



Лекция 10. Организация функционирования распределённых вычислительных систем

Ткачёва Татьяна Алексеевна

преп. Кафедры вычислительных систем
Сибирский государственный университет
телекоммуникаций и информатики

Created by:

Пазников Алексей Александрович
к.т.н. доцент Кафедры вычислительных систем

Объект курса?

Распределённые вычислительные системы

Предмет курса?

*Модели и алгоритмы организации
функционирования*

Что хочет пользователь?

- Сокращение времени решения задач.
- Решение задач, требующих огромных объёмов памяти.

В то же время

совершенствование средств ВТ на основе модели вычислителя *не даст кардинального улучшения технических характеристик.*

Распределённая ВС – мультипроцессорные ВС с MIMD-архитектурой, в которых нет единого ресурса.

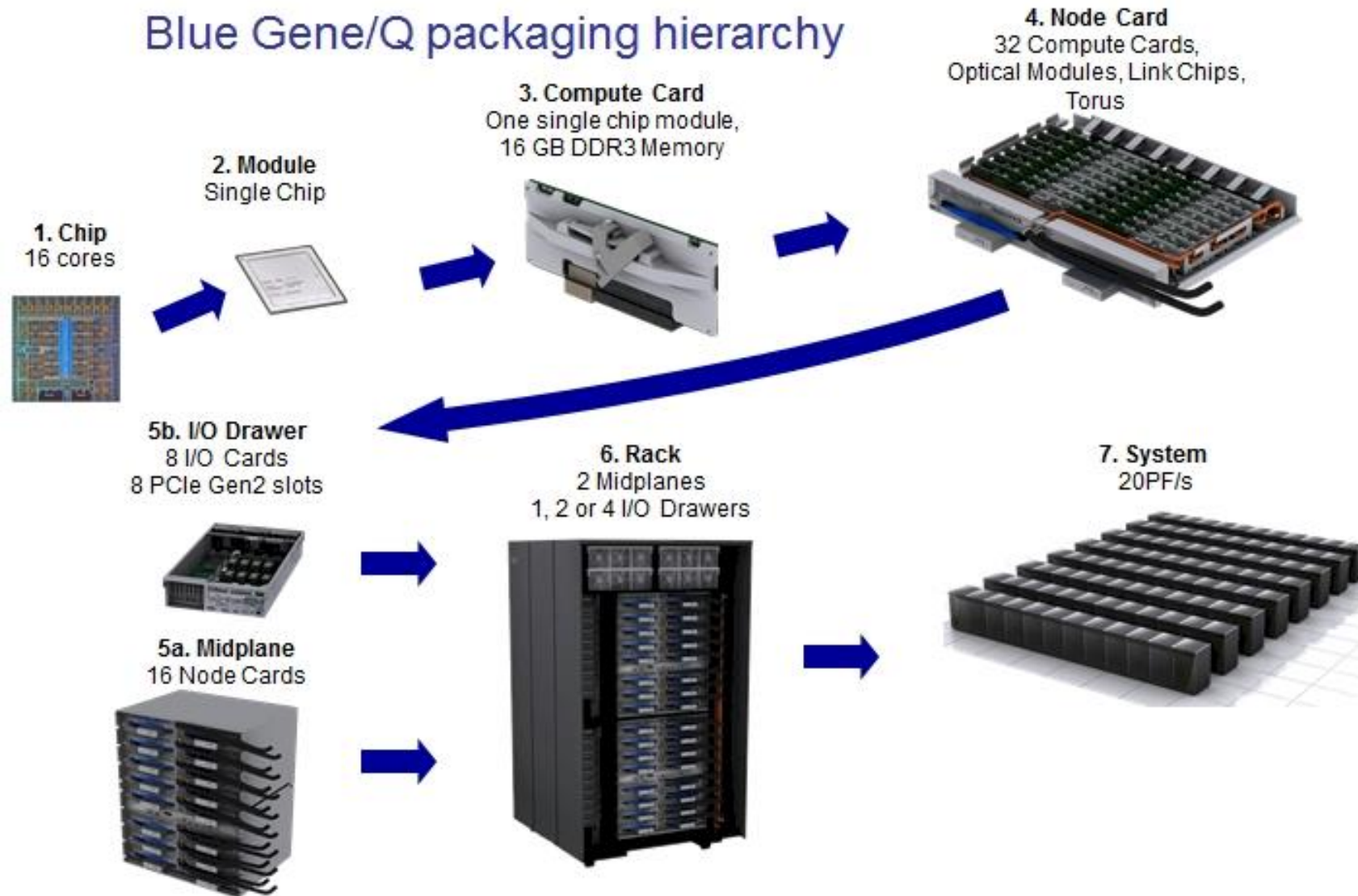
Представляется множеством взаимодействующих элементарных машин, оснащенных средствами коммуникаций и внешними устройствами.

