



Уральский
федеральный
университет

Параллельные вычисления

Введение в параллельные вычисления

Созыкин Андрей Владимирович

К.Т.Н.

Заведующий кафедрой высокопроизводительных компьютерных технологий
Институт математики и компьютерных наук

Многопоточные и параллельные вычисления

Многопоточные вычисления:

- Несколько задач, которые нужно выполнять одновременно
- Производители и потребители
- Работа программы и отслеживание действий пользователя
- Запросы к Web-серверу

Параллельные вычисления:

- Одна задача, которую нужно разбить на несколько частей для ускорения
- Умножение матриц
- Быстрое преобразование Фурье

Архитектура параллельных систем

Системы с общей разделяемой памятью:

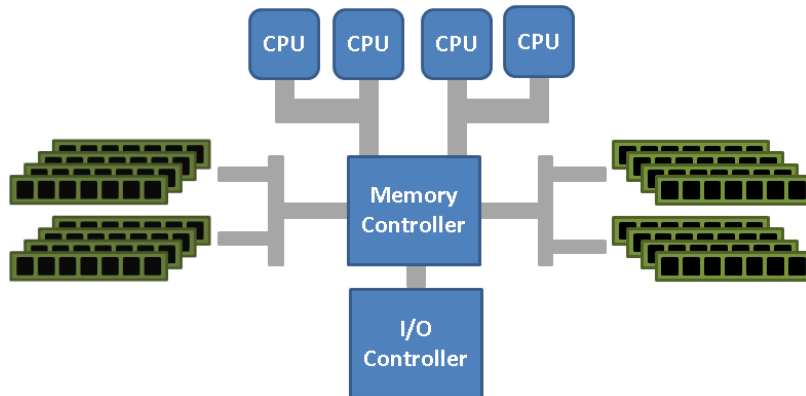
- Мультипроцессоры

Системы с распределенной памятью:

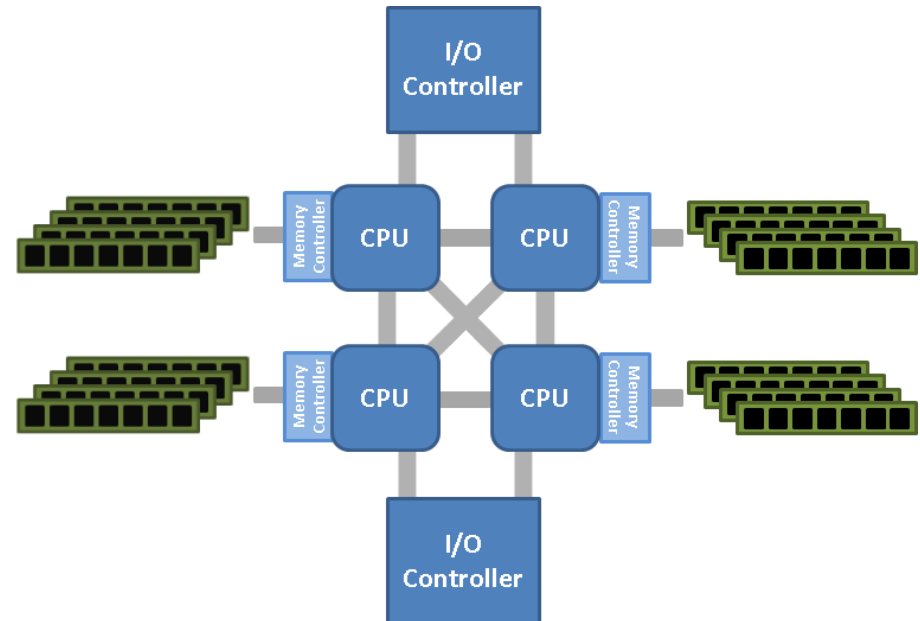
- Мультикомпьютеры

Гибридные системы

Системы с общей памятью



Симметричные
мультипроцессоры (SMP)



Неоднородный доступ к памяти
(NUMA)

