Параллельные алгоритмы 20250225_05

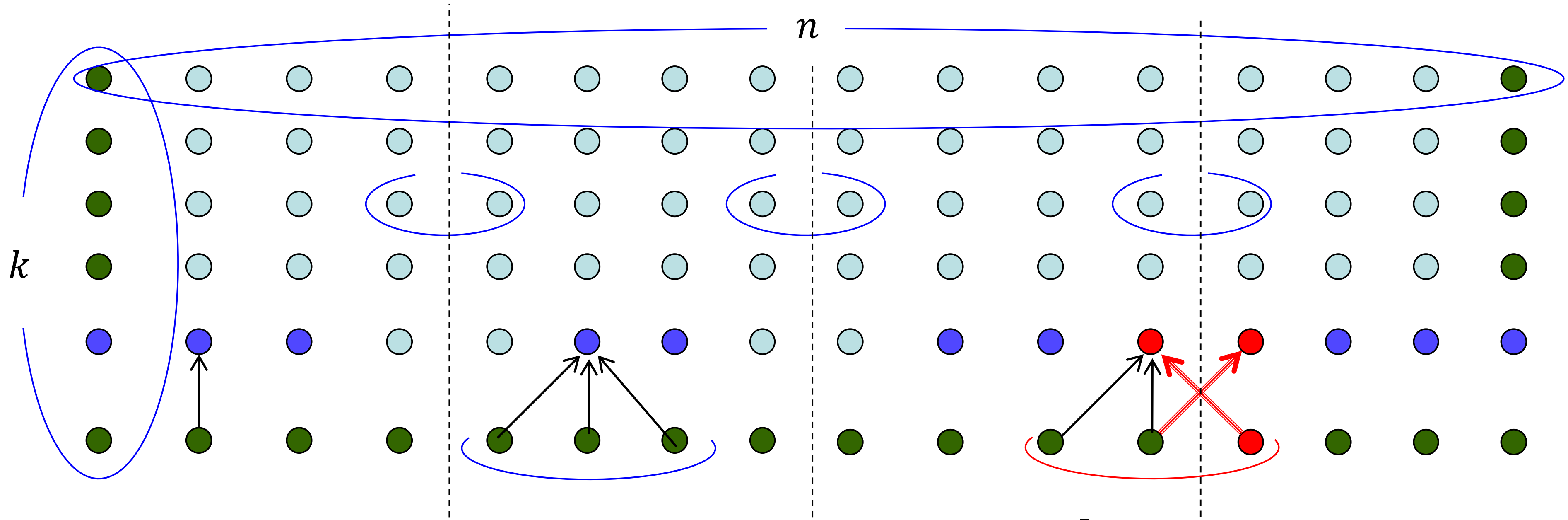
Базовые параллельные алгоритмы (2)

Якобовский Михаил Владимирович

Основные характеристики параллельной программы

- Ускорение $S_p = \frac{T_1}{T_p}$
- Эффективность $E_p = \frac{S_p}{p}$
- Предел масштабируемости – минимальное число процессоров при котором достигается максимальное ускорение
  – Число выполняемых операций
  – Время выполнения
  – Объём обрабатываемых данных

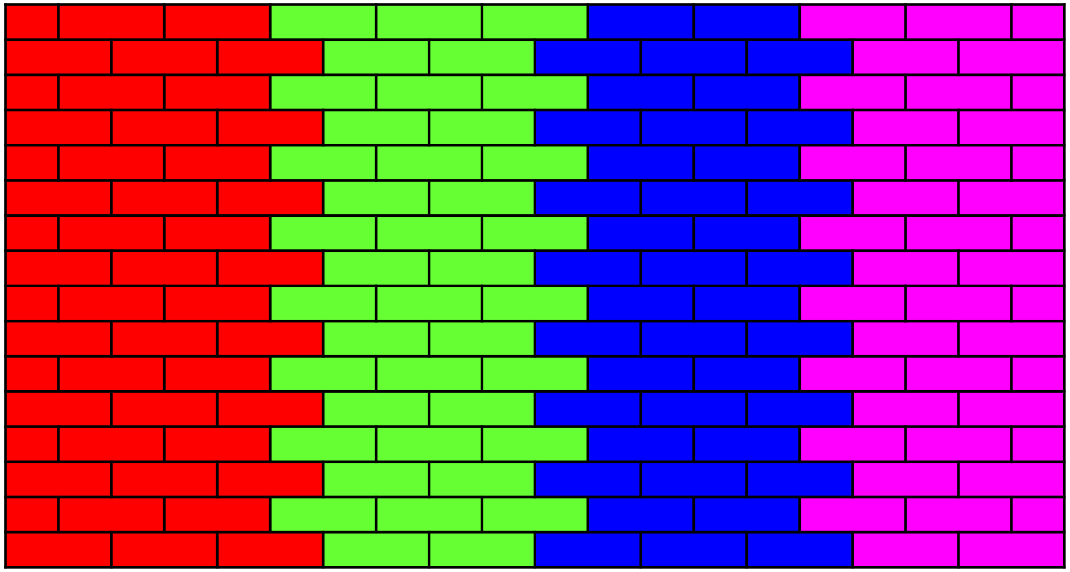
Метод геометрического параллелизма p – число процессов



$$T_1(kn) = \tau_c kn \qquad T_p(kn) = \tau_c \frac{kn}{p} + 4k\tau_s$$

$$S_p(kn) = p\frac{1}{1 + 4\frac{p}{n}\frac{\tau_s}{\tau_c}} \qquad E_p(kn) = \frac{1}{1 + 4\frac{p}{n}\frac{\tau_s}{\tau_c}}$$

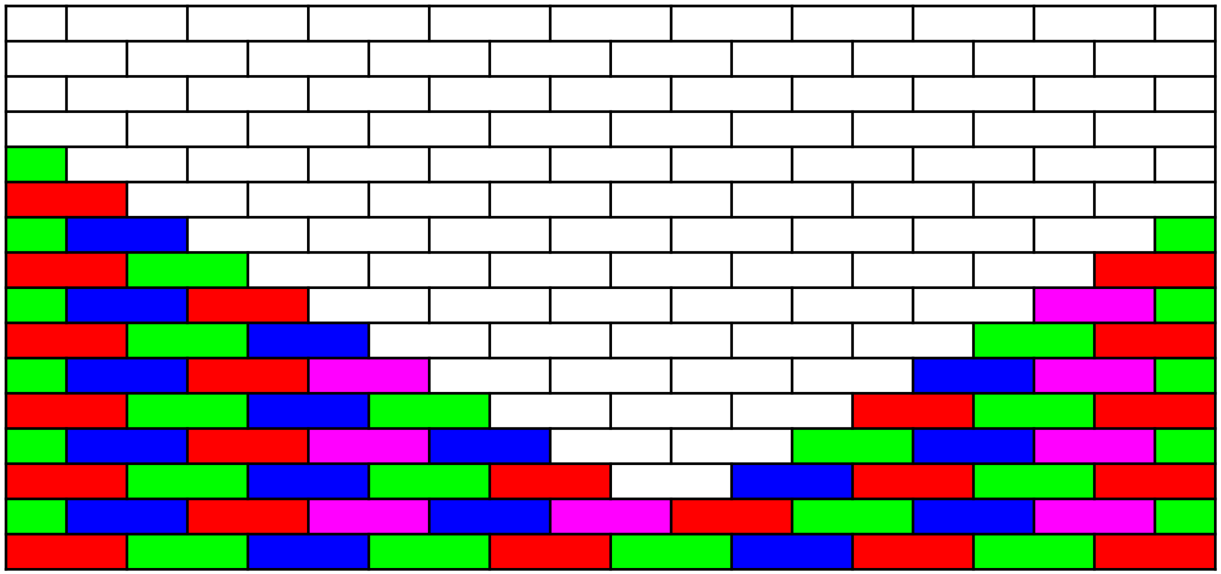
Возможные причины потери эффективности?



$$T_1(kn) = \tau_c kn \qquad T_p(kn) = \tau_c \frac{kn}{p} + 4k\tau_s$$

$$S_p(kn) = p \frac{1}{1 + 4\frac{p}{n}\frac{\tau_s}{\tau_c}} \qquad E_p(kn) = \frac{1}{1 + 4\frac{p}{n}\frac{\tau_s}{\tau_c}}$$

Возможные причины потери эффективности?



n – ширина стены
$к$ – высота стены

Содержание лекции

- Методы построения параллельных алгоритмов и их свойства:
  - Статическая балансировка
    - метод сдваивания
    - геометрический параллелизм
    - конвейерный параллелизм
  - Динамическая балансировка
    - коллективное решение
    - диффузная балансировка загрузки

Метод коллективного решения

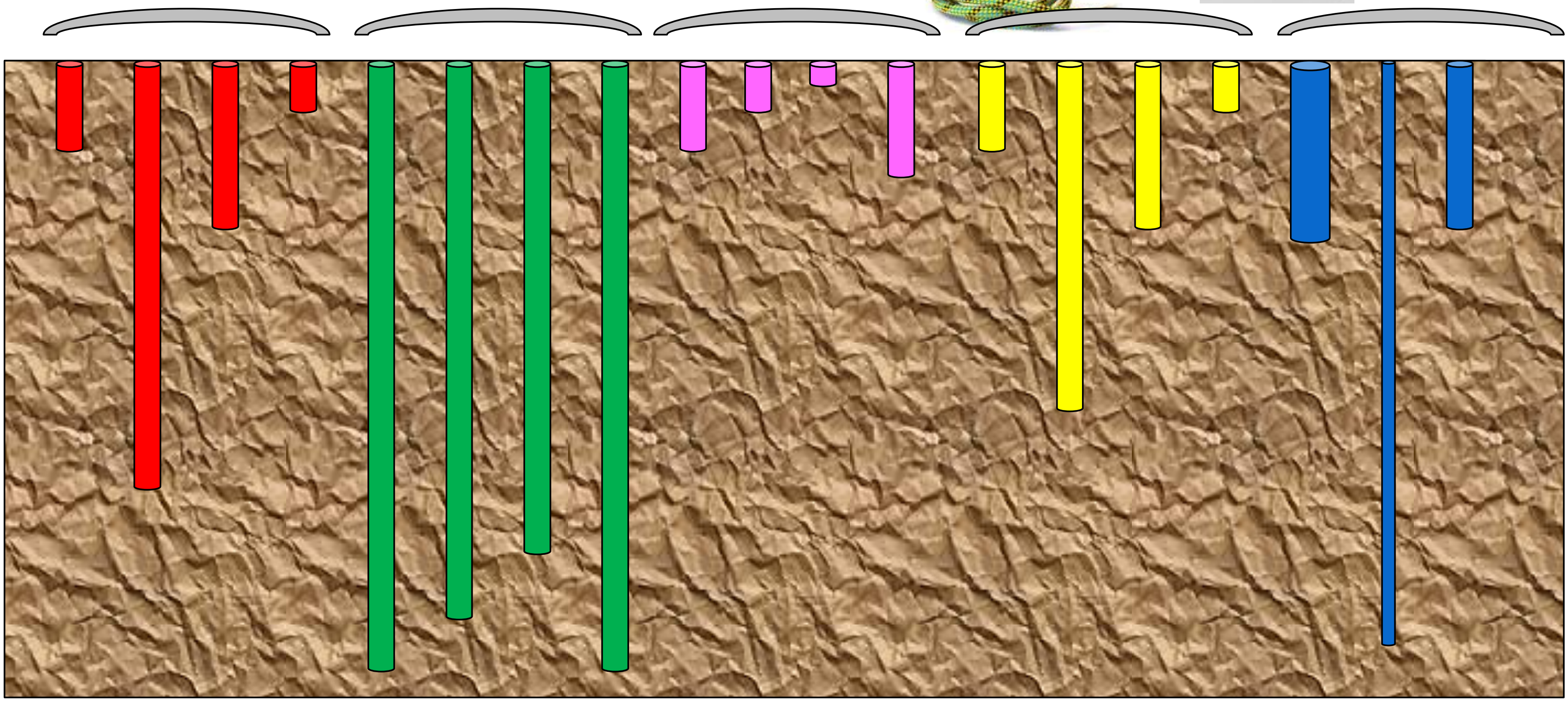
Решение множества независимых друг от друга заданий, в таких прикладных областях как:

