



**Нижегородский государственный университет  
им. Н.И.Лобачевского**

*Факультет Вычислительной математики и кибернетики*

# **Microsoft Compute Cluster Server 2003**



# Содержание

---

- ❑ Общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003
- ❑ **Упражнения:**
  - Компиляция программы для запуска в CCS 2003,
  - Запуск последовательной задачи,
  - Запуск параллельного задания,
  - Запуск множества задач,
  - Запуск потока задач,
  - Задача определения характеристик сети передачи данных



# Общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003...

## □ Основные понятия, используемые в CCS 2003:...

- **Задание (job)** – запрос на выделение вычислительных ресурсов кластера для выполнения задач. Каждое задание может содержать одну или несколько задач,
- **Задача (task)** – команда или программа (в том числе, параллельная), которая должна быть выполнена на кластере. Задача не может существовать вне некоторого задания, при этом задание может содержать как несколько задач, так и одну,
- **Планировщик заданий (job scheduler)** – сервис, отвечающий за поддержание очереди заданий, выделение системных ресурсов, постановку задач на выполнение, отслеживание состояния запущенных задач,
- **Узел (node)** – вычислительный компьютер, включенный в кластер под управлением CCS 2003,



# Общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003...

- **Важнейшие понятия, используемые в CCS 2003:**
  - **Процессор (processor)** – один из, возможно, нескольких вычислительных устройств узла,
  - **Очередь (queue)** – список заданий, отправленных планировщику для выполнения на кластере. Порядок выполнения заданий определяется принятой на кластере политикой планирования,
  - **Список задач (task list)** – эквивалент очереди заданий для задач каждого конкретного задания. Порядок запуска задач определяется правилом FCFS (первыми будут выполнены задачи, добавленные в список первыми), если пользователь специально не задаст иной порядок



# Общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003...

- ❑ Планировщик заданий CCS 2003 работает как с последовательными, так и с параллельными задачами. *Последовательной* называется задача, которая использует ресурсы только одного процессора. *Параллельной* называется задача, состоящая из нескольких процессов (или потоков), взаимодействующих друг с другом для решения одной задачи.
- ❑ В случае использования MS MPI в качестве интерфейса передачи сообщений параллельные задачи необходимо запускать с использованием специальной утилиты **mpiexec.exe**, осуществляющей одновременный запуск нескольких экземпляров параллельной программы на выбранных узлах кластера



# Общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003...

- Непосредственным запуском задач занимается планировщик, а пользователь может лишь добавить задачу в очередь, так как время ее запуска выбирается системой автоматически в зависимости от того, какие вычислительные ресурсы свободны и какие задания ожидают в очереди выделения им ресурсов



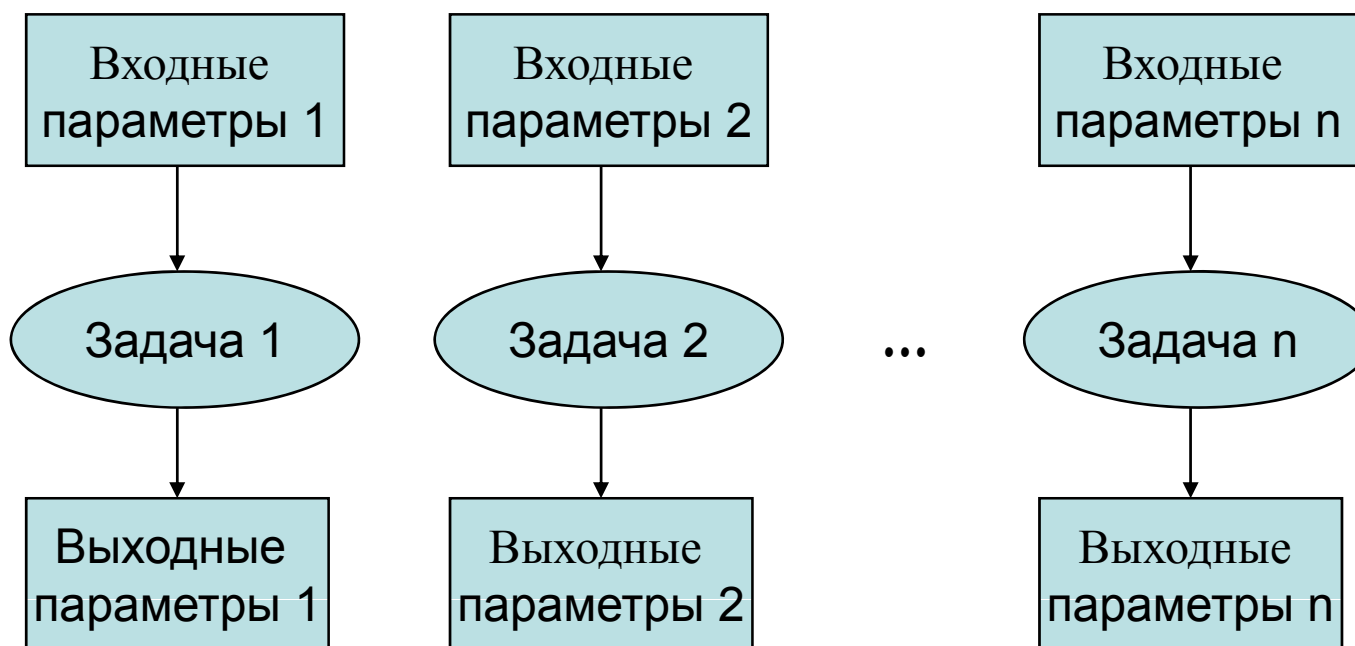
# Общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003...