Общая разделяемая память

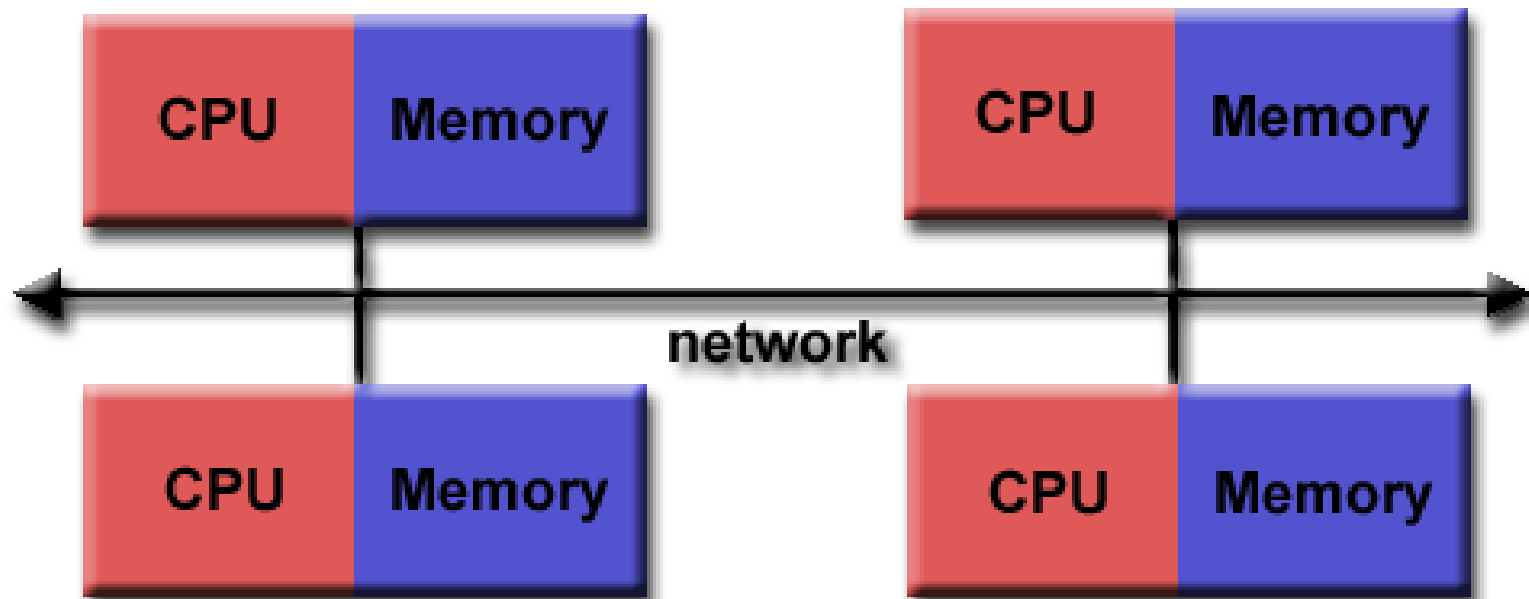
Преимущества:

- Высокая скорость доступа к данным
- Привычная модель программирования

Проблемы:

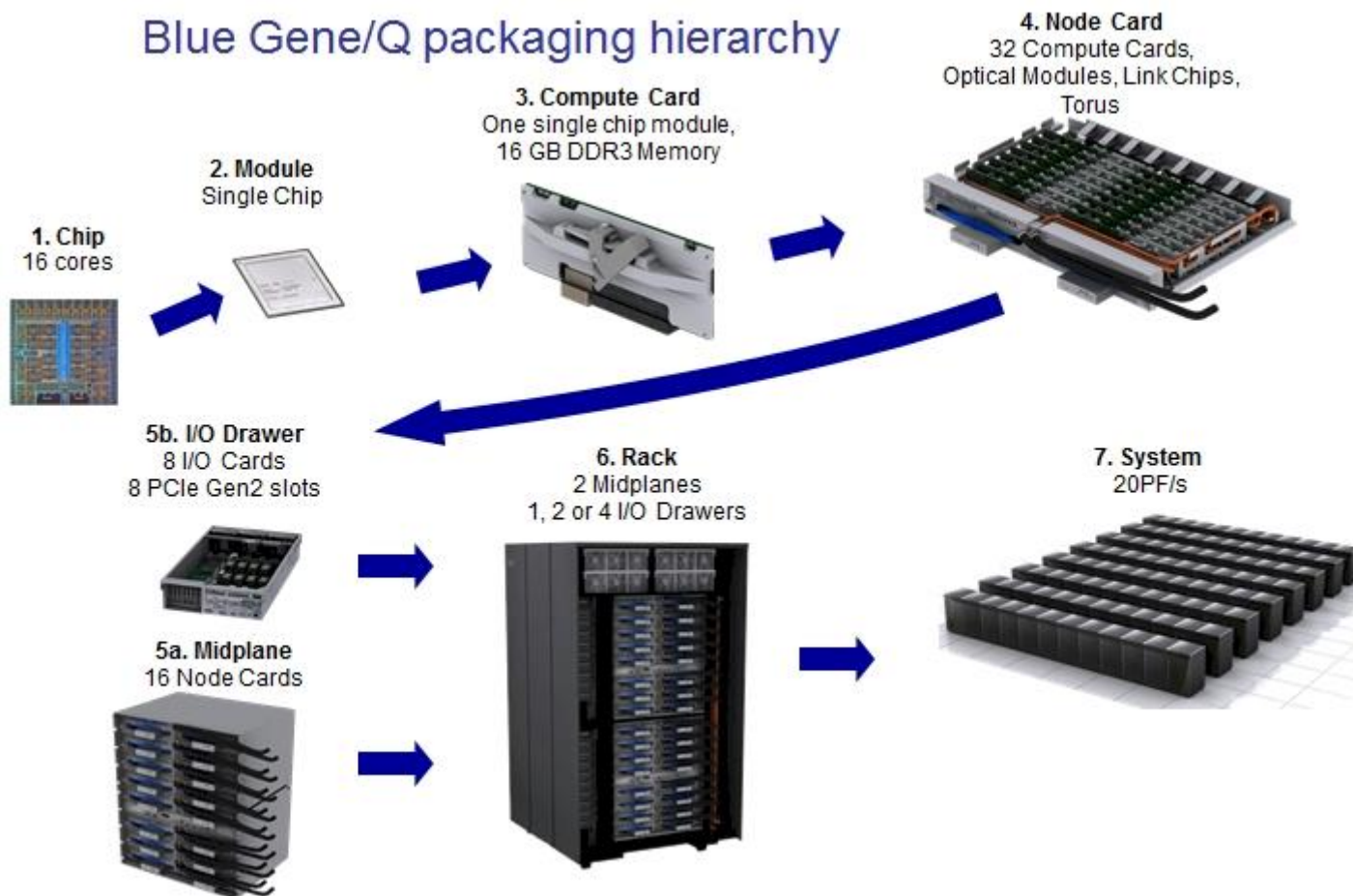
- Условия гонок при работе с общими данными
- Необходимо обеспечить когерентность кэшей
- Плохая масштабируемость

Системы с распределенной памятью

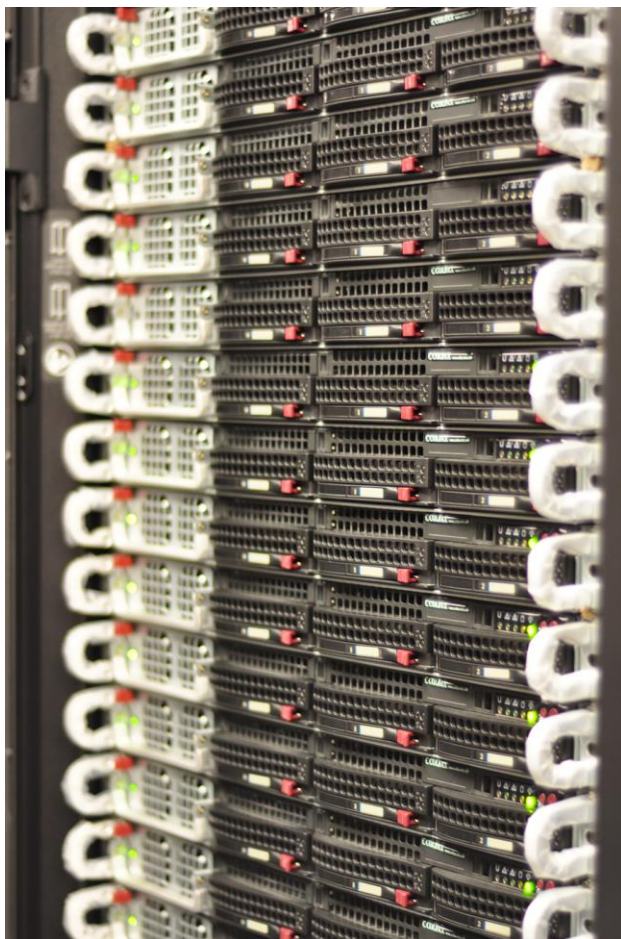


Massively parallel systems (MPP)

Blue Gene/Q packaging hierarchy



Вычислительные кластеры



Кластер из рабочих станций (CoW)