Архитектурные особенности

- Иерархическая структура
- Мультиархитектурная организация
- Разнородность состава

Структура современных ВС

Blue Gene/Q packaging hierarchy



- **Распределённая ОС (GNU/Linux)**
- **Средства разработки параллельных программ**
(MPI, OpenMP, CUDA, Cray Chapel)
- **Системы управления ресурсами (TORQUE, SLURM)**
- **Системы обеспечения отказоустойчивости**
(самоконтроль, самодиагностика, контрольные точки)
- **Диспетчеры пространственно-распределённых ВС**
(GridWay, Pegasus)

Общая память

- POSIX Threads
- OpenMP
- Intel TBB
- Intel Cilk
- CUDA
- OpenCL

Распределённая память

- Sockets
- MPI
- PVM

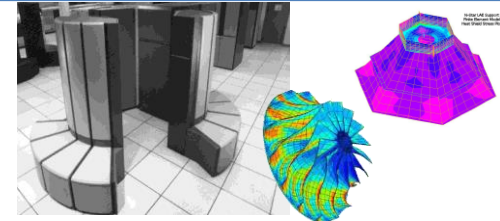
PGAS (*partitioned global address space*)

- Cray Chapel
- IBM X10
- Unified Parallel C
- HPF

Этапы развития распределённых ВС

1 ГФЛОПС – 1988: Cray Y-MP; 8 процессоров

- Задачи гидродинамики
- Fortran77 + векторизация



1 ТФЛОПС – 1998: Cray T3E; 1 024 процессоров

- Моделирование процессов магнетизма
- Fortran + MPI (Message Passing Interface)



1 ПФЛОПС – 2008: Cray XT5; 150 000 процессоров

- Моделирование сверхпроводимости
- C/C++/Fortran + MPI + векторизация



~143 ПФЛОПС – 2019: SUMMIT; 2397824 процессора

- Моделирование сверхпроводимости и прочие сложные штуки
- C/C++/Fortran + MPI + CUDA/OpenCL/OpenMP/OpenACC

Цели организации функционирования:

- Минимум времени решения задачи
- Максимальная надёжность
- Максимум прибыли
- Минимизации энергопотребления
- ...

Могут быть составные показатели.

I Монопрограммный режим

Решение одной сложной задачи – для решения задачи используются все ресурсы ВС.

II Мультипрограммный режим

Обработка набора задач – учитывается не только количество задач, но их параметры: число ветвей, время решения и др.

