узел имеет 16 процессоров
- Процессоры IBM RS/6000 POWER3
  - 64-разрядная архитектура
  - конвейерная организация
  - 2 устройства по обработке команд с плавающей запятой
  - 3 устройства по обработке целочисленных команд
  - устройства способны выполнять до 8 команд за тактовый цикл и до 4 операций с плавающей запятой за такт, тактовая частота 375 MHz
- Оперативная память – 4 Тб
- Емкость дискового пространства 180 Тб
- Операционная система – IBM AIX (версия UNIX ),
- Программное обеспечение
  - Поддерживается гибридная модель программирования – передача сообщений между узлами и нити внутри SMP-узла
  - Поддерживаются MPI, OpenMP, нити POSIX и транслятор директив IBM, имеется параллельный отладчик IBM.

# BlueGene

17

- 2004 – 1 место в Top500
- 2007– расширенный вариант, 1 место в Top500
  - 212 992 2-ядерных 32-битных процессора PowerPC 4400.7 ГГц
  - пиковая производительность – 600 Тфлопс
  - Linpack-производительность – 478 Тфлопс

# RoadRunner

18

- ❑ Самый быстрый суперкомпьютер (2008), первый в мире, производительность которого превысила рубеж 1 Пфлопс (1000 Тфлопс)
  - ❑ 12 960 процессоров IBM PowerXCell 8i
  - ❑ 6480 2-ядерных процессоров AMD Opteron
  - ❑ пиковая производительность около 1700 Тфлопс
  - ❑ Linpack-производительность – 1026 Тфлопс