- Табулирование функций
- Решение задач методами Монте-Карло
- Численное интегрирование гладких многомерных функций
- …

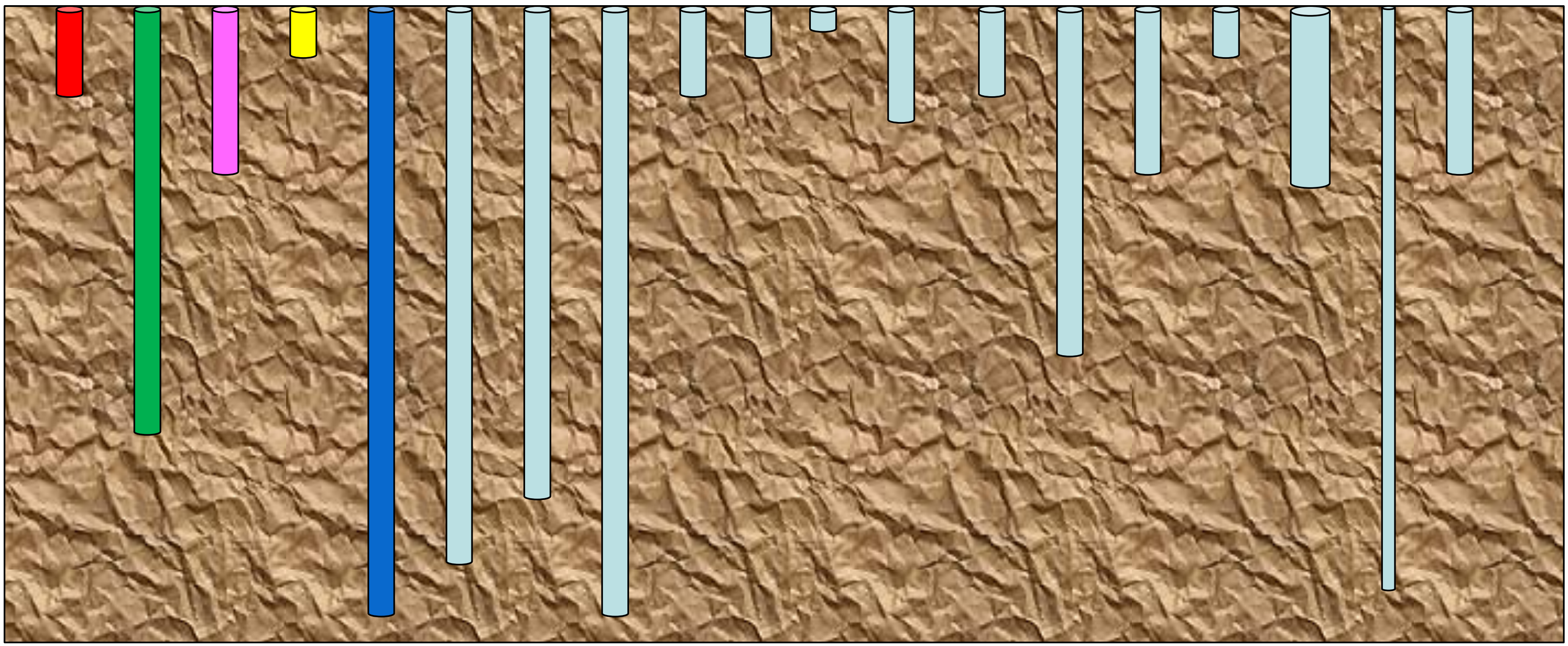
Найти самую глубокую скважину

Найти самую глубокую скважину

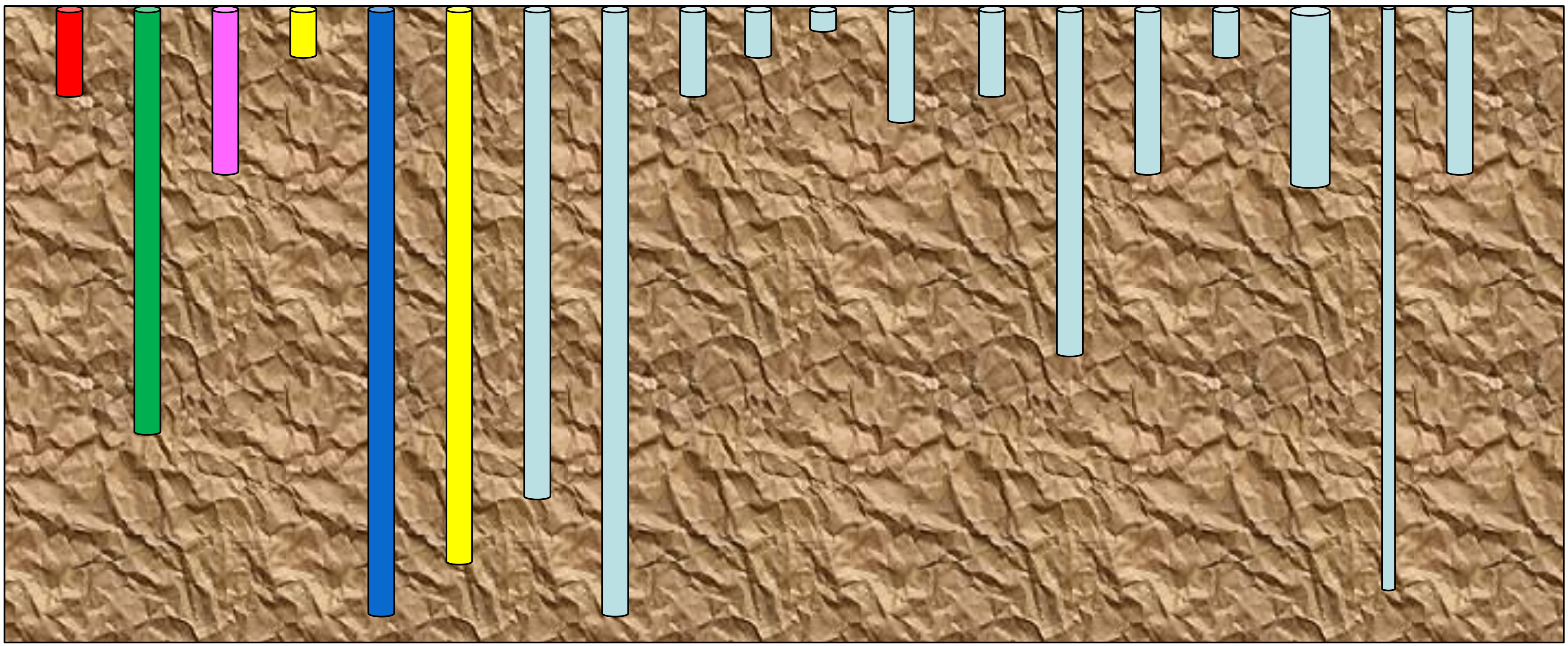
Найти самую глубокую скважину



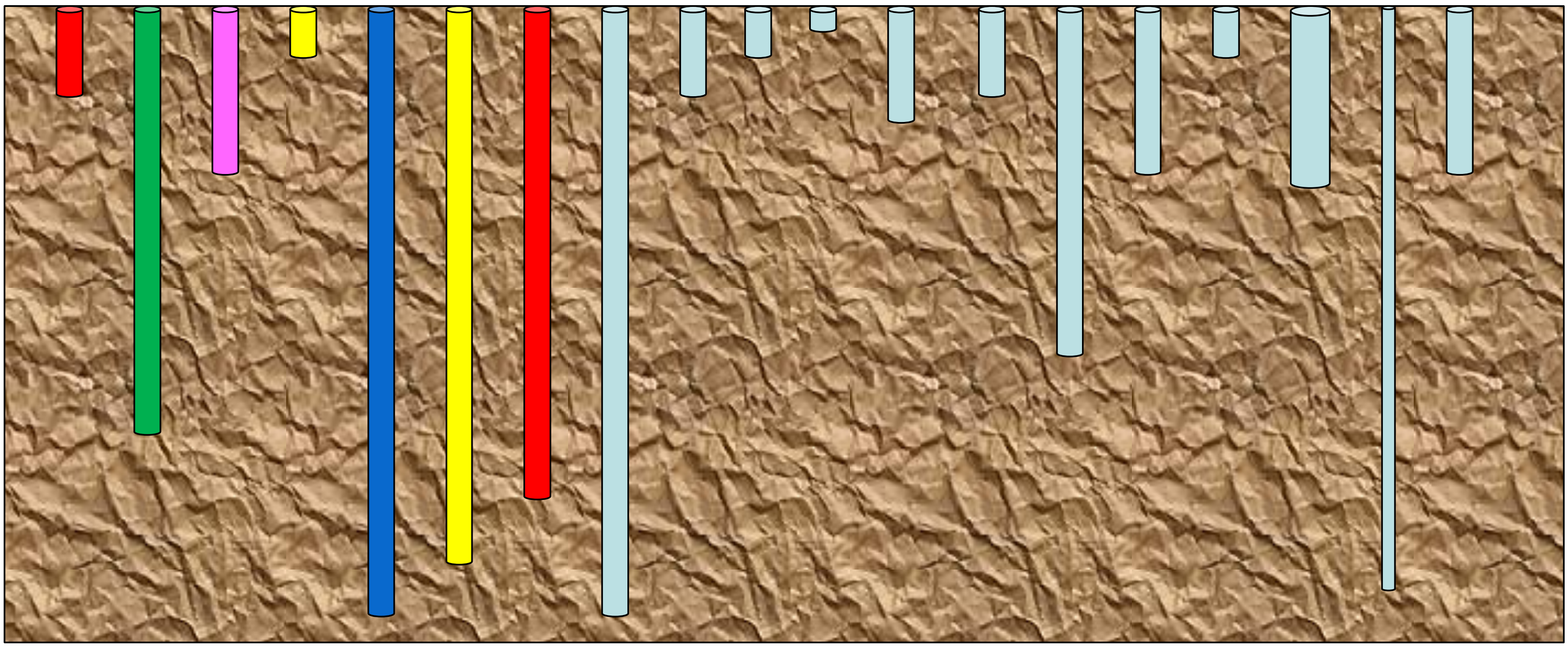
Найти самую глубокую скважину



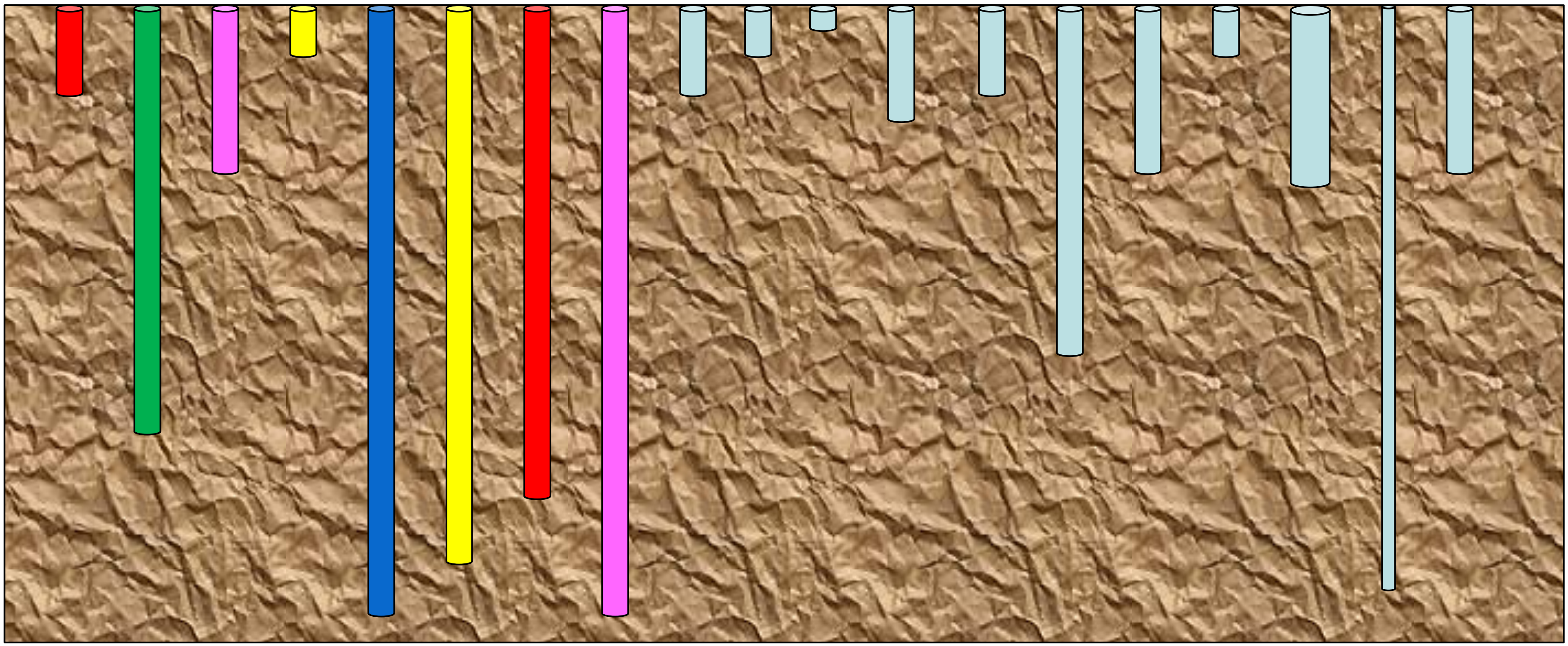