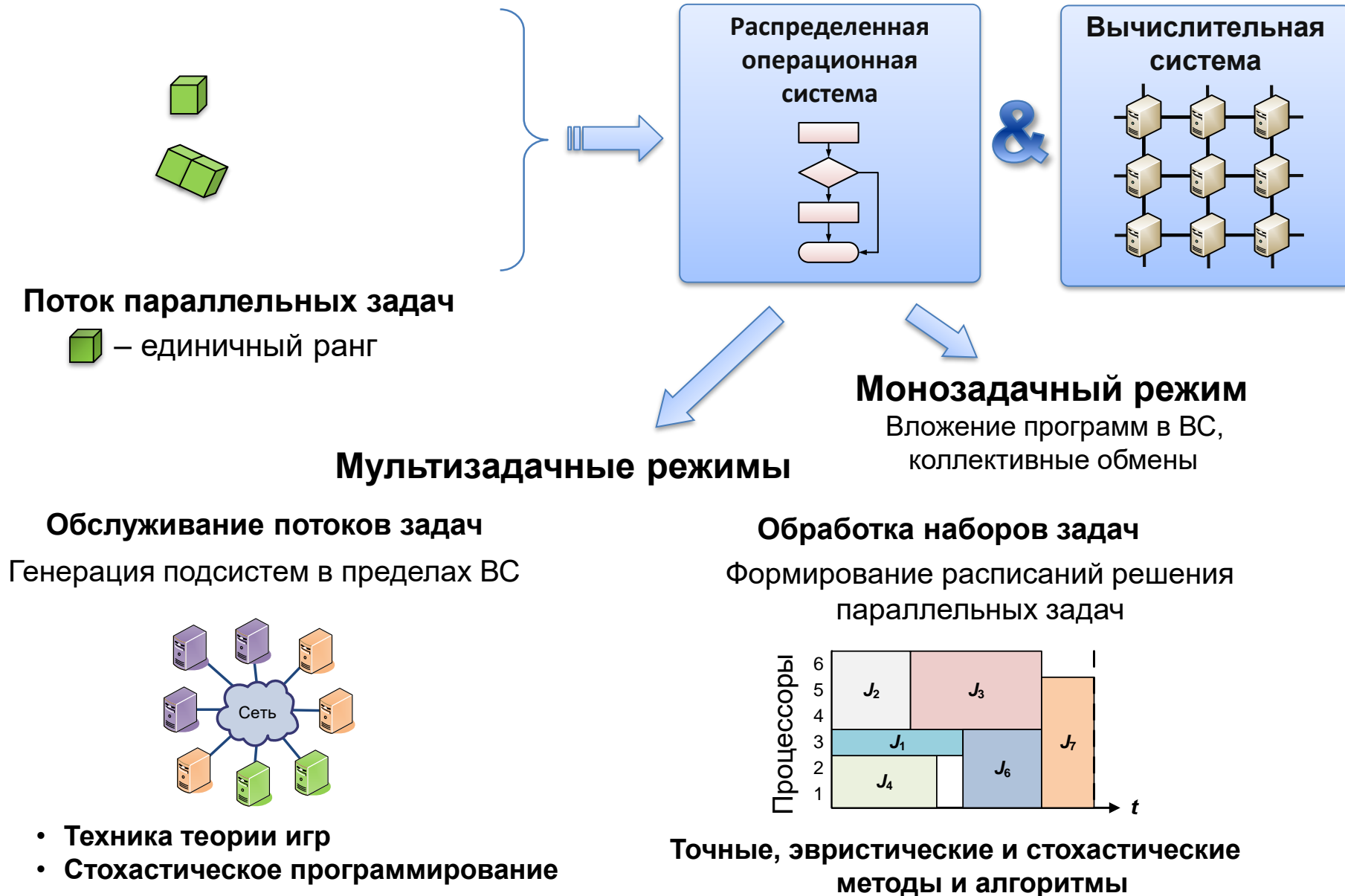
Обслуживание потока задач – задачи поступают в случайные моменты времени, их параметры случайны.

Сеанс работы пользователей в пакетном режиме

1. Поставить задачу в очередь.
2. Проверить состояние задачи.
3. Внести коррективы в задачу (её параметры).
4. Получить результат решения задачи.

Системы управления ресурсами распределённых ВС
(RMS - Resource Management System)

Параллельное мультипрограммирование



I Монопрограммный режим

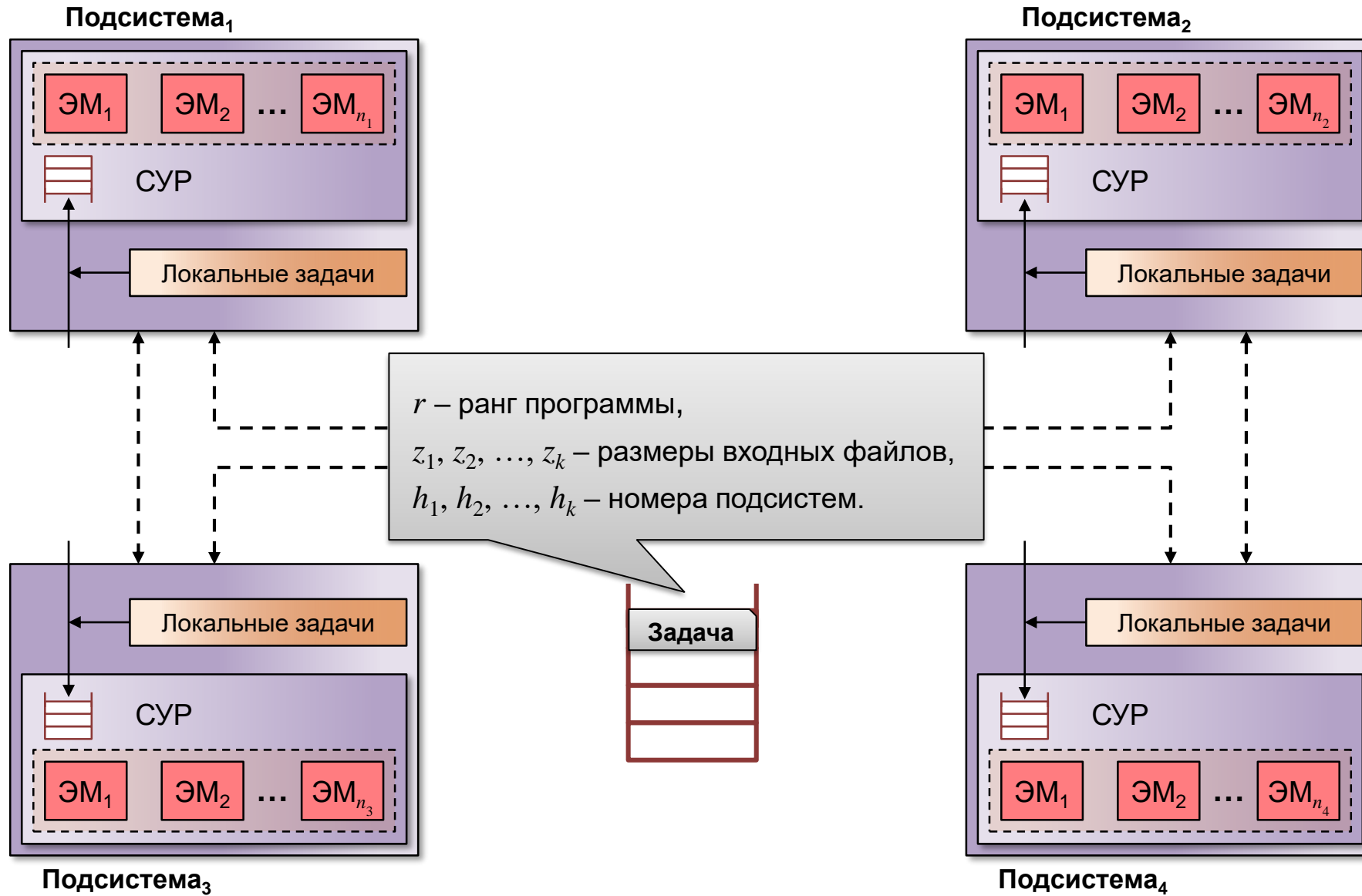
Решение одной сложной задачи – для решения задачи используются все ресурсы ВС.