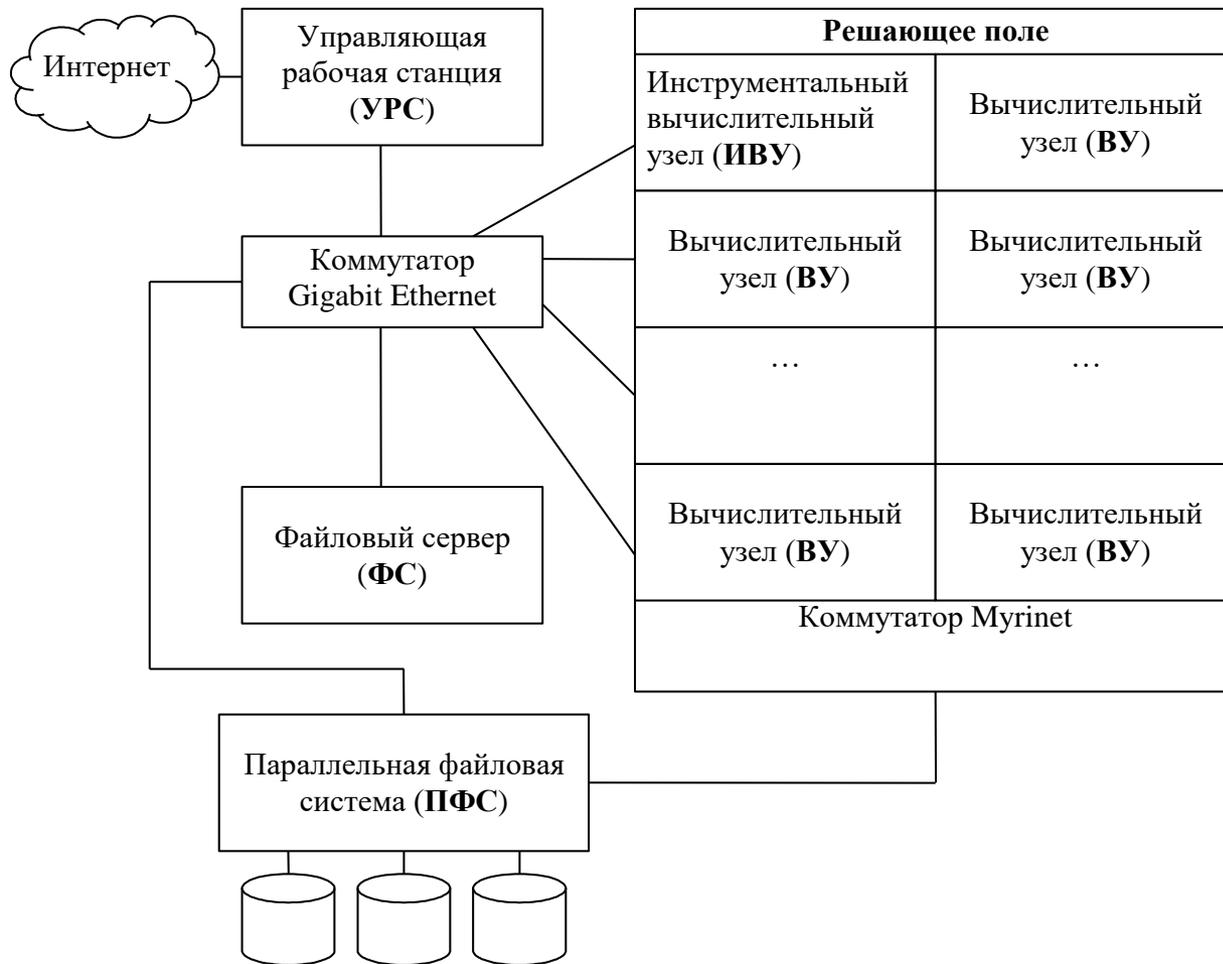
# MBC-1500

19

- 2005, Межведомственный суперкомпьютерный центр РАН
  - ▣ 276 узлов, 552 процессора
  - ▣ Узел:
    - 2 процессора IBM PowerPC 970 с тактовой частотой 2.2 ГГц, кэш L1 96 Кб и кэш L2 512 Кб
    - 4 Гб оперативной памяти на узел
    - 40 Гб жесткий диск IDE
  - ▣ Операционная система SuSe Linux Enterprise Server 8 для платформ x86 и PowerPC
  - ▣ Пиковая производительность 4857.6 Gflops
  - ▣ Linpack-производительность – 3052 GFlops.