Распределенная память

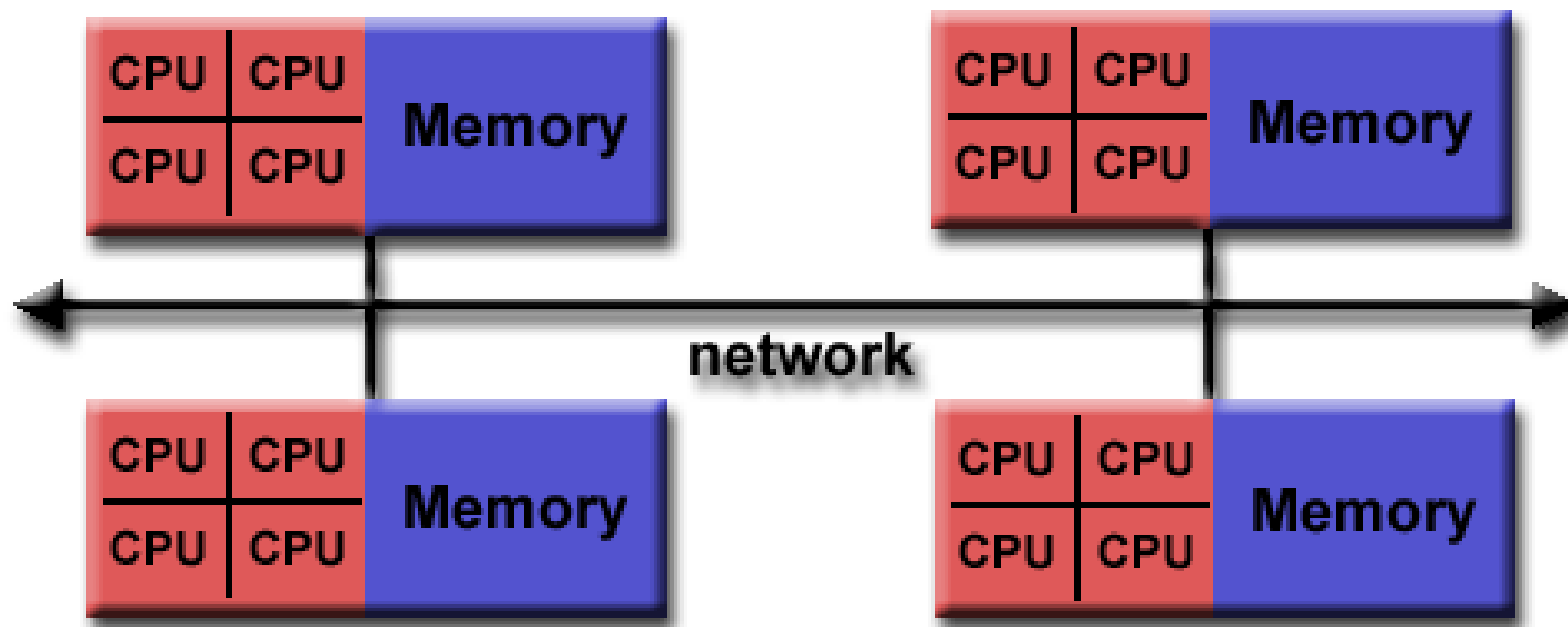
Преимущества:

- Высокая масштабируемость
- Нет условий гонок
- Не нужно обеспечивать когерентность кэшей

Проблемы:

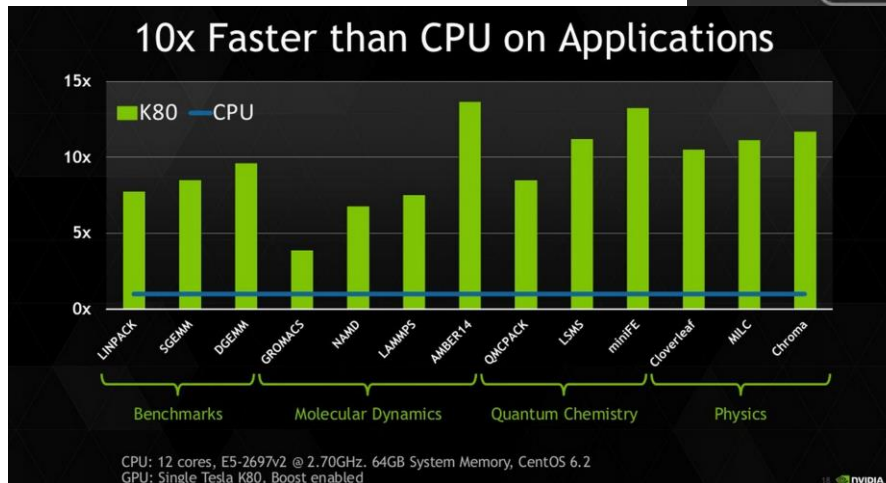
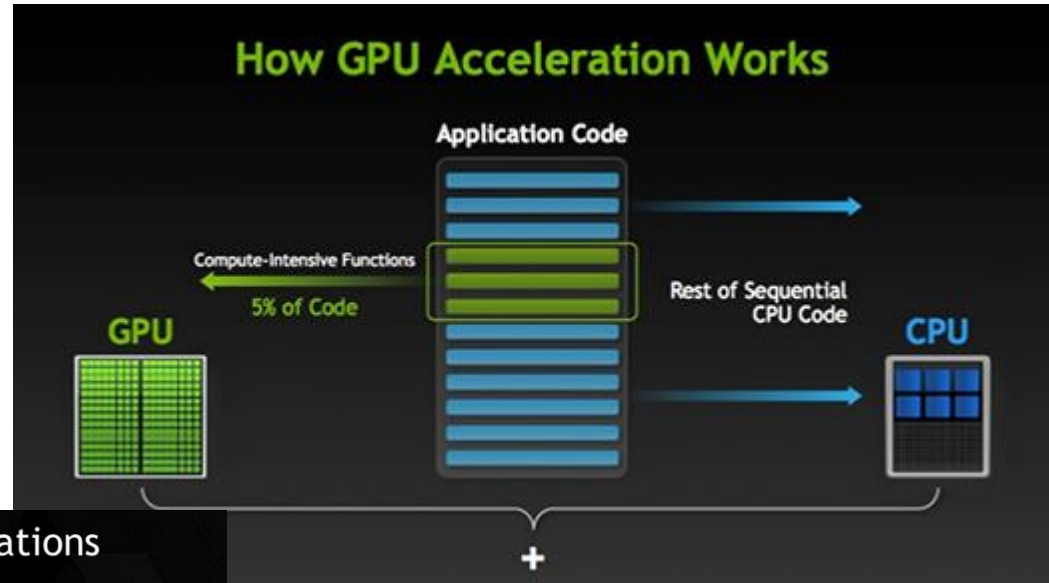
- Нет возможности напрямую обращаться к памяти другого узла
- Модель программирования на основе передачи сообщений
- Высокие задержки при передаче данных через сеть

Гибридные системы

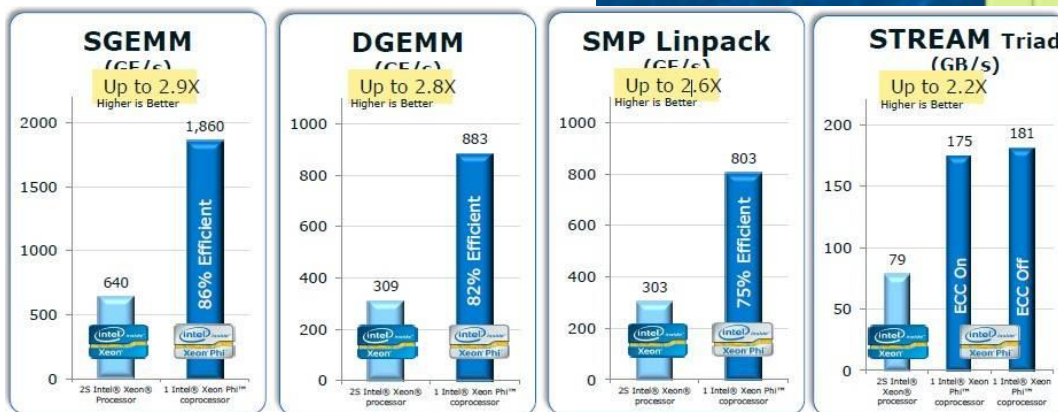
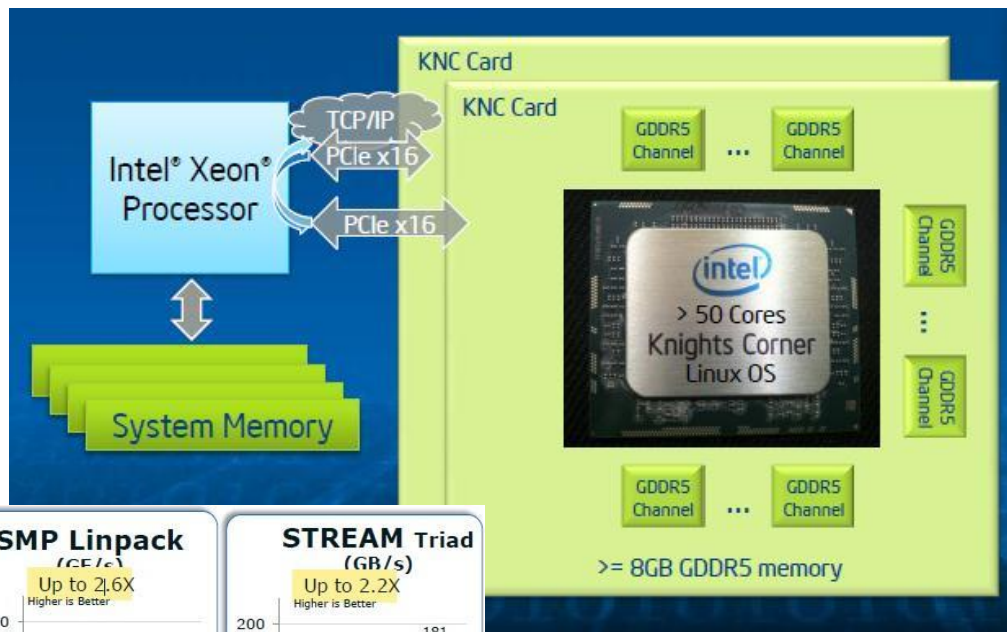