Найти самую глубокую скважину



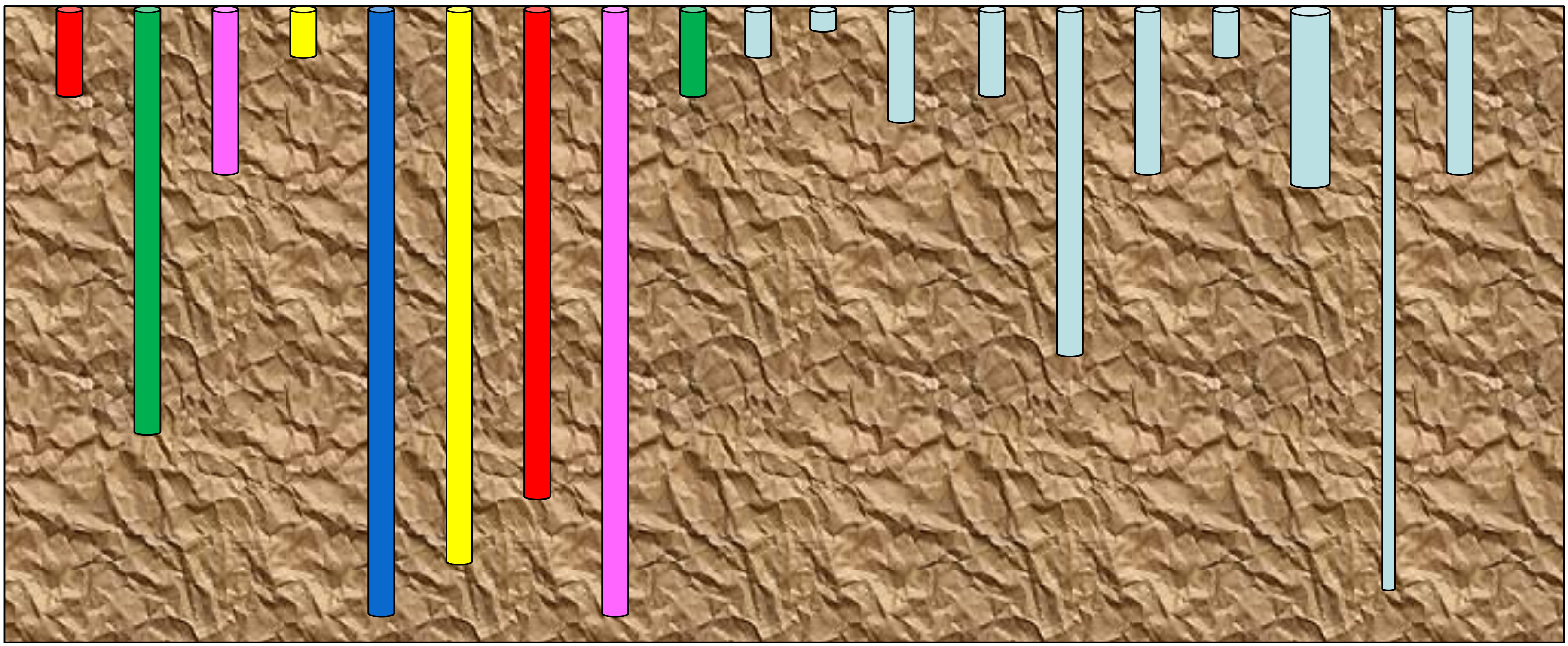
Найти самую глубокую скважину



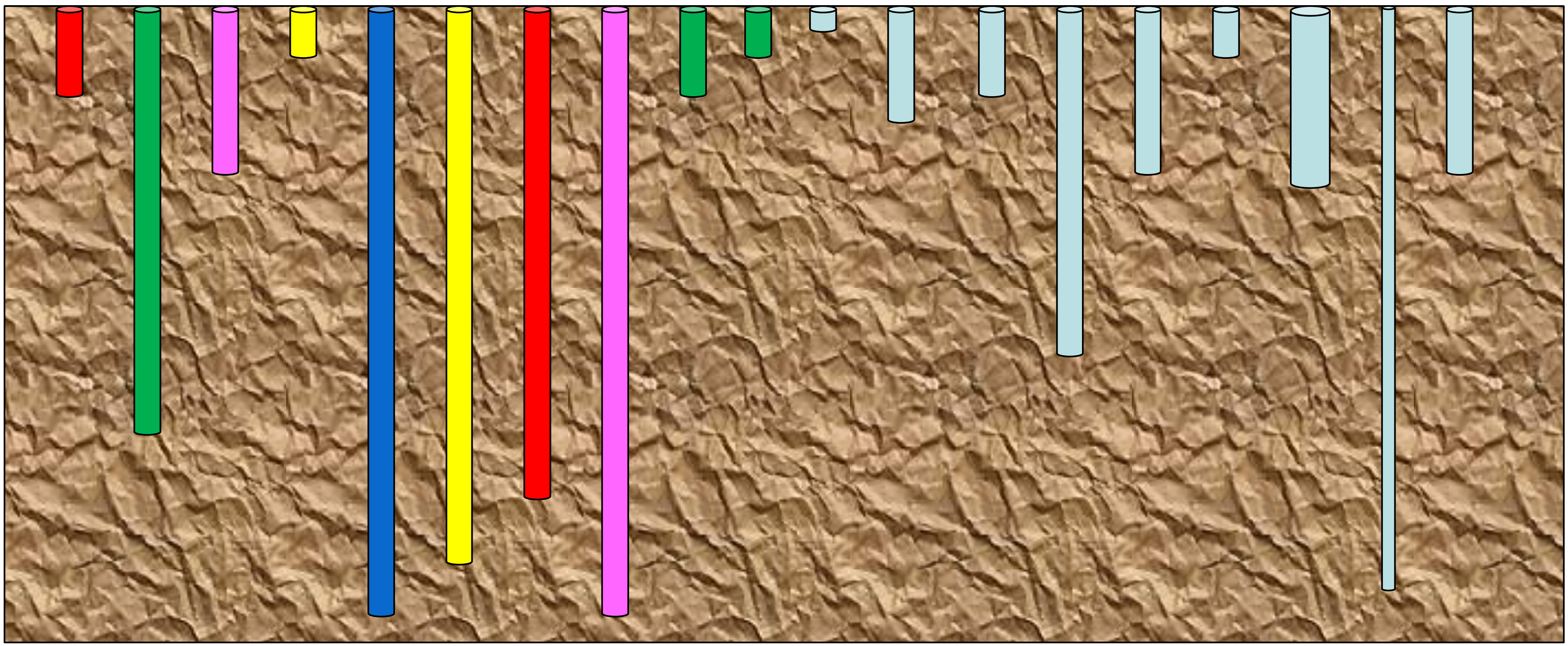
Найти самую глубокую скважину



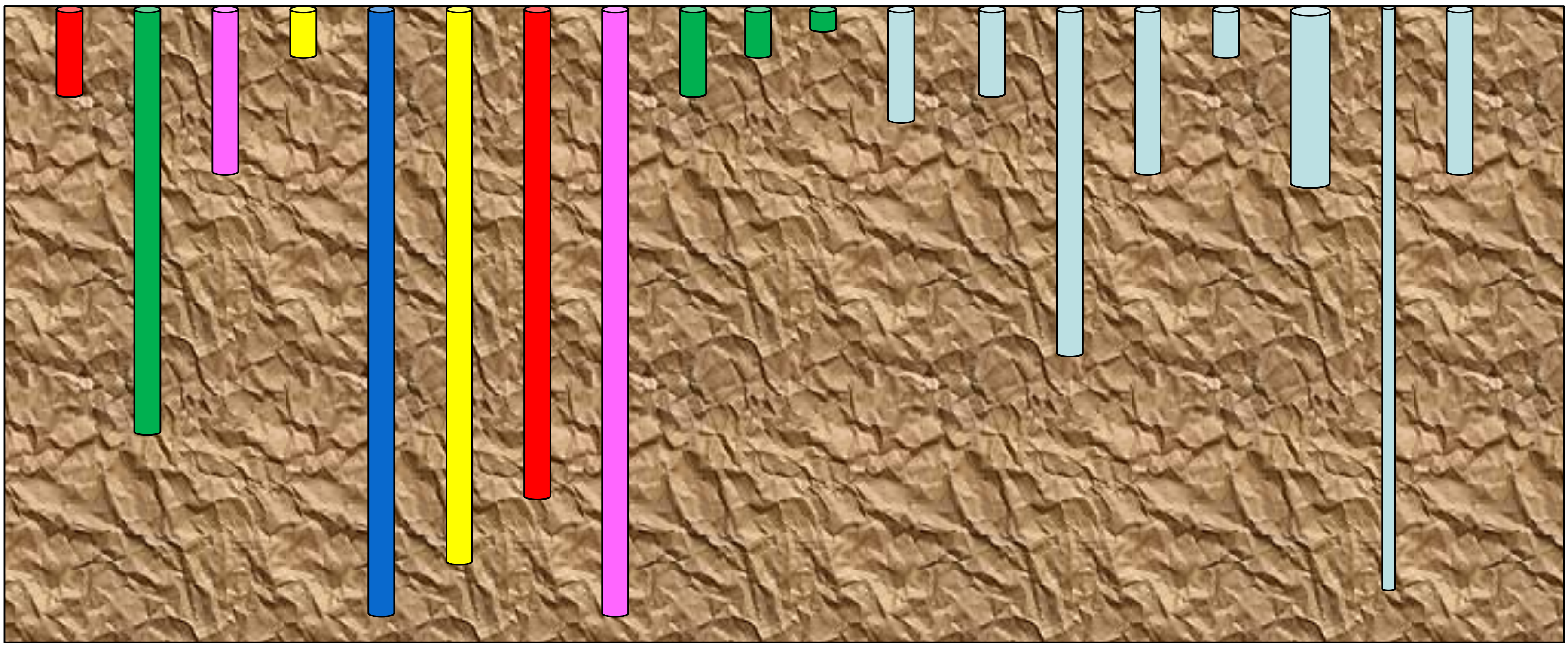
Найти самую глубокую скважину



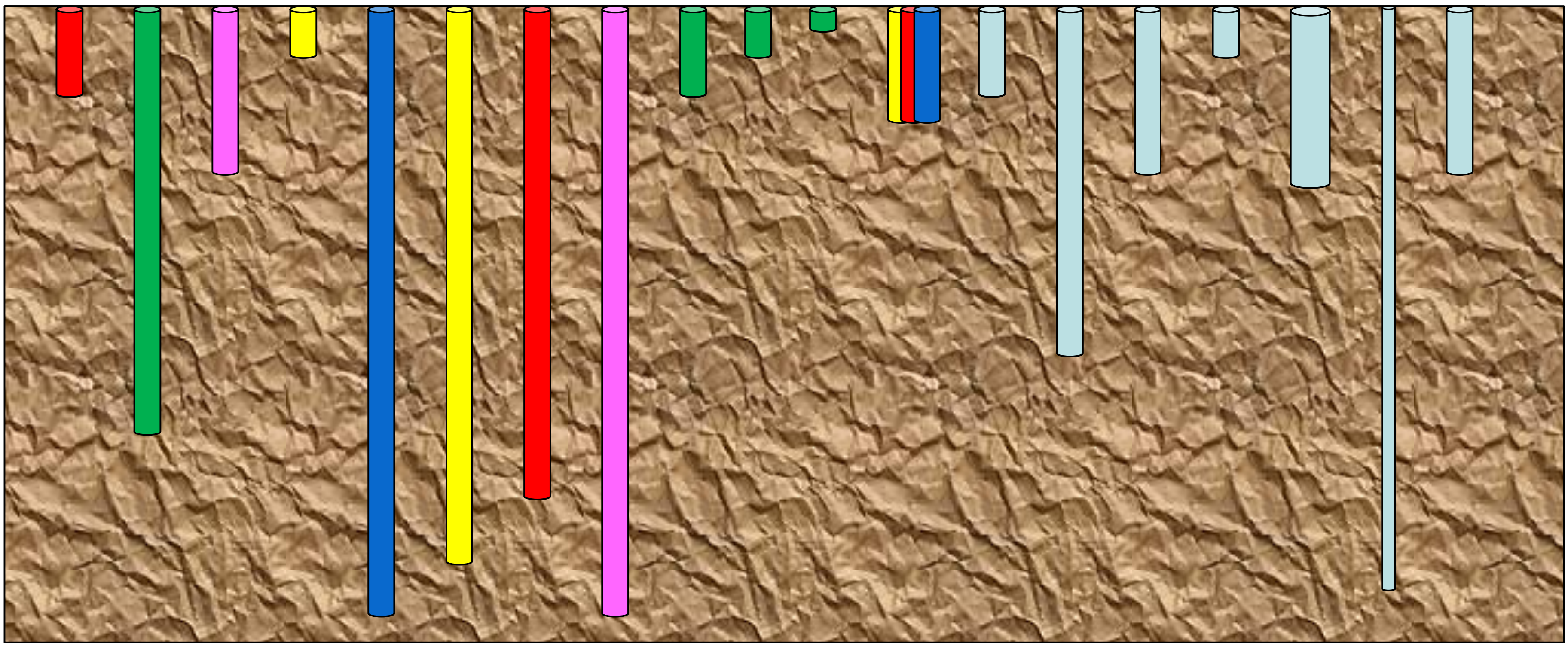
Найти самую глубокую скважину



Найти самую глубокую скважину

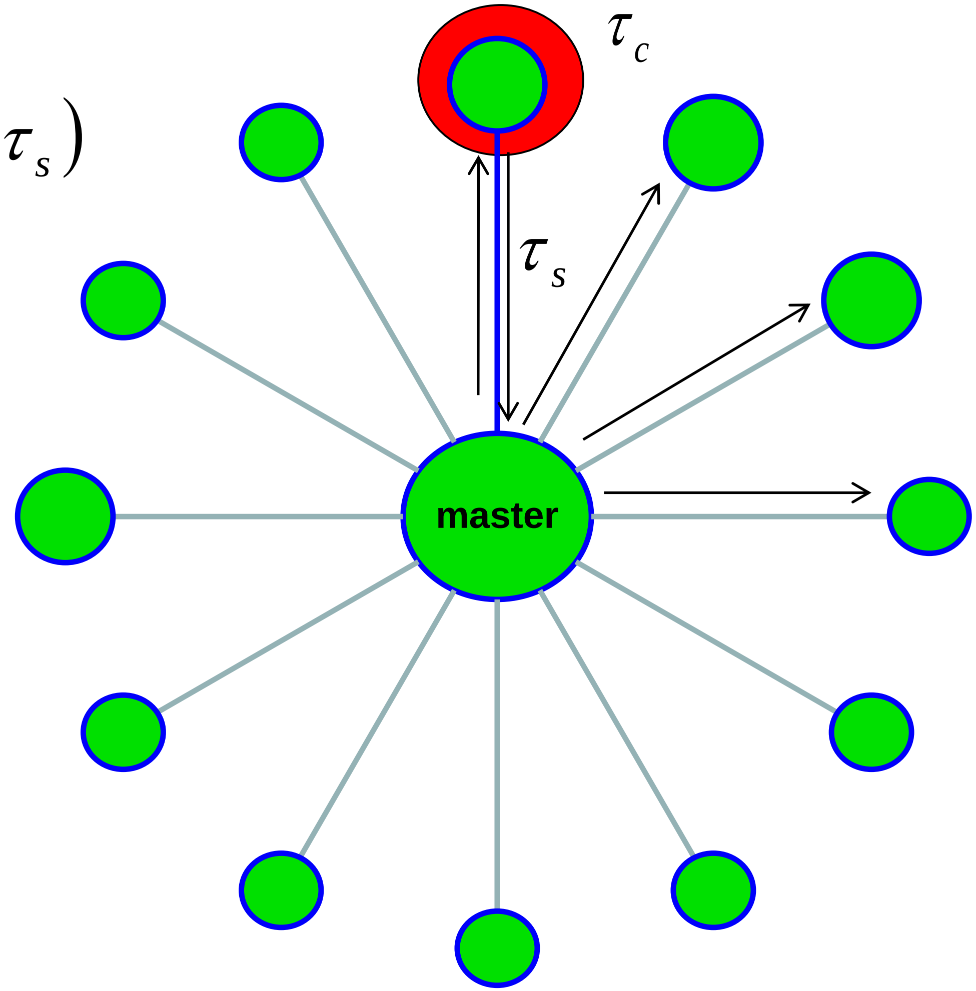


