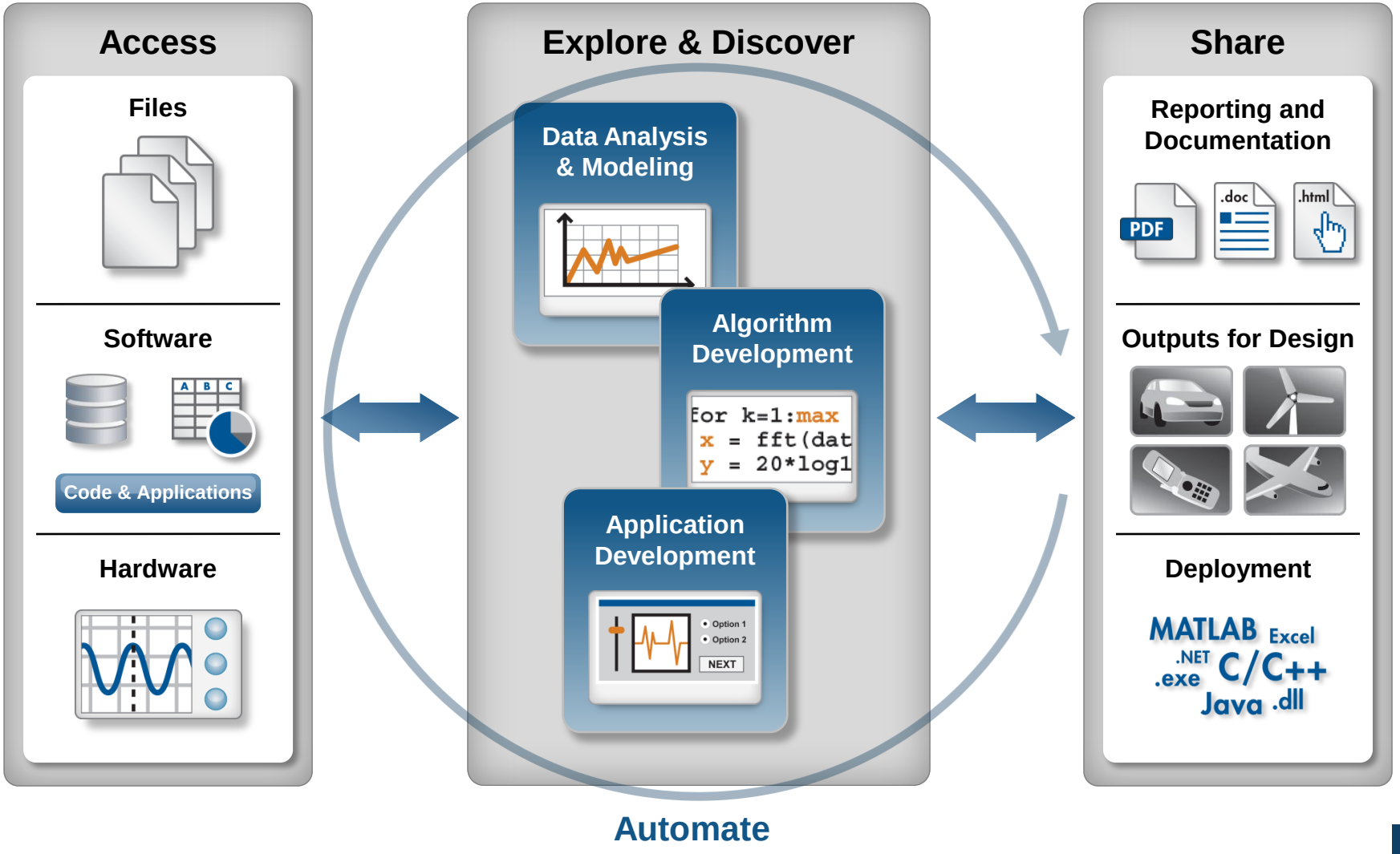
MathWorks

MATLAB&SIMULINK

Вычисления на GPU с помощью MATLAB

Денис Жегалин
Технический маркетинг
Департамент MathWorks, SoftLine

Data Analysis Tasks



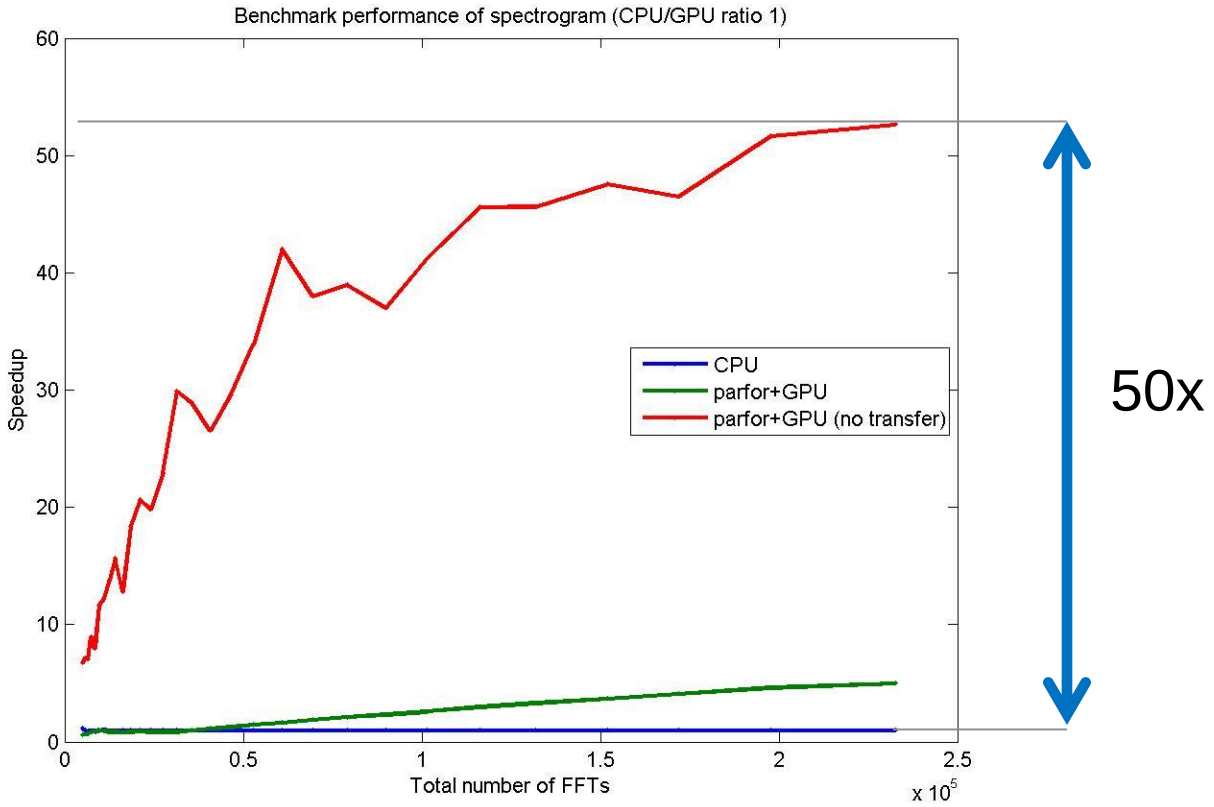
NEWS