https://computing.llnl.gov/tutorials/parallel_comp/

Ускорители вычислений



Ускорители вычислений: Xeon Phi



TOP500 суперкомпьютеров мира

www.top500.org

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
2	Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P , NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	17,808
3	Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
4	Gyokou - ZettaScaler-2.2 HPC system, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 700Mhz , ExaScaler Japan Agency for Marine-Earth Science and Technology Japan	19,860,000	19,135.8	28,192.0	1,350
5	Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
6	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890

Измерение производительности

FLOPS:

- Floating-point operations per second
- Количество операций с плавающей точкой в секунду

Производительность:

- Пиковая – максимально возможная для процессора, на практике не достижима
- Реальная – производительность на реальных задачах (тест Linpack – перемножение матриц)

Теория параллельных вычислений

Производительность:

- Сокращение времени выполнения одной задачи (ускорение)
- Увеличение количества выполняемых за единицу времени задач (пропускная способность)
- Сокращение энергопотребления

Ускорение

$$S_p = \frac{T_1}{T_p}$$

- T_1 – время выполнение программы на одном исполнительном устройстве
- T_p – время выполнение программы на P исполнительных устройств

Эффективность

$$E_p = \frac{S_p}{P} = \frac{T_1}{P T_p}$$

- S_p – ускорение выполнения программы на P исполнительных устройств
- T_1 – время выполнение программы на одном исполнительном устройстве
- T_p – время выполнение программы на P исполнительных устройств

Ускорение

$$E = 1$$

- Идеальный случай

$$E < 1$$

- Типичный случай
- Последовательные части алгоритма
- Накладные расходы

$$E > 1$$

- Суперлинейное ускорение

Ускорение

Какое максимальное ускорение может быть достигнуто?