Найти самую глубокую скважину



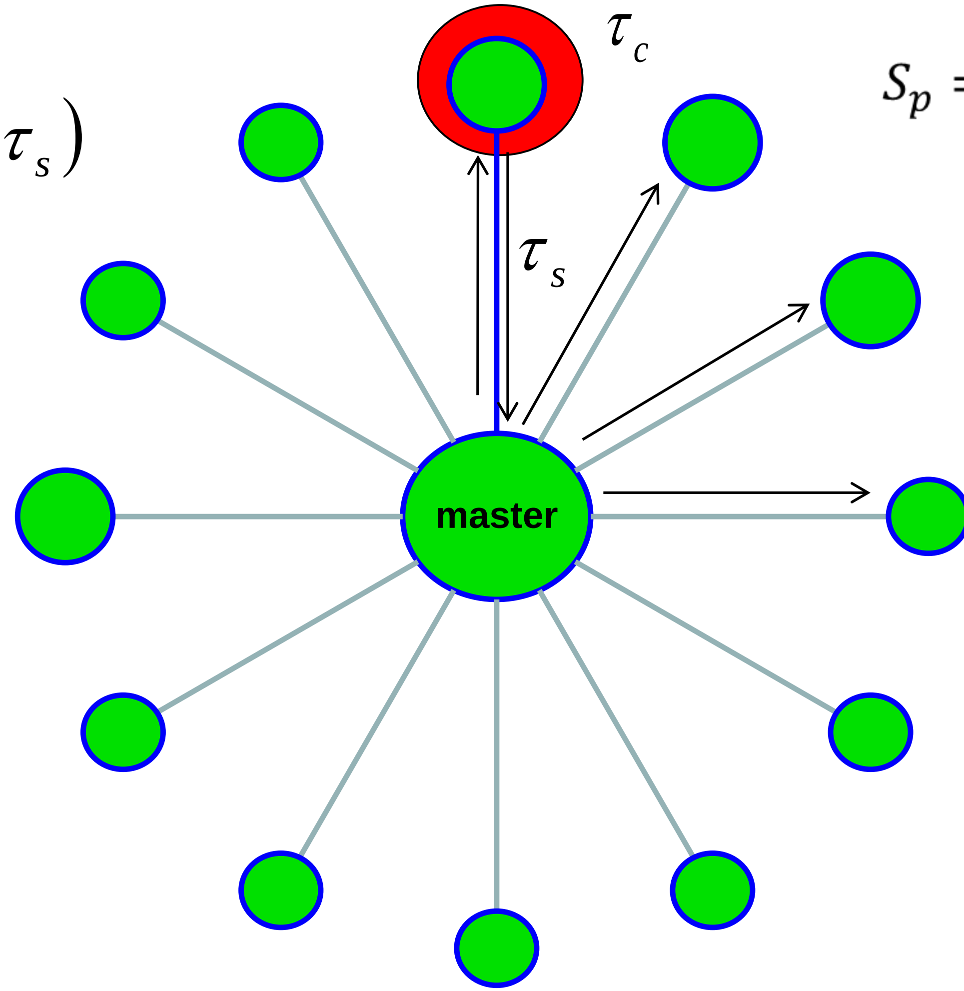
Метод коллективного решения

$$T_p = \frac{N}{p}\left(\tau_c + \tau_s\right)$$



Метод коллективного решения

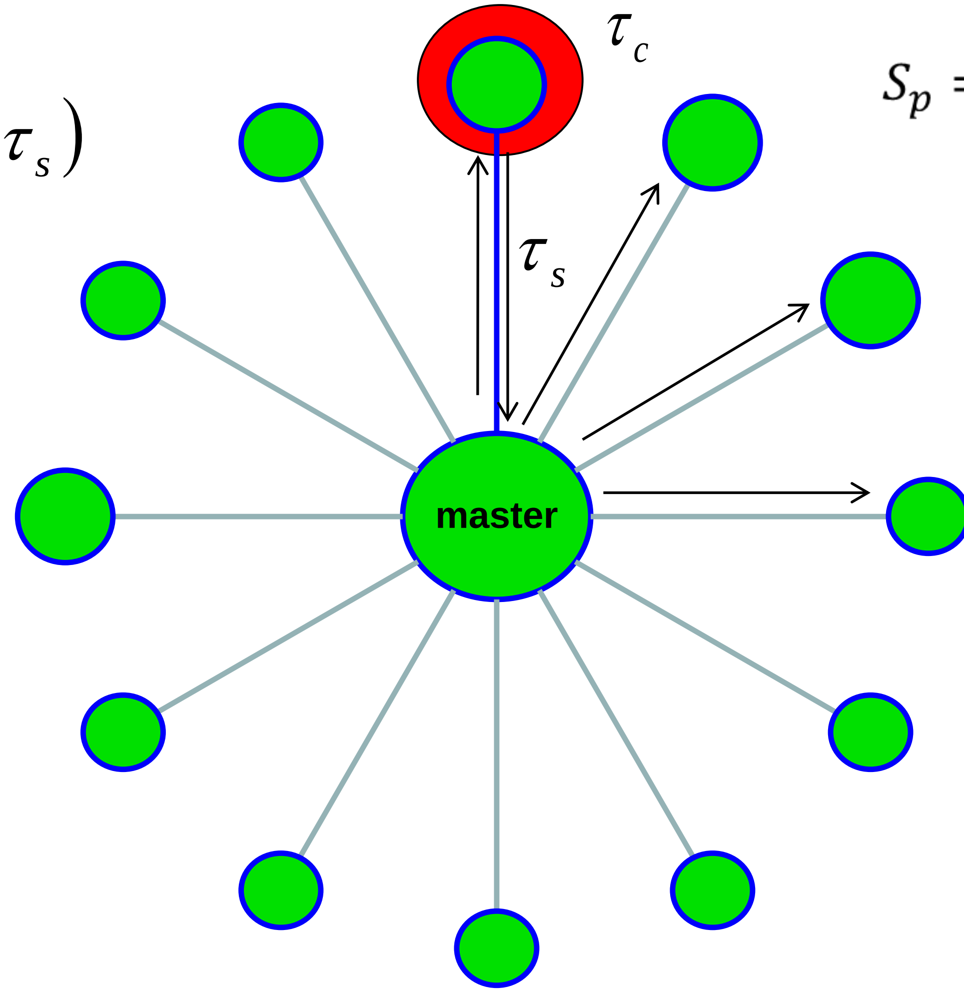
$$T_p = \frac{N}{p}\left(\tau_c + \tau_s\right)$$



$$S_p = \frac{N\tau_c}{\frac{N}{p}(\tau_c + \tau_s)} = p\frac{1}{1 + \frac{\tau_s}{\tau_c}}$$

Метод коллективного решения