MathWorks предоставляет поддержку GPU в MATLAB

- Портирование алгоритмов на графический процессор NVIDIA

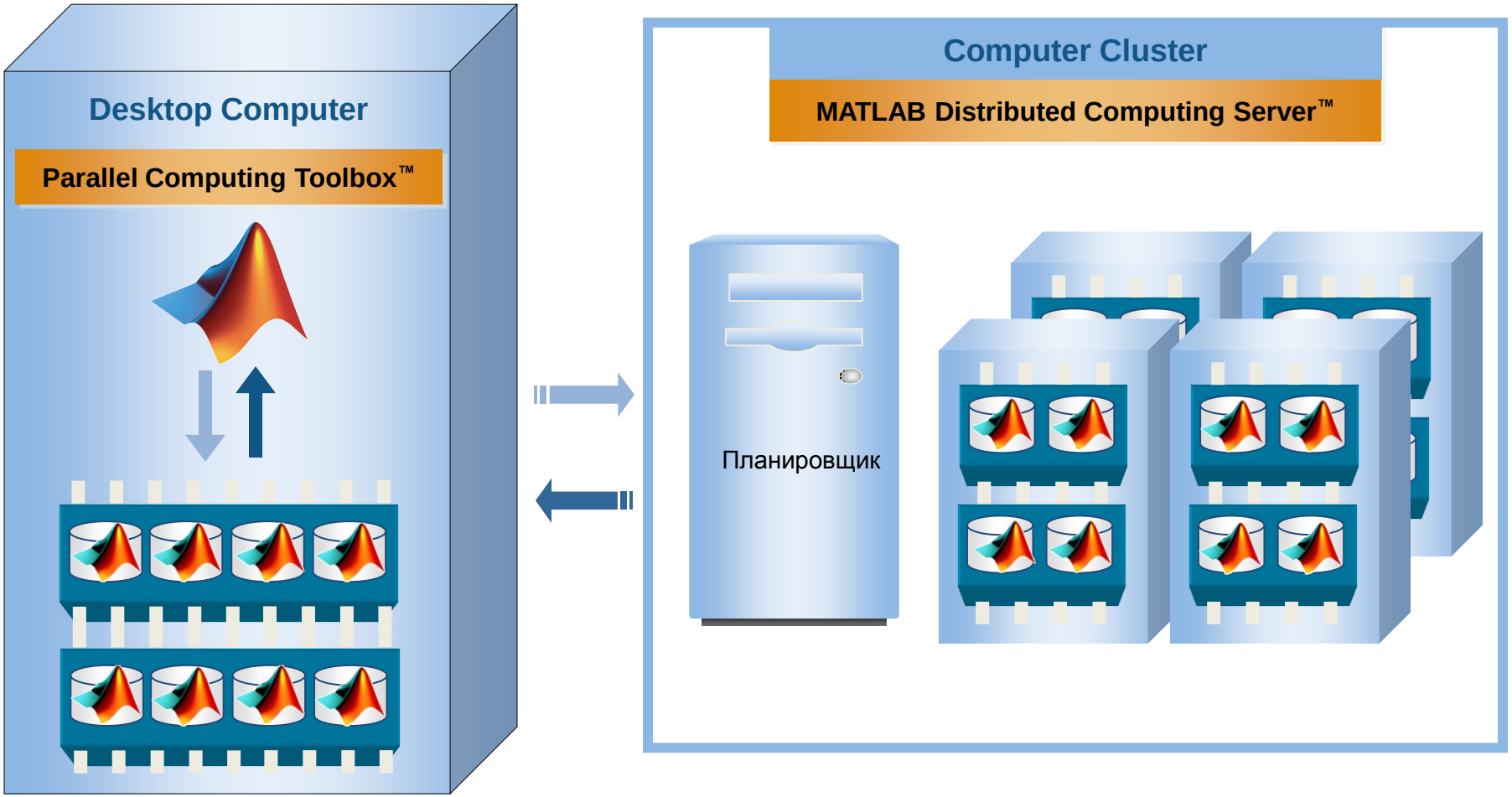
- Ускорение вычислений на GPU кластерах (локальных и распределенных) с помощью специальных инструментов для распараллеливания.