Закон Амдала

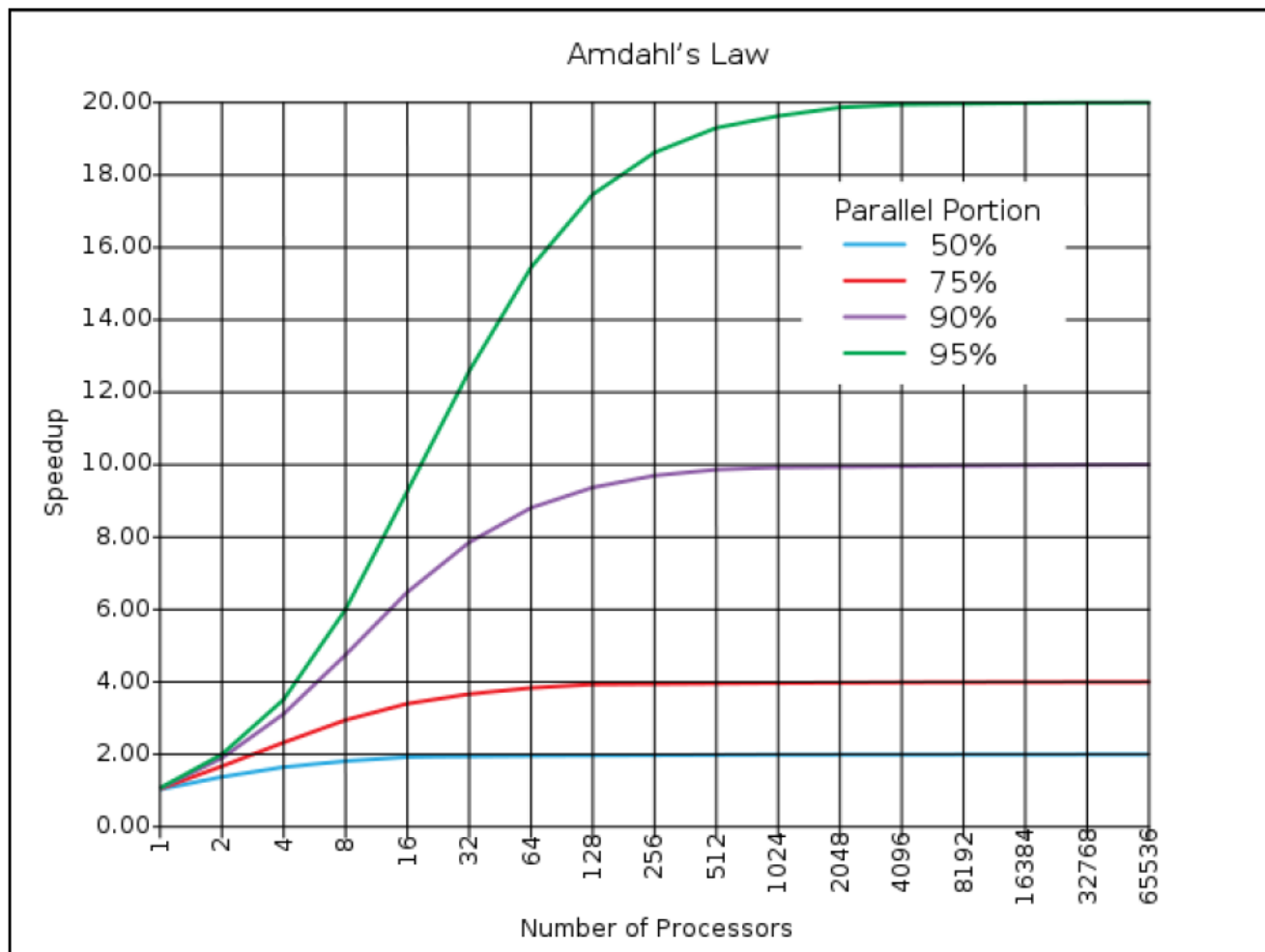
$$T_1 = W_{\text{ser}} + W_{\text{par}}$$