- Для выполнения программы в CCS 2003 необходимо выполнить следующие действия:
  - *Создать* задание с описанием вычислительных ресурсов, необходимых для его выполнения,
  - *Определить* задачу. Задача задается при помощи той или иной команды, выполнение которой приводит к запуску на кластере последовательных или параллельных программ. Например, параллельная задача описывается при помощи команды **mpiexec.exe** с соответствующими параметрами (список узлов для ее запуска, имя параллельной программы, аргументы командной строки программы и др.),
  - *Добавить* задачу к созданному ранее заданию



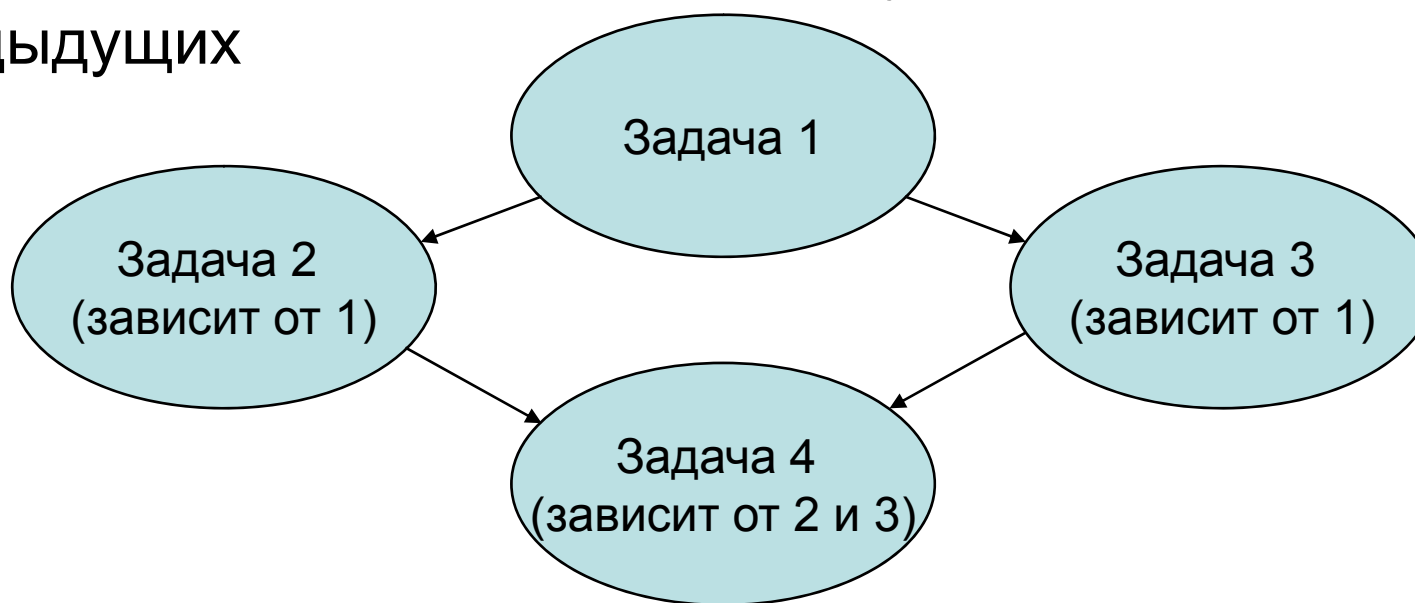
# Общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003...