Спектрограмма показывает 50ти кратное увеличение скорости вычислений на GPU кластере

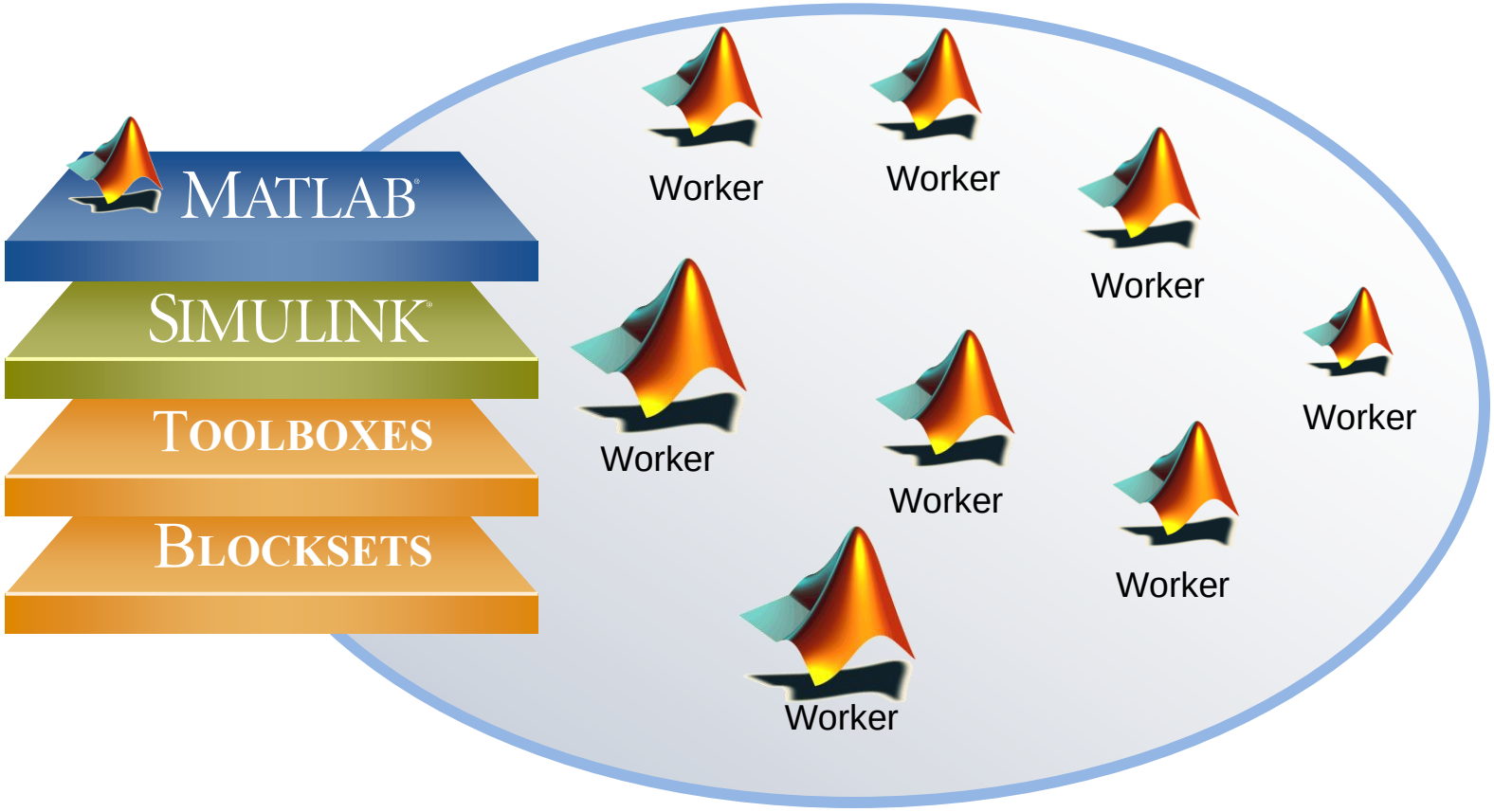


Параллельные вычисления в MATLAB

Инструменты и терминология



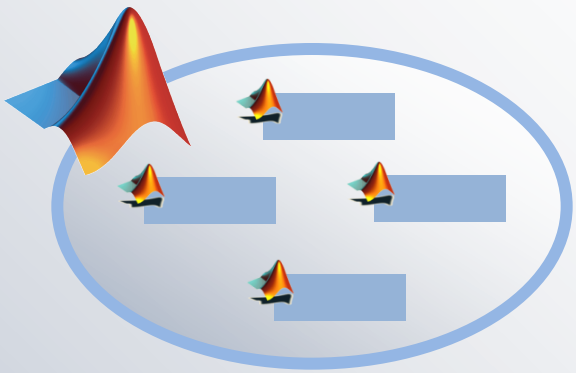
MATLAB: Дополнительные работники



Parallel Computing позволяет инженерам …

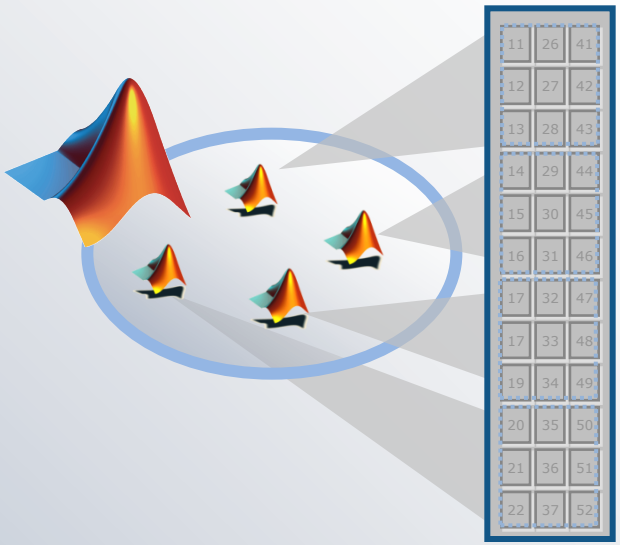
Larger Compute Pool

Увеличить скорость вычислений



Larger Memory Pool

Работать с большими объемами данных



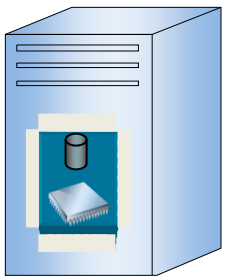
Возможности

Встроенная поддержка для расширений MATLAB

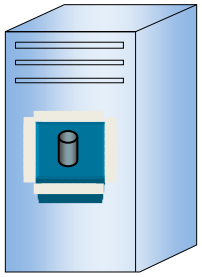
Поддержка высокоуровневых программных конструкций:
parfor, distributed arrays, batch