$$T_p = \frac{N}{p}\left(\tau_c + \tau_s\right)$$

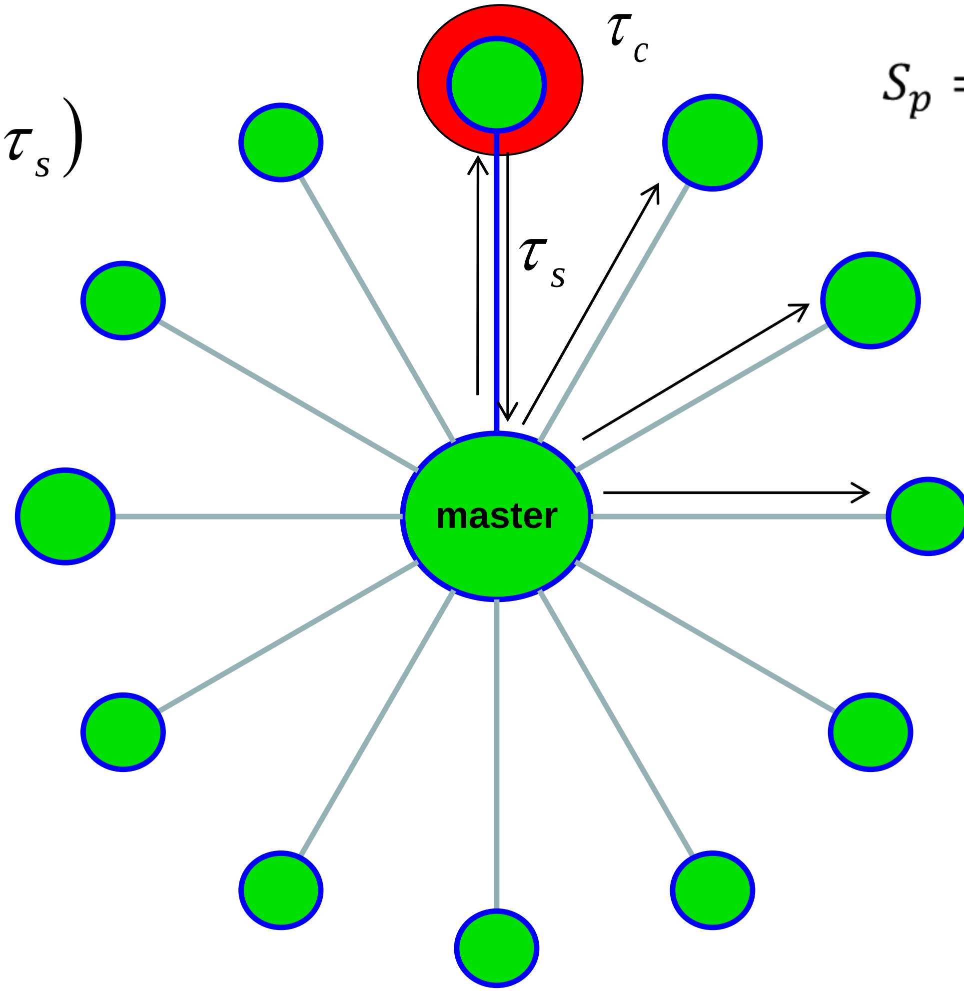


$$S_p = \frac{N\tau_c}{\frac{N}{p}(\tau_c + \tau_s)} = p\frac{1}{1 + \frac{\tau_s}{\tau_c}}$$

$$S_p \approx p$$

Метод коллективного решения

$$T_p = \frac{N}{p}\left(\tau_c + \tau_s\right)$$

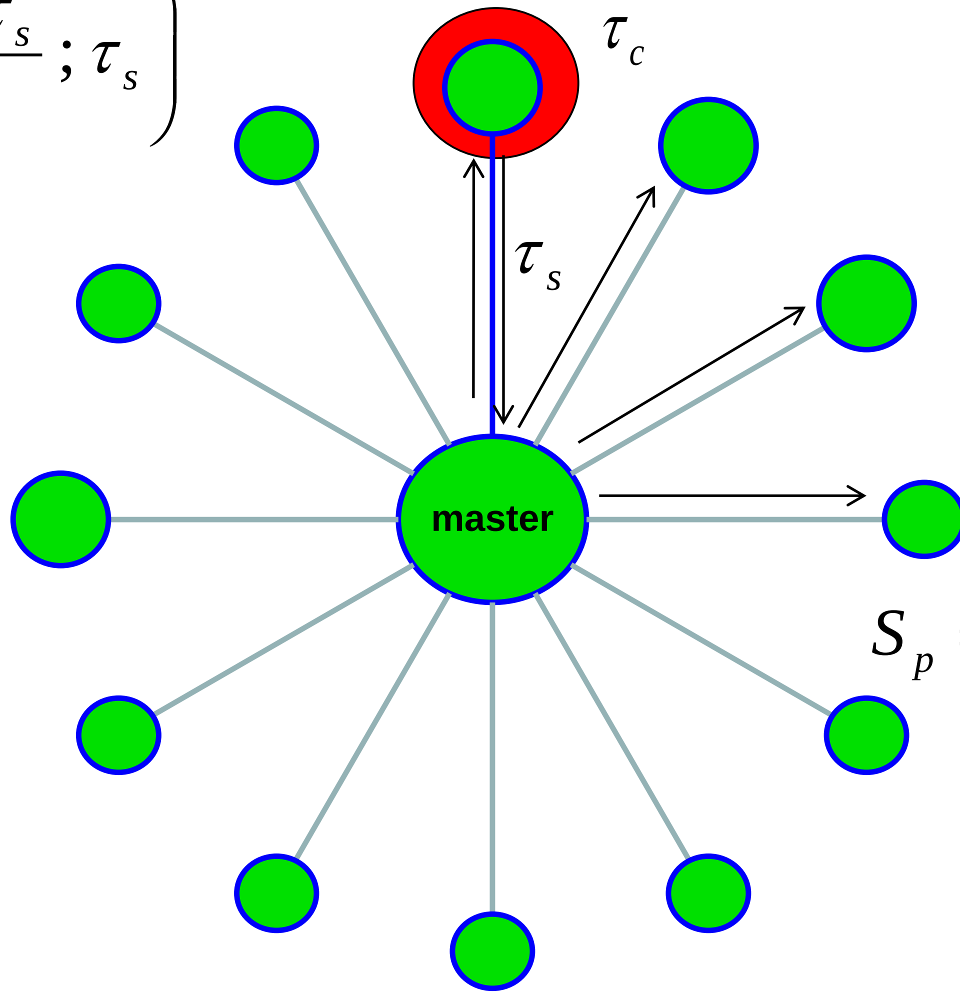


$$S_p = \frac{N\tau_c}{\frac{N}{p}(\tau_c + \tau_s)} = p\frac{1}{1 + \frac{\tau_s}{\tau_c}}$$