- **Параметрическое множество задач (parametric sweep)** – одна и та же программа (последовательная или параллельная), несколько экземпляров которой запускается (возможно, одновременно) с разными входными параметрами и разными файлами вывода,



# Общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003

- **Поток задач (task flow)** – несколько задач (возможно, одна и та же программа с разными входными параметрами) запускаются в определенной последовательности. Последовательность запуска объясняется, например, зависимостью некоторых задач последовательности от результатов вычислений предыдущих



# Упражнение 1: Компиляция программы для запуска в CCS 2003...

- Задание 1 – Установка Microsoft Compute Cluster Pack SDK:
  - Для компиляции параллельных программ, работающих в среде MS MPI, необходимо установить **SDK (Software Development Kit)** – набор интерфейсов и библиотек для вызова MPI-функций



# Упражнение 1: Компиляция программы для запуска в CCS 2003...

- **Задание 2** – Настройка интегрированной среды разработки Microsoft Visual Studio 2005:...
  - **Путь до заголовочных файлов объявлений MPI**
    - Выберите пункт меню **Project->Project Properties**.
    - В пункте **Configuration Properties->C++->General->Additional Include Directories** введите путь до заголовочных файлов MS MPI: **<Директория установки CCS SDK>\Include**,
  - **Библиотечный файл с реализацией функций MPI**
    - Выберите пункт меню **Project->Project Properties**.
    - В пункте **Configuration Properties->C++->Linker->Input ->Additional Dependencies** введите название библиотечного файла **msmpi.lib**,



# Упражнение 1: Компиляция программы для запуска в CCS 2003...

- **Задание 2** – Настройка интегрированной среды разработки Microsoft Visual Studio 2005:
  - **Путь до библиотечного файла msmapi.lib.**
    - Выберите пункт меню **Project->Project Properties.**
    - В пункте **Configuration Properties->C++->Linker->General**  
**->Additional Library Directories** введите путь до библиотечного файла **msmapi.lib**:  
**<Директория установки CCS SDK>\Lib\i386** или  
**<Директория установки CCS SDK>\Lib\AMD64**  
в зависимости от используемой Вами архитектуры процессоров



# Упражнение 1: Компиляция программы для запуска в CCS 2003

- **Задание 3** – Компиляция параллельной программы в Microsoft Visual Studio 2005:
  - Скомпилируем в Microsoft Visual Studio 2005 параллельную программу вычисления числа Пи,
  - Создайте новый проект: выберите пункт меню **File->New->Project**. В окне выбора нового проекта выберите **консольное Win32 приложение (Other Languages->Visual C++->Win32->Win32 Console Application)**. Проведите настройки проекта из предыдущего задания



## Упражнение 2: Запуск последовательной задачи...

- **Задание 1** – Запуск программы через графический пользовательский интерфейс:...
  - Для запуска программы через графический интерфейс (программа **Compute Cluster Job Manager**) необходимо выбрать пункт меню **File->Submit Job**,
  - При задании параметров задания указывают **минимальное** и **максимальное** числа процессоров, необходимых для выполнения задания. Под максимальным здесь понимается оптимальное для задания число процессоров (именно столько будет выделено в случае низкой загрузки кластера). Гарантируется, что задание не начнет выполняться, если на кластере доступно менее минимального числа процессоров



## Упражнение 2: Запуск последовательной задачи...

- Задание 1 – Запуск программы через графический пользовательский интерфейс:
  - В случае запуска последовательной задачи необходимо создать одно задание и включить в него единственную последовательную задачу,
  - Ход выполнения заданий можно наблюдать в окне программы **Compute Cluster Job Manager**