Низкоуровневый контроль вычислений:
Jobs/Tasks, spmd, MPI-interface

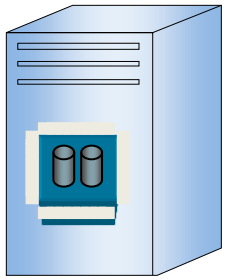
Эволюция вычислительных комплексов



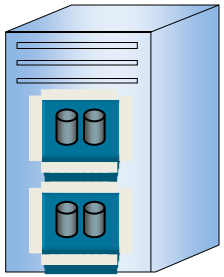
GPU



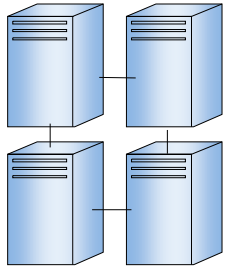
Single processor



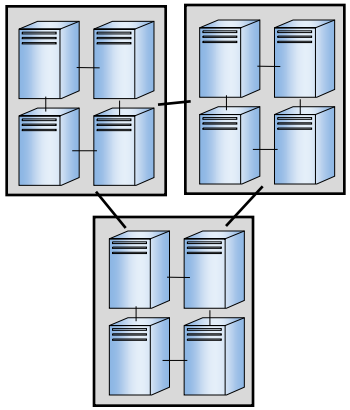
Multicore



Multiprocessor

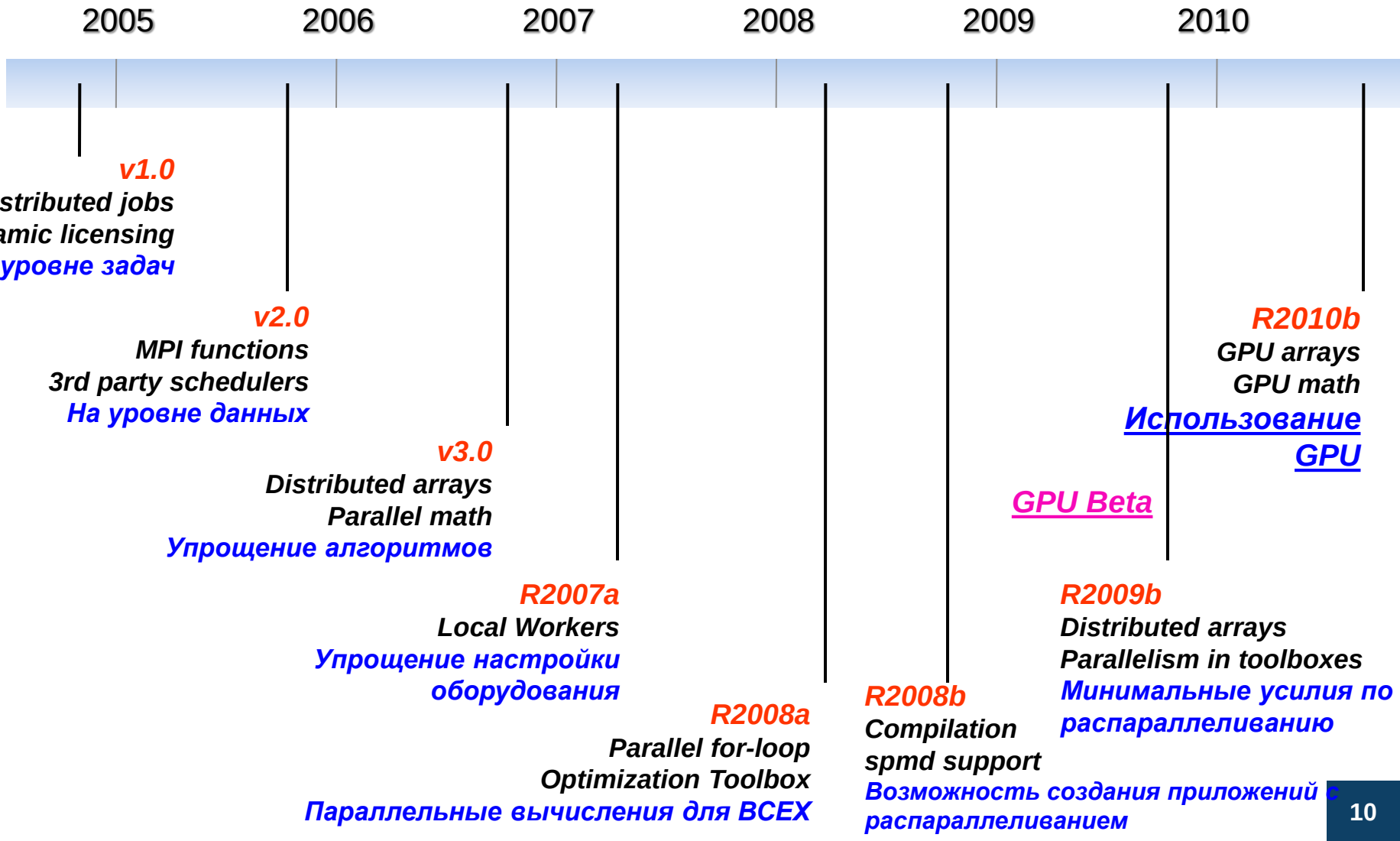


Cluster



Grid,
Cloud

Развитие параллельных вычислений в MATLAB



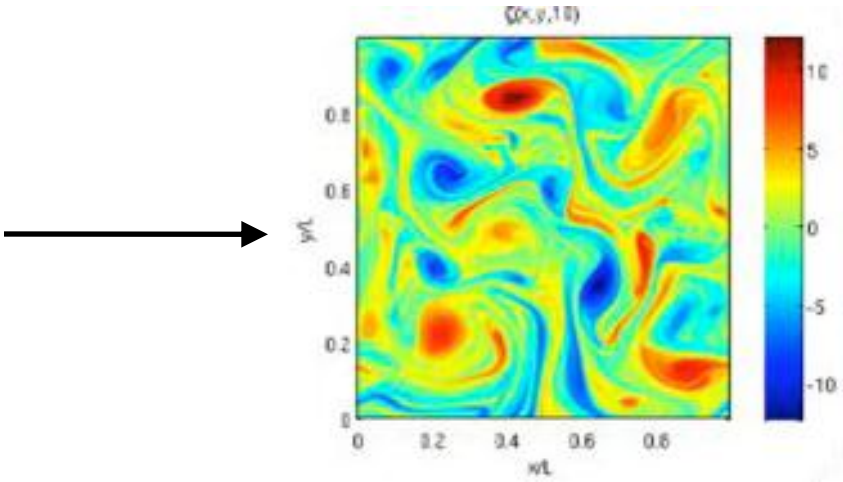
Что такое GPU ?

- Массив вычислителей
  - Сотни ядер на одном графическом процессоре
  - Ядра GPU дополняют ядра CPU
- Выделенная высокоскоростная память

* Parallel Computing Toolbox требует NVIDIA GPU с вычислительной способностью 1.3 ивыше, включая NVIDIA Tesla 10-серий и 20-серий.

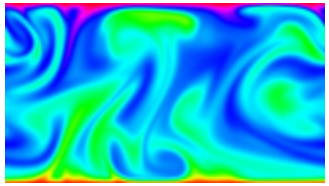
История развития GPU

3D Gaming & CAD ⟶ Scientific Computing



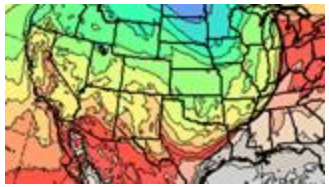
Области применения GPU

Список задач из CUDA Community Showcase:

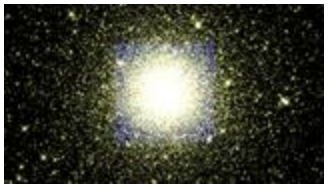


Динамика жидкостей

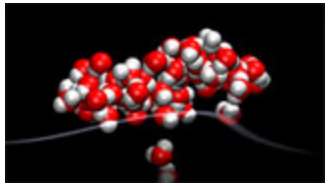
Вычислительные финансы



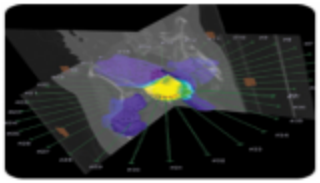
Моделирование погодных условий



Решение задачи N тел



Молекулярное моделирование



Цифровая обработка сигналов

http://www.nvidia.com/object/cuda_showcase_html.html