# MBC-1500

20

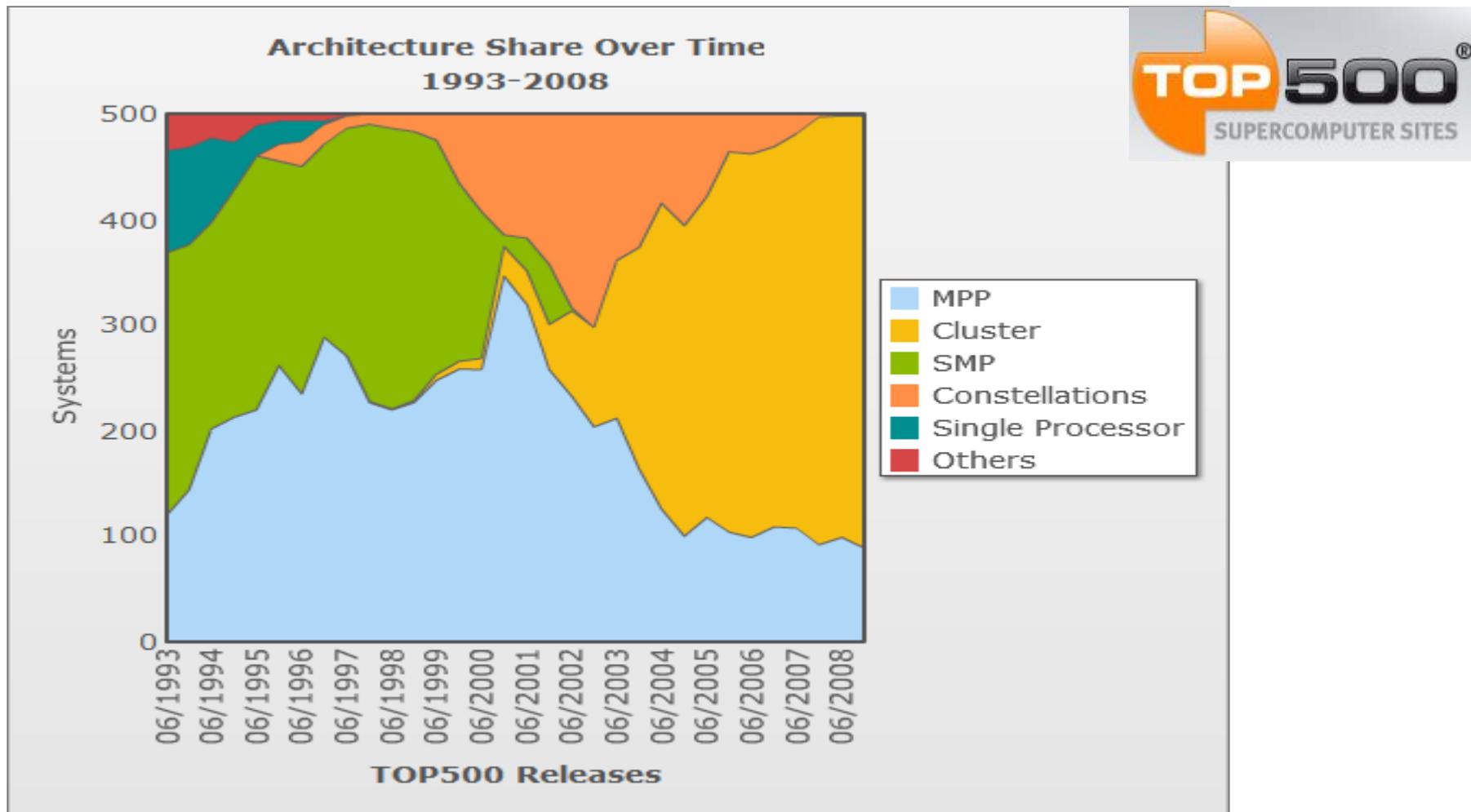


# Кластеры

- *Кластерная вычислительная система (кластер)* – набор рабочих станций или персональных компьютеров общего назначения, объединенных в систему с помощью одной из стандартных сетевых технологий (Fast или Gigabit Ethernet, Myrinet и др.) на базе шинной архитектуры или коммутатора.
- Кластер предполагает более высокую надежность и эффективность, чем ЛВС, и существенно более низкую стоимость в сравнении с другими видами параллельных вычислительных систем (за счет использования типовых аппаратных и программных решений).