II Мультипрограммный режим

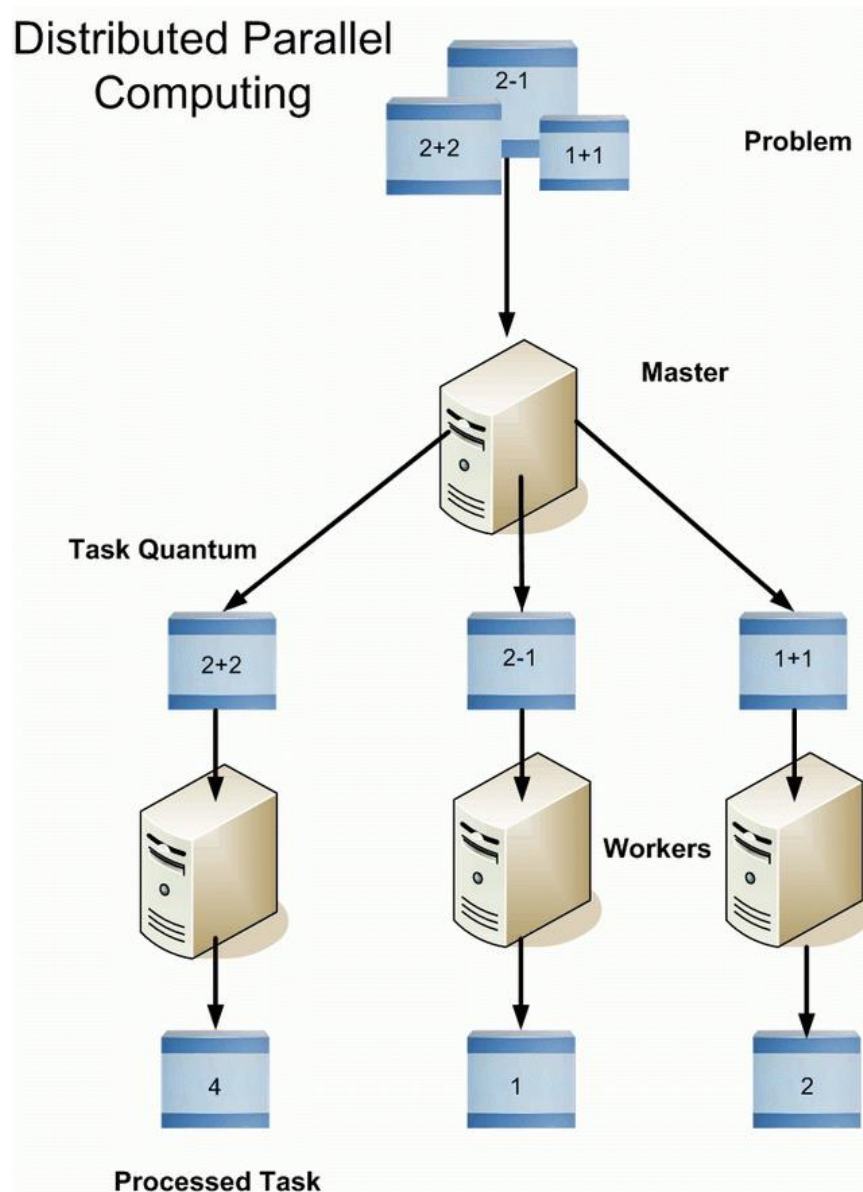
Обработка набора задач – учитывается не только количество задач, но их параметры: число ветвей, время решения и др.