$$T_p \geq W_{\text{ser}} + W_{\text{par}} / P$$

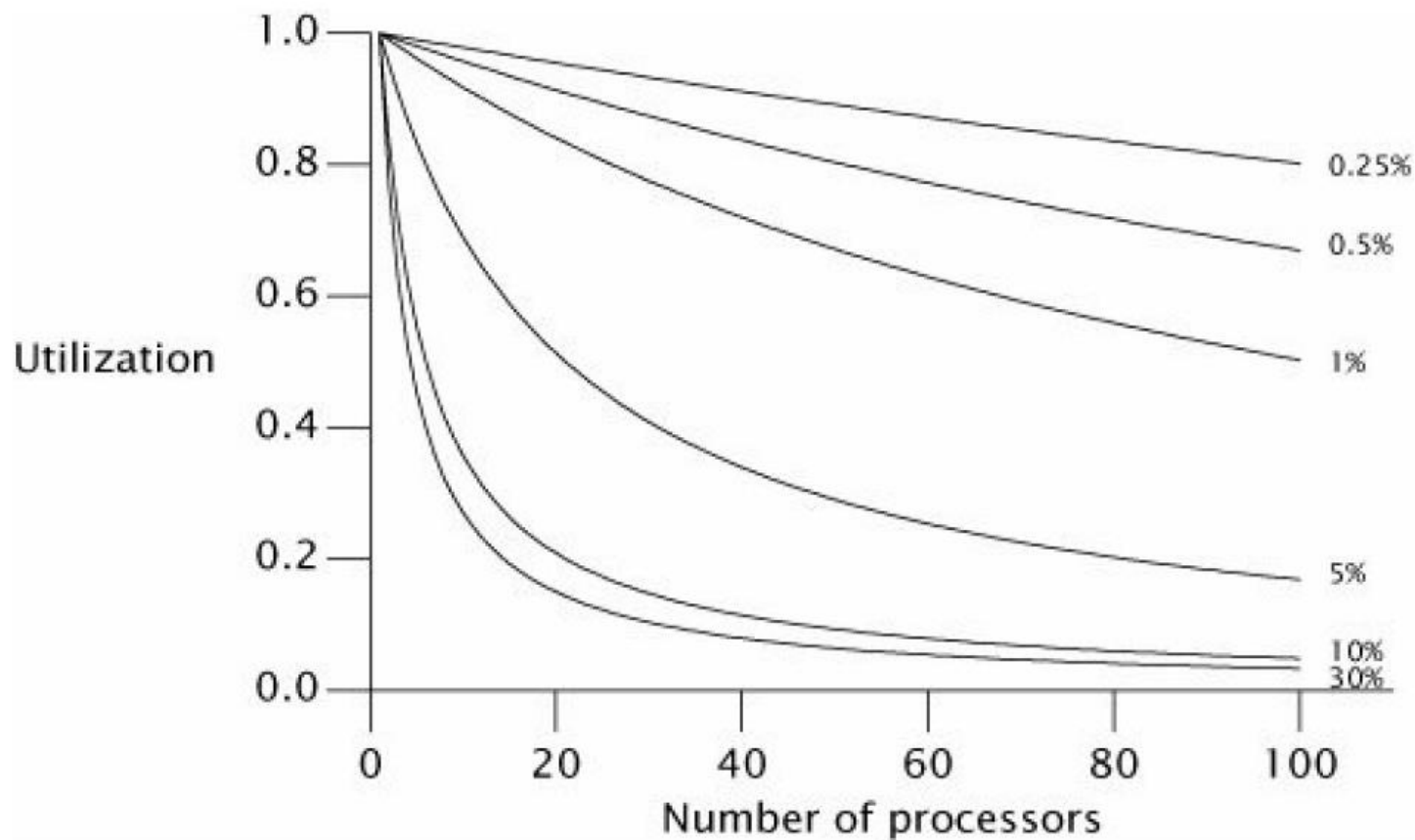
$$S_P \leq \frac{W_{\text{ser}} + W_{\text{par}}}{W_{\text{ser}} + W_{\text{par}}/P}.$$

$$\begin{aligned} W_{\text{ser}} &= f T_1, \\ W_{\text{par}} &= (1 - f) T_1. \end{aligned} \quad S_P \leq \frac{1}{f + (1 - f)/P}. \quad S_{\infty} \leq \frac{1}{f}.$$

Закон Амдала. Ускорение



Закон Амдала. Эффективность



Закон Густавсона-Барсиса

Джон Густавсон в Sandia National Labs получил ускорение для некоторых алгоритмов в 1000 раз

Закон Густавсона-Барсиса

- Speedup should be measured by scaling the problem to the number of processors, not by fixing the problem size.
- Ускорение нужно измерять, увеличивая объем данных с количеством процессоров, а не фиксируя размер задачи

Умножение матриц

- I/O – N^2
- Вычисления – N^3

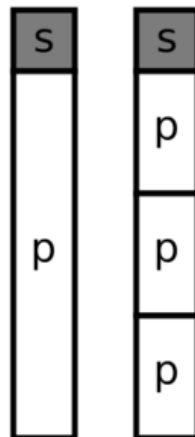
Закон Густавсона-Барсиса

Закон Густавсона-Барсиса

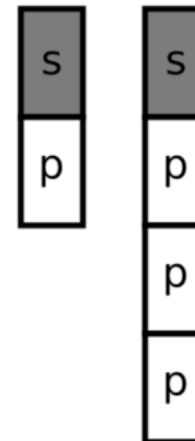
- $S_n = s + (1 - s) n$

Где

- s – доля последовательной части в программе
- n – количество процессоров



Закон Амдала



Закон Густавсона-Барсиса

Что будем изучать

OpenMP

- Системы с общей памятью

MPI

- Системы с распределенной памятью

Вопросы?