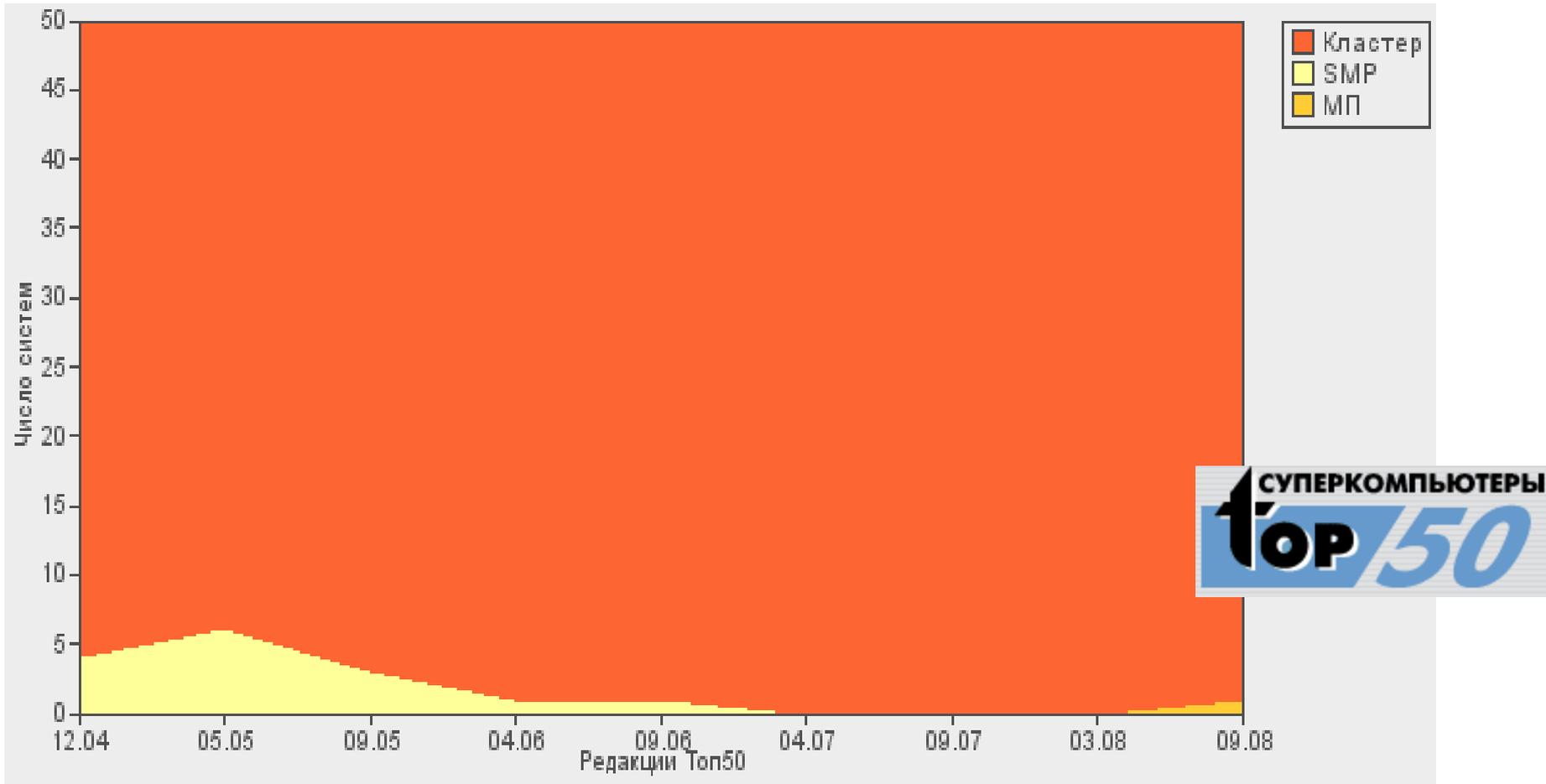
# Кластеры

22



# Кластеры

23



# СКИФ МГУ

24

- 2008, НИВЦ МГУ
  - 625 2-процессорных узлов (1250 4-ядерных процессоров Intel Xeon E5472 3.0 ГГц)
  - Общий объем оперативной памяти – 5,5 Тбайт
  - Объем дисковой памяти узлов – 15 Тбайт
  - Операционная система Linux
  - Пиковая производительность 60 Тфлопс
  - Linpack-производительность 47 Тфлопс.



# Infinity

25

- Число вычислительных узлов/процессоров: 26/52
- Тип процессора:  
Intel Xeon EM64T 3.2 ГГц
- Оперативная память: 58 Гб
- Дисковая память: 2400 Гб
- Тип системной сети:  
InfiniBand (PCI-Express 4x)
- Тип управляющей (вспомогательной) сети: Gigabit Ethernet
- Пиковая производительность: 333 GFlops
- Производительность на тесте Linpack:  
270 GFlops
- Операционная система:
  - ▣ Linux Gentoo 2006.1
  - ▣ Windows Compute Cluster Server 2003
- Система бесперебойного электропитания: 15 kVA



# СКИФ Урал

26

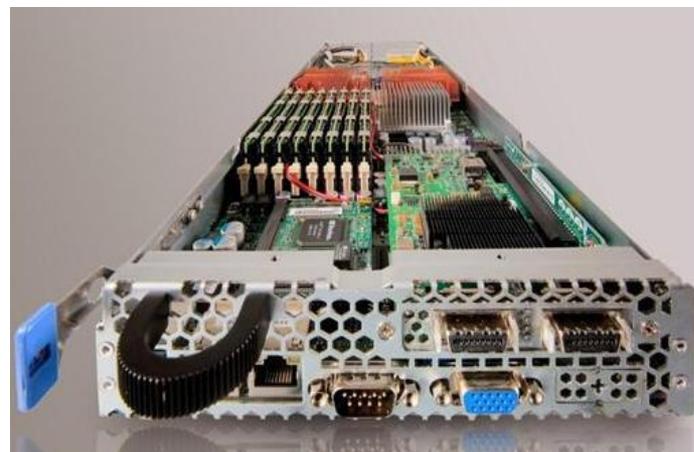
- ❑ Число вычислительных узлов/процессоров/ядер:  
166/332/1328
- ❑ Оперативная память: 1.33 Тб
- ❑ Тип процессора:  
Intel Xeon E5472 (4 ядра по 3.0 ГГц)
- ❑ Дисковая память: 49.29 Тб
- ❑ Параллельная система хранения данных:  
T-Platforms ReadyStorage – 20ТВ
- ❑ Тип системной сети:  
InfiniBand (20 Гбит/s, макс. задержка 2 мкс)
- ❑ Тип управляющей сети: Gigabit Ethernet
- ❑ Сервисная сеть: СКИФСervNet
- ❑ Пиковая производительность: 16 Тфлопс
- ❑ Производительность на тесте Linpack: 12.2 Тфлопс
- ❑ Операционная система:
  - ❑ SUSE Linux Enterprise Server 10
  - ❑ Windows HPC Server 2008
- ❑ Система бесперебойного электропитания:  
APC Symmetra 160 kVA