$$T_p \geq N\tau_s$$

Метод коллективного решения

$$T_p = N \max\left(\frac{\tau_c + \tau_s}{p}; \tau_s \right)$$



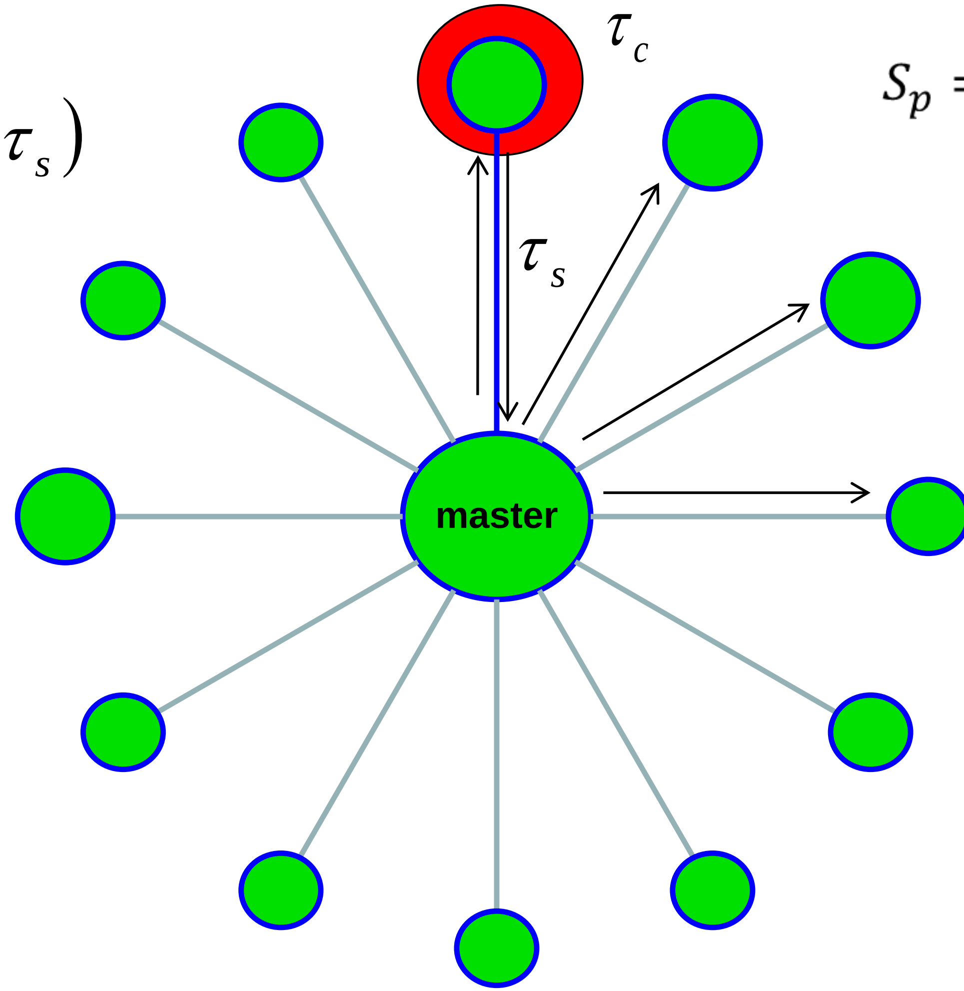
$$T_1 = N\tau_c$$

$$T_p \geq N\tau_s$$

$$S_p = \min\left(p\frac{1}{1+\dfrac{\tau_s}{\tau_c}}; \frac{\tau_c}{\tau_s} \right)$$

Метод коллективного решения

$$T_p = \frac{N}{p}\left(\tau_c + \tau_s\right)$$



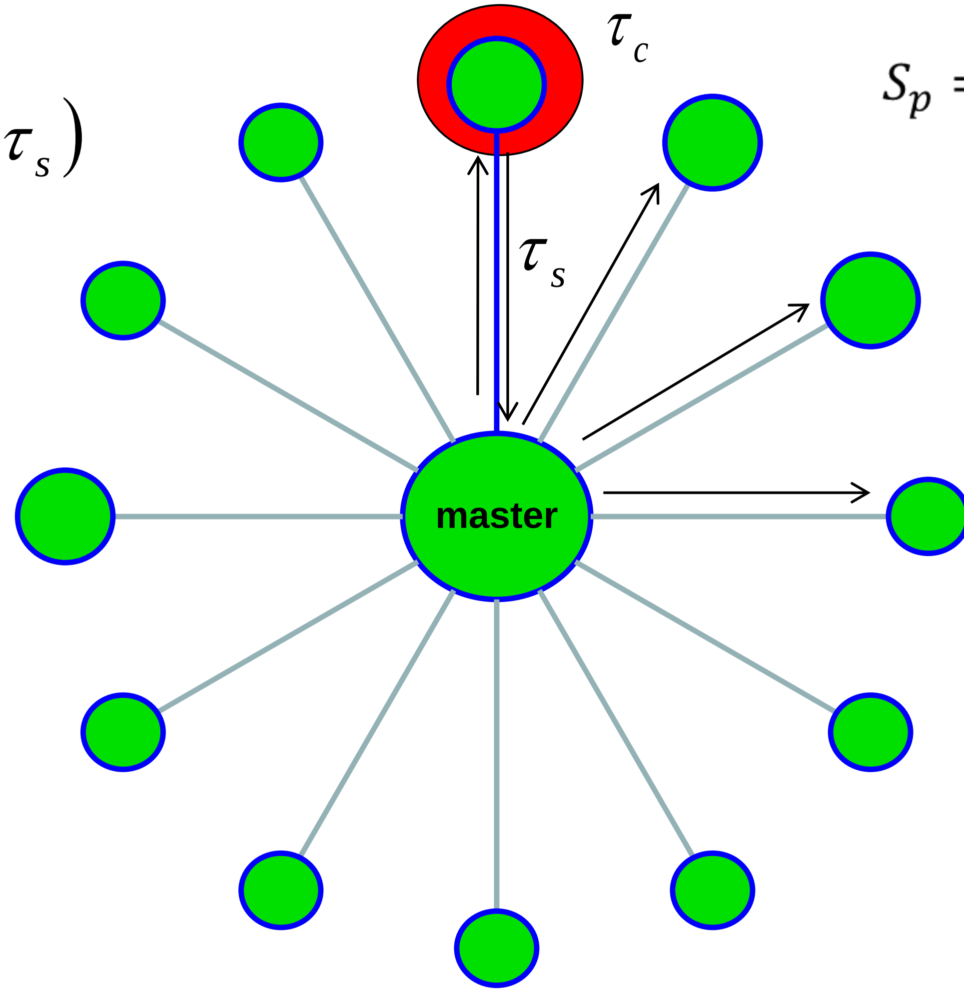
$$S_p = \frac{N\tau_c}{\frac{N}{p}(\tau_c + \tau_s)} = p\frac{1}{1 + \frac{\tau_s}{\tau_c}}$$

$$T_p \geq N\tau_s$$

$$S_p \leq \frac{N\tau_c}{N\tau_s} = \frac{\tau_c}{\tau_s}$$

Метод коллективного решения

$$T_p = \frac{N}{p}(\tau_c + \tau_s)$$



$$S_p = \frac{N\tau_c}{\frac{N}{p}(\tau_c + \tau_s)} = p\frac{1}{1+\frac{\tau_s}{\tau_c}}$$

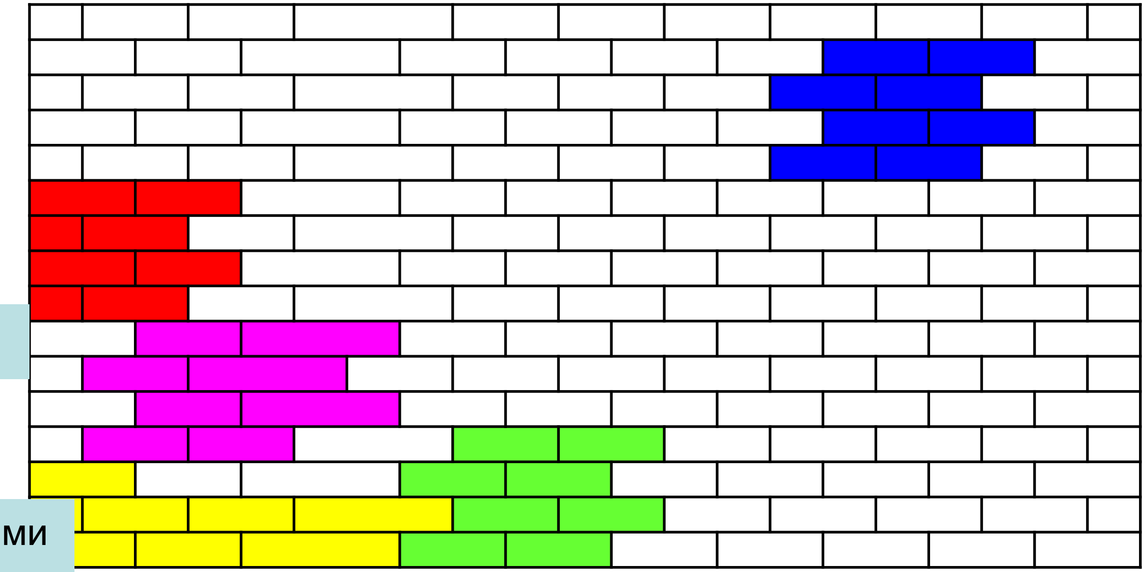
$$T_p \geq N\tau_s$$

$$S_p \leq \frac{N\tau_c}{N\tau_s} = \frac{\tau_c}{\tau_s}$$

$$S_p \approx p \leq p_{max}$$

$$p_{\max} = \frac{\tau_c}{\tau_s}$$

Пример модельной задачи: Укладка паркета



$$T_1(kn) = \tau_c kn$$

$$T_p(kn) = \frac{kn}{rp}(r\tau_c + \tau_s)$$

Число порций

Обработка порции

Обмен данными

$$S_{p=\frac{r\tau_c}{\tau_s}}(kn) = \left(\frac{r\tau_c}{\tau_s}\right)^2 \frac{1}{1+\frac{r\tau_c}{\tau_s}} = p_{\max}\frac{1}{1+\frac{1}{p_{\max}}}$$

$$p_{\max} = \frac{r\tau_c}{\tau_s}$$

$$E_{p=\frac{r\tau_c}{\tau_s}}(kn) = \frac{1}{1+\frac{\tau_s}{r\tau_c}}$$

r – размер порции

Как правильно?