Почему GPU и почему сейчас?

- Вычисления
  - С двойной и одинарной точностью (single/double)
- Операции соответствуют стандартам IEEE
- Кроссплатформенность

Расчеты на GPU с помощью MATLAB

На одном или нескольких GPU:

1) Использование GPU массивов и встроенных функций MATLAB

2) Разработка собственных алгоритмов на GPU

3) Создание CUDA ядер на базе написанного C кода

Пример использования:

GPU массив:

```
>> A = someArray(1000, 1000);
>> G = gpuArray(A); % Push to GPU memory
…
>> F = fft(G);
>> x = G\b;
…
>> z = gather(x); % Bring back into MATLAB
```

+100 функций, поддерживающих GPU массивы

- `fft, fft2, ifft, ifft2`
- Умножение матриц (`A*B`)
- Левое деление матриц (`A\b`)
- LU разложение
- `'  .'`
- `abs, acos, …, minus, …, plus, …, sin, …`

Демонстрация

Пример:

CPU:

```
D = data;
iterations = 2000; % # of parallel iterations
stride = iterations*step; %stride of outer loop

M = ceil((numel(x)-W)/stride);%iterations needed
o = cell(M, 1); % preallocate output

for i = 1:M
    % What are the start points
    thisSP = (i-1)*stride:step: …
        (min(numel(x)-W, i*stride)-1);

    % Move the data efficiently into a matrix
    X = copyAndWindowInput(D, window, thisSP);

    % Take lots of fft's down the colmuns
    X = abs(fft(X));

    % Return only the first part to MATLAB
    o{i} = X(1:E, 1:ratio:end);

end
```