# Вычислительный узел СКИФ Урал

27

- Тип процессора:  
Intel Xeon E5472 (4 ядра по 3.0ГГц)
- Оперативная память: 8 Гб
- Дисковая память: 320 Гб
- Операционная система:
  - ▣ SUSE Linux Enterprise Server 10
  - ▣ Windows Computer Cluster Server



# Хост-машина СКИФ Урал

28

- Тип процессора:  
Intel Xeon E5472 (4 ядра по 3.0 ГГц)
- Оперативная память: 16 Гб
- Дисковая память: 1.25 Тб RAID
- Операционная система:
  - SUSE Linux Enterprise Server 10
  - Windows Computer Cluster Server



# Параллельная система хранения данных СКИФ Урал

29

- Производительное хранилище данных Panasas ActiveStorage 5100 (20 Тб) с архитектурой NAS, разработанное специально для кластеров
- Пропускная способность более 60 Гбит/сек
- Параллельный доступ к данным всех узлов кластера



# Программное обеспечение СКИФ Урал

30

- LSTC LS-DYNA – многоцелевой конечно-элементный комплекс разработки Livermore Software Technology Corp. для анализа высоконелинейных и быстротекущих процессов в задачах механики твердого и жидкого тела.
- TECIS FlowVision – комплексное решение в области моделирования трехмерных турбулентных течений жидкости и газа.
- ANSYS – многоцелевой конечно-элементный пакет для проведения анализа в широком круге инженерных дисциплин (прочность, теплофизика, динамика жидкостей и газов и электромагнетизм).
- Dassault Systemes ABAQUS – программный комплекс для решения задач в области конечно-элементных прочностных расчетов для сложных линейных и нелинейных инженерных проблем.
- SFTC DEFORM – специализированный инженерный программный комплекс, предназначенный для анализа процессов обработки металлов давлением, термической и механической обработки.
- MathWorks MATLAB – специализированный пакет для решения инженерных, научно-технических и экономических задач.
- Autodesk 3ds Max 2008 – полнофункциональная профессиональная программная система для работы с трёхмерной графикой.

# Персональные суперкомпьютеры

31

- Персональный мини-кластер T-Edge Mini [t-platforms.ru/ru/temini.php](http://t-platforms.ru/ru/temini.php)
  - 4 двухпроцессорных узла на базе 4-ядерных процессоров Intel Xeon (всего 32 ядра)
  - Оперативная память – до 128 Гбайт
  - Сеть передачи данных – Gigabit Ethernet или InfiniBand
  - Операционная система – одна из: SUSE Linux Enterprise Server, RedHat Enterprise Linux, MS Windows Compute Cluster Server 2003
  - Пиковая производительность – 384 GFlops
  - Размеры – 57×33×76 см



# Заключение

- Основные пути достижения параллелизма
  - Независимое функционирование отдельных устройств компьютера
  - Избыточность элементов вычислительной системы
  - Конвейерная реализация обрабатывающих устройств
  - Параллельный режим выполнения программы на многопроцессорной вычислительной системе
- Суперкомпьютер – вычислительная система, обладающая предельными характеристиками по производительности среди компьютерных систем, имеющихся в данное время.
  - Флопс – мера производительности суперкомпьютеров.
  - TOP500 и TOP50 – рейтинг-листы суперкомпьютеров.
  - В настоящее время подавляющее большинство суперкомпьютеров имеют кластерную архитектуру.