r – размер порции

$$p_{\max} = \frac{r\tau_c}{\tau_s}$$

или

$$p_{\max} = \frac{r\tau_c}{r\tau_s} = \frac{\tau_c}{\tau_s}$$

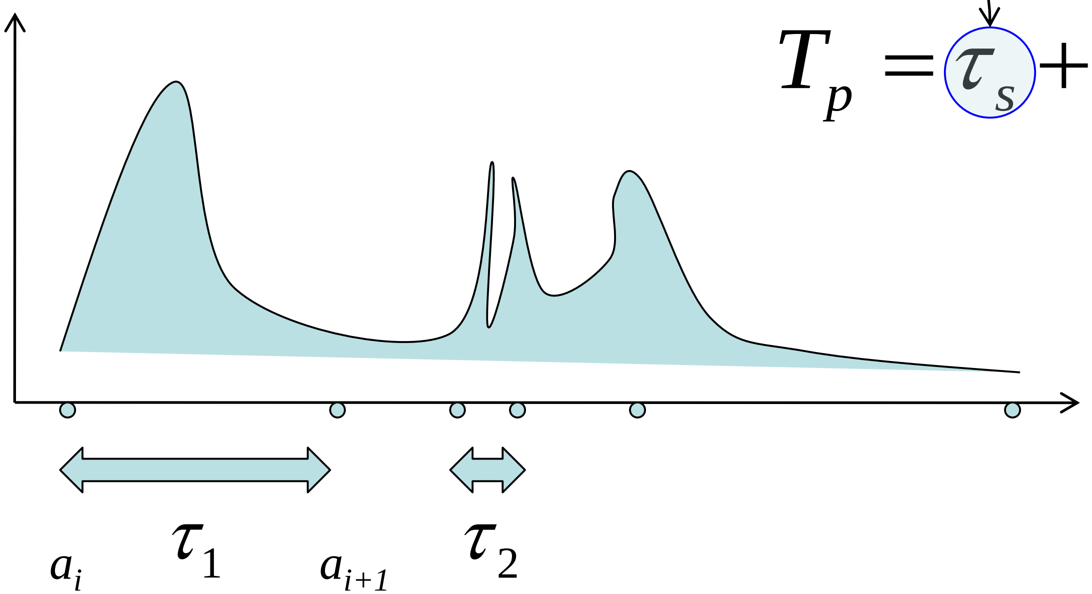
Зависит ли время передачи данных от размера порции (задания)?

Вычисление определенного интеграла

$\boxed{\text{Send}(a_i);\ \text{Send}(a_{i+1});\ \text{Recv}(s);}$

$$I = \int_A^B f(x)dx = \sum_i \int_{a_i}^{a_{i+1}} f(x)dx$$

$$T_p = \tau_s + \max_i \tau_i$$

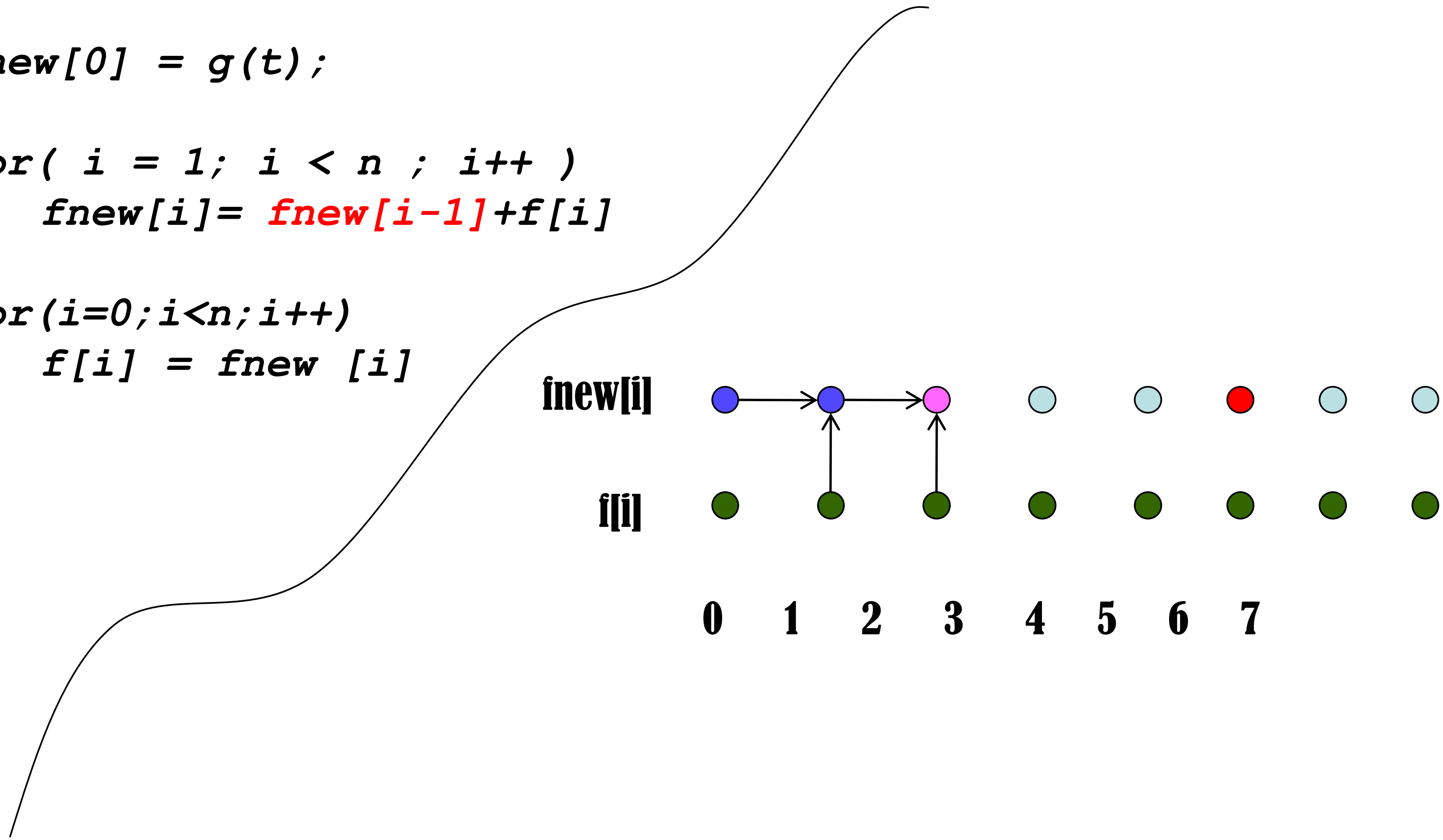


Метод конвейерного параллелизма

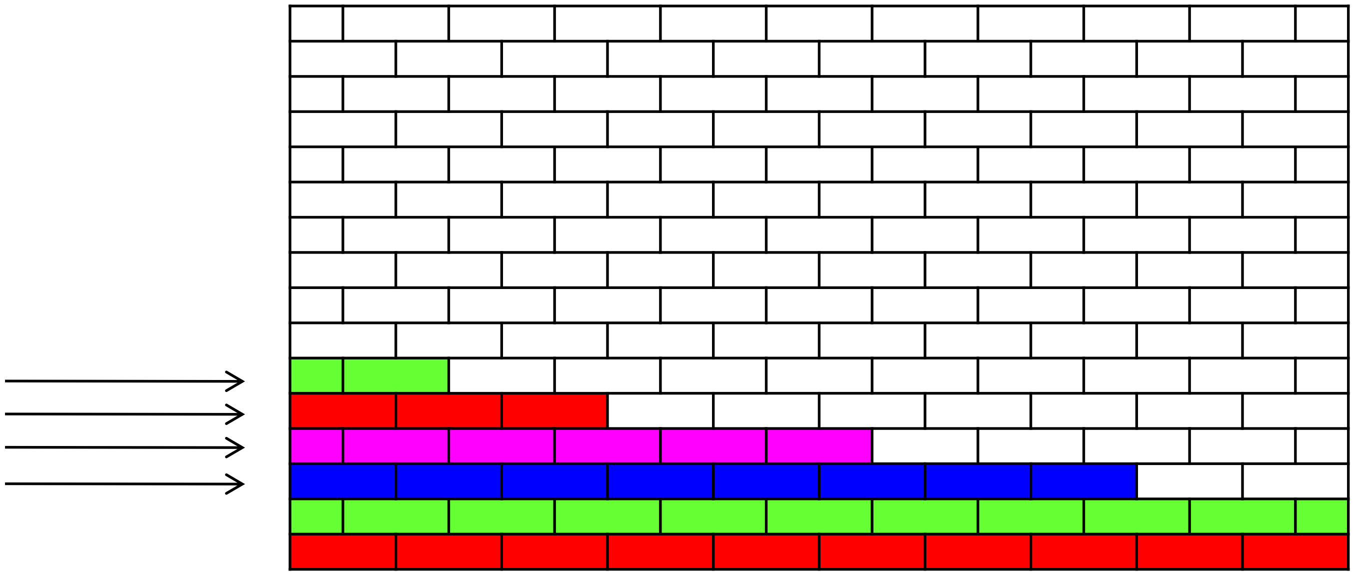
```
for( t = 0; t < tmax; t += dt )
{
    fnew[0] = g(t);

    for( i = 1; i < n ; i++ )
        fnew[i]= fnew[i-1]+f[i]

    for(i=0;i<n;i++)
        f[i] = fnew [i]
}
```