Обслуживание потока задач – задачи поступают в случайные моменты времени, их параметры случайны.

Диспетчеризация параллельных задач



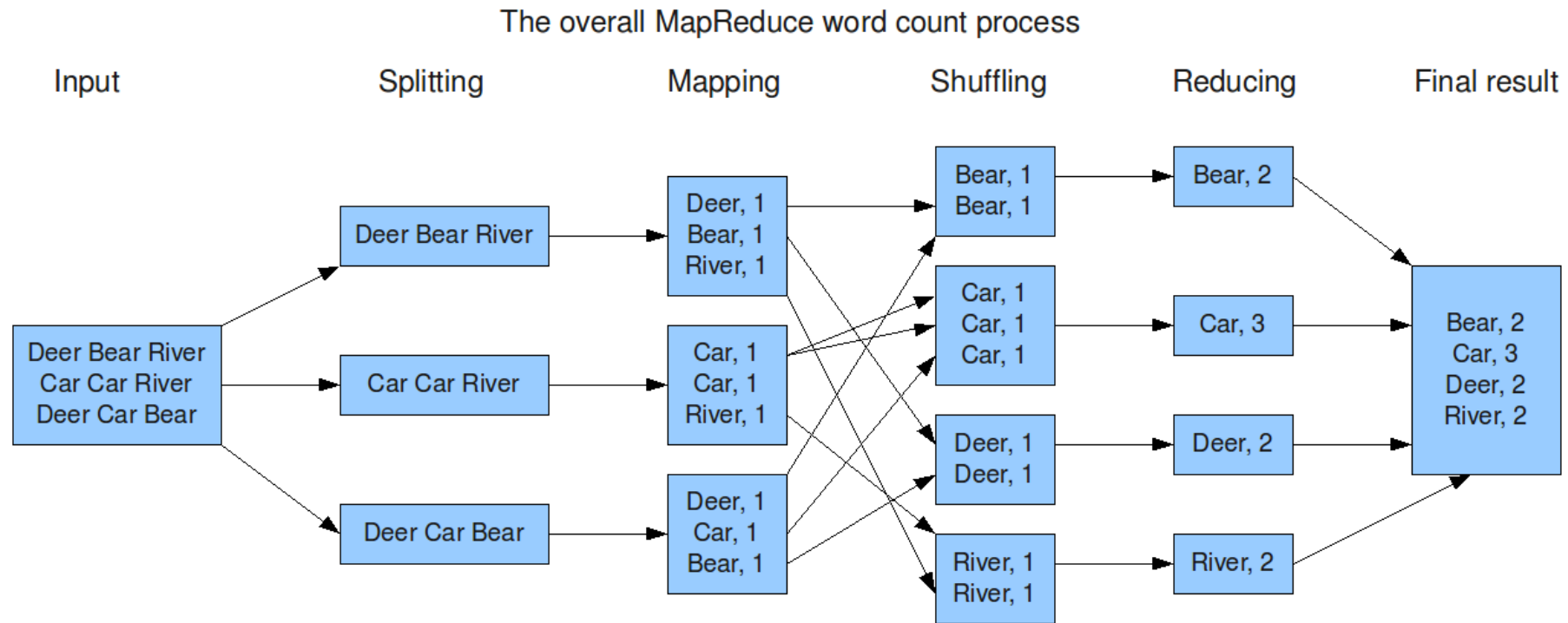
Распределённые вычисления (Вычисления высокой пропускной способности, НТС)



- живучесть, отказоустойчивость (миграция, контрольные точки)
- длительное время решения большого количества заданий
- слабо связанные задания
- большой объём задействованных ресурсов

*CONDOR, MOSIX, BOINC,
MapReduce*

Алгоритмы MapReduce



Высокопроизводительные вычисления (HPC)



- сильно связанные параллельные ветви – интенсивный обмен сообщениями – быстрые каналы связи
- единая точка доступа к ресурсам
- параллельные программы

Конец слайдов



П.Пикассо. «Герника»