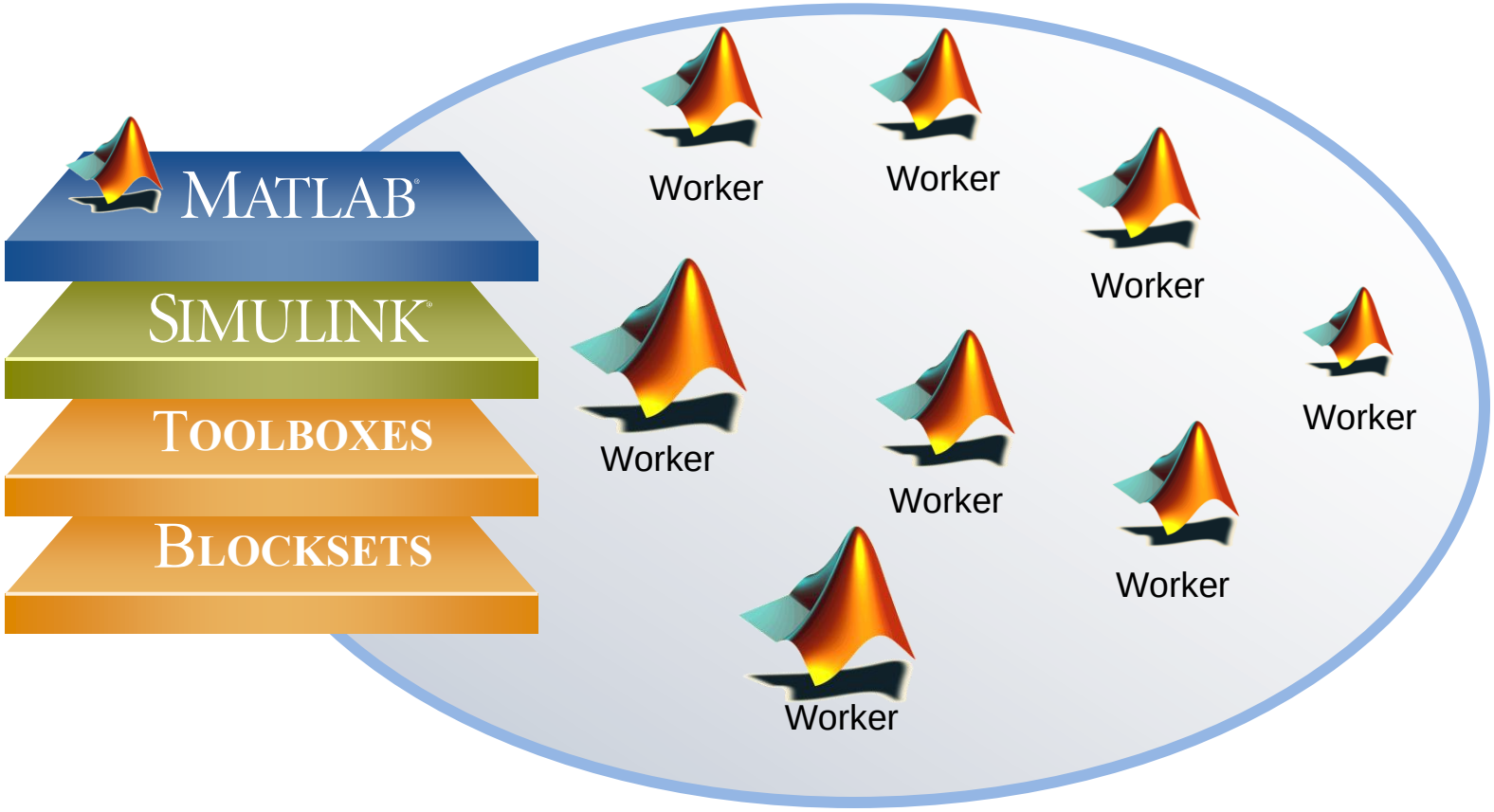
Пример:

CPU → GPU

```
D = data;
iterations = 2000; % # of parallel iterations
stride = iterations*step; %stride of outer loop

M = ceil((numel(x)-W)/stride);%iterations needed
o = cell(M, 1); % preallocate output

for i = 1:M
    % What are the start points
    thisSP = (i-1)*stride:step: …
        (min(numel(x)-W, i*stride)-1);

    % Move the data efficiently into a matrix
    X = copyAndWindowInput(D, window, thisSP);

    % Take lots of fft's down the colmuns
    X = abs(fft(X));

    % Return only the first part to MATLAB
    o{i} = X(1:E, 1:ratio:end);

end
```

```
D = gpuArray(data);
iterations = 2000; % # of parallel iterations
stride = iterations*step; %stride of outer loop

M = ceil((numel(x)-W)/stride);%iterations needed
o = cell(M, 1); % preallocate output

for i = 1:M
    % What are the start points
    thisSP = (i-1)*stride:step: ...
        (min(numel(D)-W, i*stride)-1);

    % Move the data efficiently into a matrix
    X = copyAndWindowInput(D, window, thisSP);

    % Take lots of fft's down the colmuns
    X = gather(abs(fft(X)));

    % Return only the first part to MATLAB
    o{i} = X(1:E, 1:ratio:end);

end
```

MATLAB: Дополнительные работники



Пример:

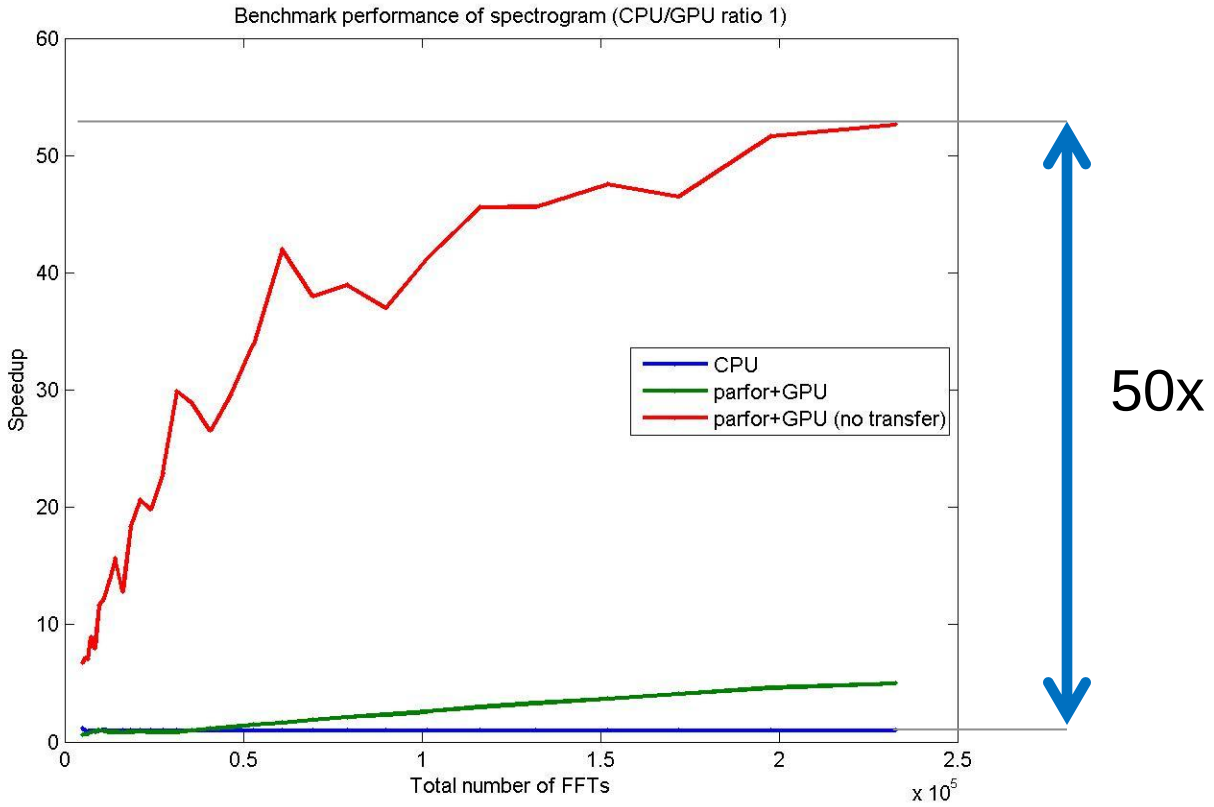
CPU → GPU → несколько GPU

```
D = gpuArray(data);
iterations = 2000; % # of parallel iterations
stride = iterations*step; %stride of outer loop

M = ceil((numel(x)-W)/stride);%iterations needed
o = cell(M, 1); % preallocate output

for i = 1:M
    % What are the start points
    thisSP = (i-1)*stride:step: ...
        (min(numel(D)-W, i*stride)-1);

    % Move the data efficiently into a matrix
    X = copyAndWindowInput(D, window, thisSP);

    % Take lots of fft's down the colmuns
    X = gather(abs(fft(X)));

    % Return only the first part to MATLAB
    o{i} = X(1:E, 1:ratio:end);

end
```

```
D = gpuArray(data);
iterations = 2000; % # of parallel iterations
stride = iterations*step; %stride of outer loop

M = ceil((numel(x)-W)/stride);%iterations needed
o = cell(M, 1); % preallocate output

parfor i = 1:M
    % What are the start points
    thisSP = (i-1)*stride:step: ...
        (min(numel(D)-W, i*stride)-1);

    % Move the data efficiently into a matrix
    X = copyAndWindowInput(D, window, thisSP);

    % Take lots of fft's down the colmuns
    X = gather(abs(fft(X)));

    % Return only the first part to MATLAB
    o{i} = X(1:E, 1:ratio:end);

end
```

Тест-Драйв

Спектрограмма показывает 50и кратное увеличение скорости вычислений на GPU кластере



Какое оборудование поддерживается?

- GPU NVIDIA соответствующие спецификации CUDA 1.3
- Список оборудования: http://www.nvidia.com/object/cuda_gpus.html

Контактная информация департамента Mathworks

- Softline: www.sl-matlab.ru
- matlab.exponenta.ru
- Mathworks: www.mathworks.com
- E-mail: matlab@softline.ru
- Phone: +7 (495) 232 00 23 доб. 0609