## Упражнение 2: Запуск последовательной задачи...

- **Задание 2** – Запуск программы с использованием шаблона:
  - В программе **Compute Cluster Job Manager** есть возможность сохранения параметров задания в *xml* файл для последующего быстрого создания копии. Таким образом, Вы можете изменять интересующие параметры задания, оставляя другие неизменными, и экономя, таким образом, время на формирование задания,
  - Для сохранения задания необходимо дважды щелкнуть по нему левой кнопкой мыши в окне **Compute Cluster Job Manager** и нажать кнопку “**Save As Template...**”,
  - Для создания задания по шаблону – выполнить команду меню **File->Submit Job with Template**



## Упражнение 2: Запуск последовательной задачи

### □ Задание 3 – Запуск программы из командной строки:

- Часто бывает удобней управлять ходом выполнения заданий из командной строки. Microsoft Compute Cluster Server 2003 имеет в своем составе текстовые утилиты, предоставляющие полный контроль за ходом выполнения заданий на кластере,
- Для запуска последовательной задачи необходимо выполнить команды:
  - **job new /jobname:SerialPiCL /scheduler:s-cw-head** – создание нового задания,
  - **job add 26 /numprocessors:1 /scheduler:s-cw-head /stdout://s-cw-head/temp/serialpi.txt /workdir://s-cw-head/temp/ serialpi.exe 1000** – добавление новой задачи в задание,
  - **job submit /id:26 /scheduler:s-cw-head** – начало планирования задания



## Упражнение 3: Запуск параллельного задания

- Задание 1 – Запуск параллельного задания:
  - Запуск задач, разработанных для MS MPI, необходимо осуществлять с использованием специальной утилиты **mpiexec.exe**, которая принимает в качестве параметров имя параллельной программы, список узлов, на которых произойдет запуск, и параметры запускаемой программы,
  - Список узлов задается параметром “**-hosts**”. При этом в случае, если узлы должны быть выделены планировщиком автоматически, следует указать в качестве параметра переменную окружения **CCP\_NODES**, которая содержит список узлов кластера



## Упражнение 4: Запуск множества задач...

### □ Задание 1 – Запуск множества задач:...

- Под *параметрическим множеством задач* понимается серия запусков одной и той же программы с разными параметрами. В качестве примера можно привести запуск серии из нескольких сотен экспериментов по вычислению числа  $\Pi$  для исследования скорости сходимости метода к точному решению,
- Для добавления множества задач на вкладке **Tasks** диалога **Submit Job** есть кнопка **Add parametric Sweep**,



## Упражнение 4: Запуск множества задач

- Задание 1 – Запуск множества задач:
  - При добавлении множества задач пользователь задает границы изменения индекса и положение индекса в параметрах задач (индекс заменяется на “\*”). Каждой задаче в задании будет передан параметр командной строки с соответствующим значением индекса



## Упражнение 5: Запуск потока задач

### □ Задание 1 – Запуск потока задач:

- Поток задач используется тогда, когда для выполнения некоторой задачи в составе задания необходимы результаты работы других задач, что выдвигает требования к порядку их выполнения,
- Задать порядок выполнения задач в задании можно на вкладке **Tasks** диалога **Submit Job**, последовательно выбирая каждую задачу и задавая для нее список предшествующих задач (кнопка **Edit**, вкладка **Task Dependencies**, кнопка **Preceding Tasks**)



## Дополнительное упражнение: Задача определения характеристик сети передачи данных...

- ❑ Особенностью разработки программ для кластерных систем является необходимость учитывать не только характеристики отдельных компьютеров (прежде всего это производительность процессора и скорость памяти), но и характеристики сети передачи данных между ними. Чаще всего эти характеристики используются для построения теоретических оценок времени работы алгоритмов, что позволяет предсказывать время работы программ в зависимости от размера входных данных. Получение характеристик сети – отдельная задача, решаемая запуском специальных тестовых программ на конкретном имеющемся оборудовании. Необходимость проведения тестов для каждого конкретного кластера объясняется тем, что данные, предоставляемые поставщиком аппаратного обеспечения, могут сильно отличаться в зависимости от используемого программного обеспечения, настроек кластера или особенностей взаимодействия оборудования разных моделей и изготовителей



## Дополнительное упражнение: Задача определения характеристик сети передачи данных...

- Описание характеристик, определяющих производительность сети:
  - **Латентностью** (задержкой) называется время, затрачиваемое аппаратной и программной частью на обработку запроса отправки сетевого сообщения. То есть, это время с момента поступления команды на пересылку информации до начала ее непосредственной передачи. Обычно латентность указывается в **микросекундах (мкс)**,
  - **Пропускной способностью** сети называется максимальное количество информации, передаваемое между узлами сети за единицу времени. Обычно пропускная способность сети указывается в **мегабайтах в секунду (Мбайт/сек)** или **мегабитах в секунду (Мбит/сек)**,



## Дополнительное упражнение: Задача определения характеристик сети передачи данных...

### □ Общая характеристика алгоритма:...

- Идея алгоритма определения характеристик сети, используемого в тестах данной работы, состоит в последовательной пересылке между 2 узлами сообщений различной длины, используя средства установленной реализации MPI, и измерения времени, затрачиваемого на пересылку,
- Имея эти данные, пропускную способность можно определить, поделив длину переданного сообщения на затраченное на передачу время,
- Для минимизации погрешности передачу повторяют  $N$  раз и усредняют результат,



## Дополнительное упражнение: Задача определения характеристик сети передачи данных...

### □ Общая характеристика алгоритма:

- Оценка пропускной способности обычно увеличивается с ростом длины сообщения, стремясь к некоторому предельному значению. Обычно именно это предельное значение (или значение, полученное при пересылке большого сообщения) целесообразно использовать как оценку пропускной способности,
- За латентность обычно принимается время, затрачиваемое для пересылки сообщений нулевой длины



## Дополнительное упражнение: Задача определения характеристик сети передачи данных...

### □ Компиляция программы:...

- Последнюю версию тестового пакета **IMB** в составе **Intel Cluster Tools** Вы можете скачать на сайте Intel по адресу <http://www.intel.com/cd/software/products/asmo-na/eng/cluster/mapi/219848.htm>,
- Для того, чтобы скомпилировать IMB для ОС Windows, Вам придется самостоятельно создать проект в Microsoft Visual Studio 2005 аналогично проектам, создание которых было описано в данной работе. Либо Вы можете воспользоваться заготовкой проекта, которая входит в состав лабораторной работы (папка **IMB\_2\_3**),



## Дополнительное упражнение: Задача определения характеристик сети передачи данных...

### □ Компиляция программы:

- Скачать **тесты НИВЦ МГУ** можно по адресу <http://parallel.ru/ftp/tests/mpi-bench-suite.zip> .
- Для компиляции тестов в ОС Windows Вам также придется самостоятельно создать проект в Microsoft Visual Studio 2005, либо воспользоваться заготовкой проекта, входящей в состав лабораторной работы (папка **MGU\_tests**)



## Дополнительное упражнение: Задача определения характеристик сети передачи данных

### □ Выполнение программы:

- Тесты необходимо запускать на двух узлах сети по одному процессу на каждом узле. Таким образом, для запуска тестов в CCS 2003 необходимо указать в качестве требований для задания суммарное число процессоров на 2 узлах и выбрать эти узлы вручную



# Заключение

- ❑ Рассмотрена общая схема выполнения заданий под управлением Microsoft Compute Cluster Server 2003:
  - Компиляция программы для запуска в CCS 2003,
  - Запуск последовательной задачи,
  - Запуск параллельного задания,
  - Запуск множества задач,
  - Запуск потока задач
- ❑ Представлена схема выполнения экспериментов для определения характеристик сети передачи данных



# Контрольные вопросы

- ❑ Дайте определения терминам задание (job) и задача (task). В чем основные отличия?
- ❑ Какие основные настройки Microsoft Visual Studio 2005 необходимо произвести при компиляции параллельной программы для использования в среде MS MPI?
- ❑ В чем особенность запуска параллельных задач (скомпилированных для MS MPI) на кластере?
- ❑ Что такое параметрическое множество задач (parametric sweep)? Что такое поток задач (work flow)?
- ❑ Какие характеристики, определяющие производительность сети, Вы знаете? Дайте их определения