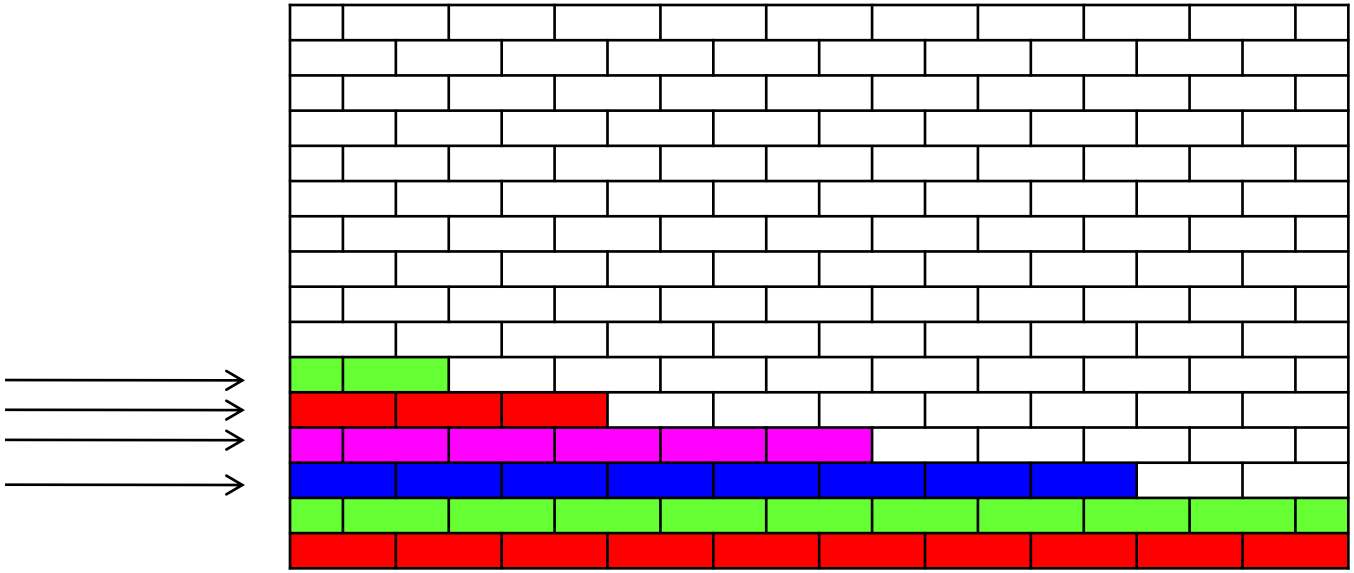


Метод конвейерного параллелизма



$$T_1(kn) = \tau_c kn \quad T_p(kn) = \tau_c \frac{kn}{p} + ?$$

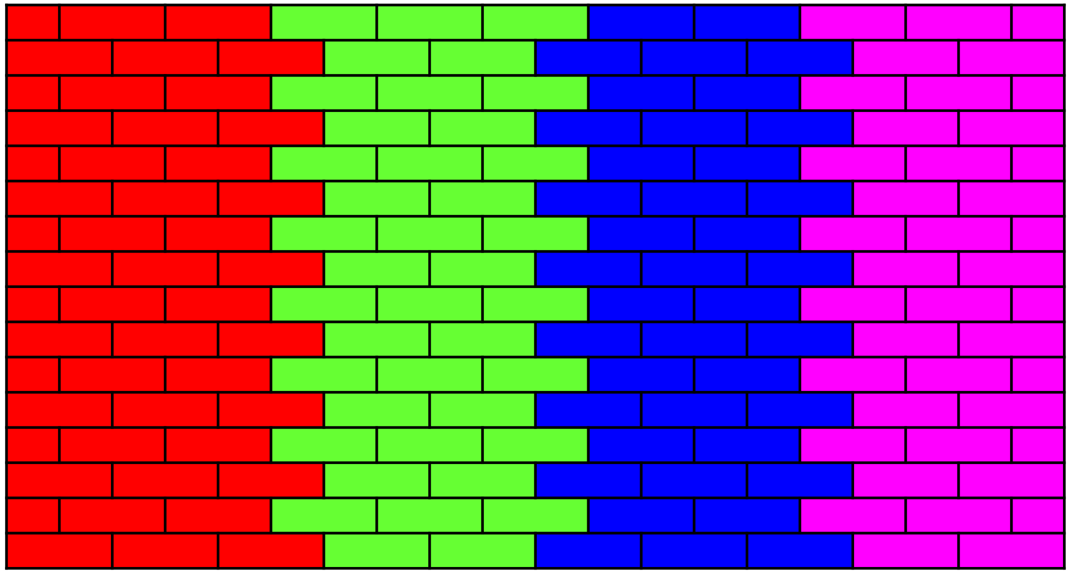
Метод конвейерного параллелизма



$$T_1(kn) = \tau_c kn \qquad T_p(kn) = \tau_c \frac{kn}{p} + k\frac{n}{p}\tau_s$$

$$S_p(kn) = p\frac{1}{1+\dfrac{\tau_s}{\tau_c}} \qquad E_p(kn) = \frac{1}{1+\dfrac{\tau_s}{\tau_c}}$$

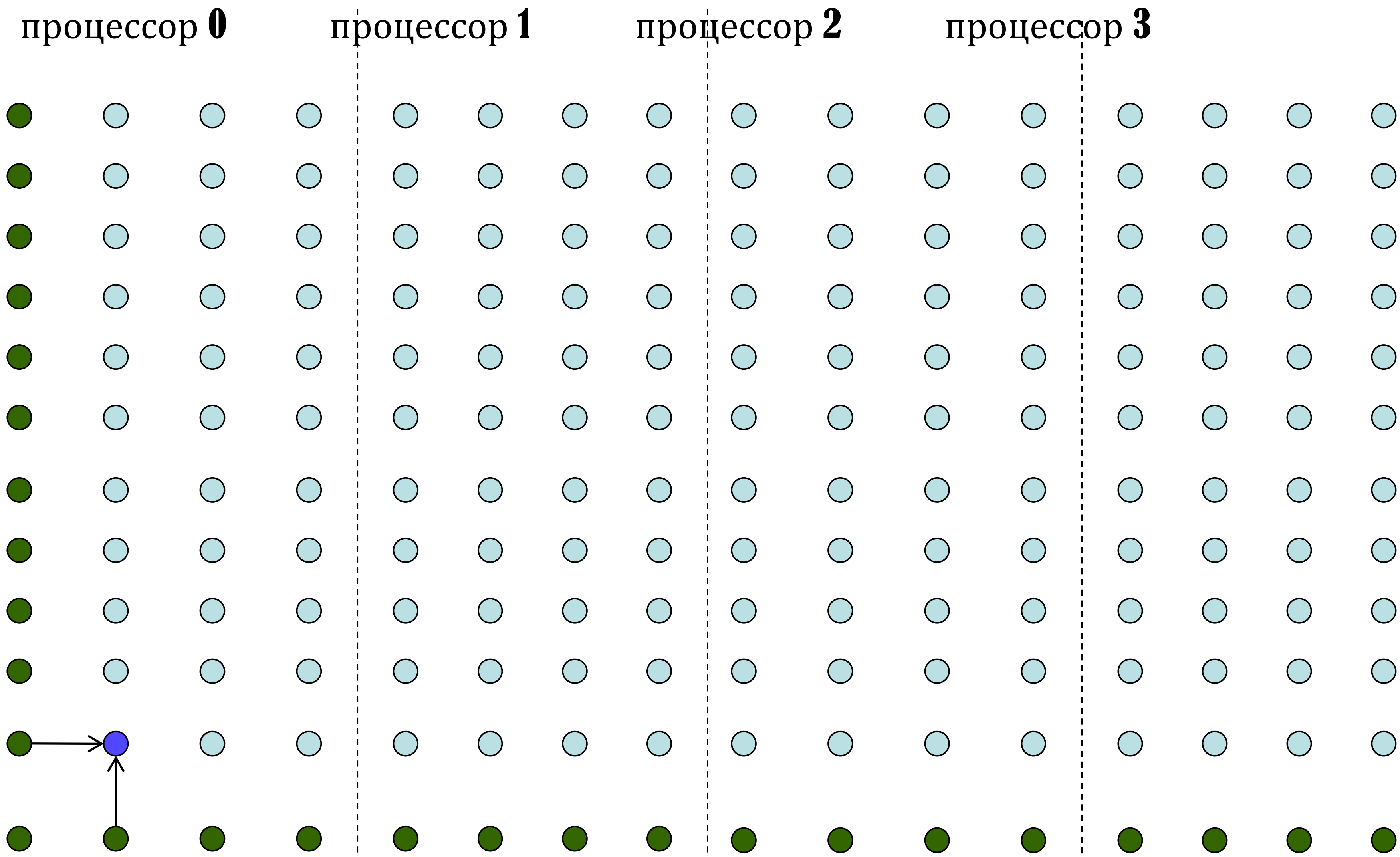
Метод геометрического параллелизма



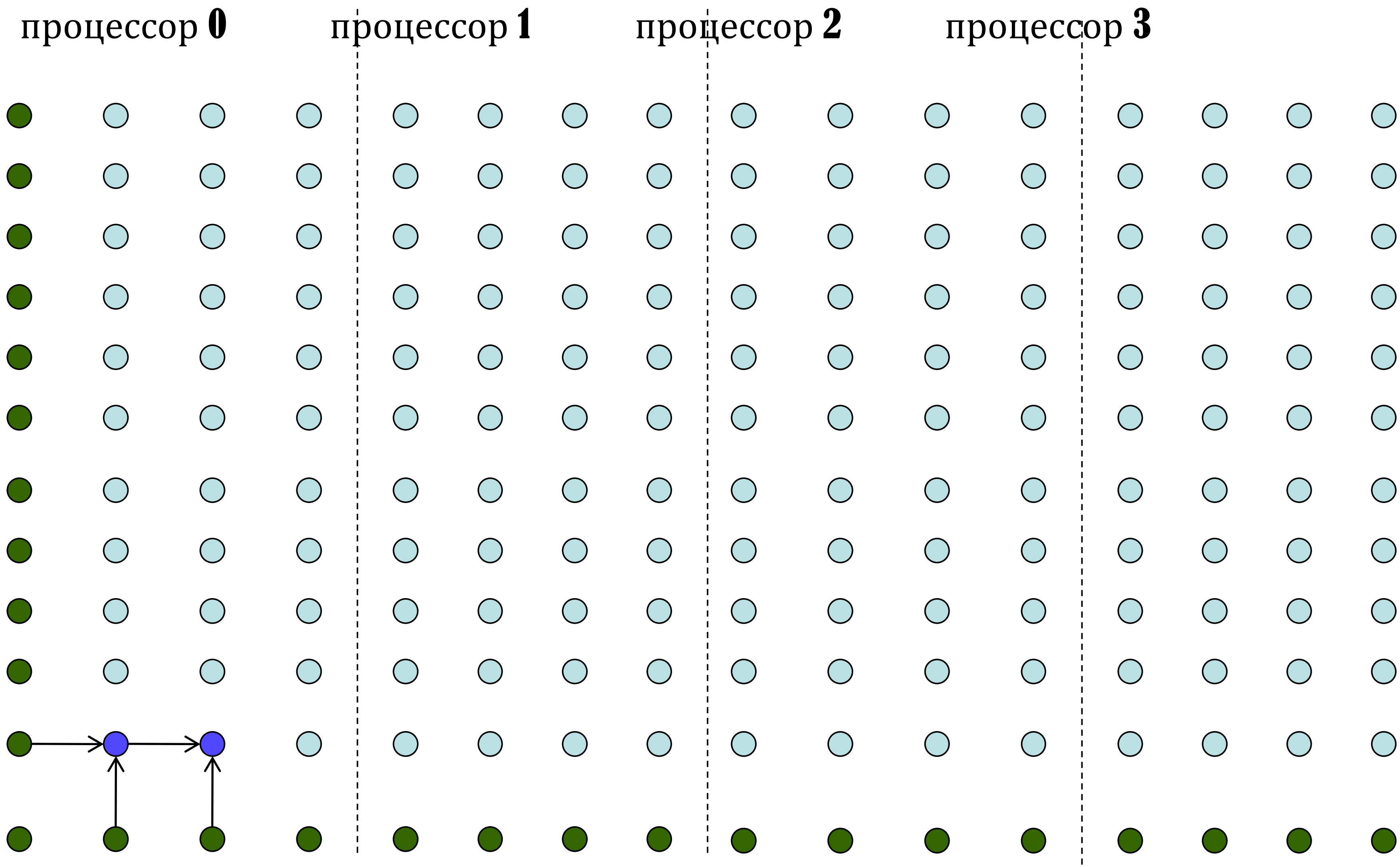
$$T_1(kn) = \tau_c kn \qquad T_p(kn) = \tau_c \frac{kn}{p} + 4k\tau_s$$

$$S_p(kn) = p \frac{1}{1 + 4\frac{p}{n}\frac{\tau_s}{\tau_c}} \qquad E_p(kn) = \frac{1}{1 + 4\frac{p}{n}\frac{\tau_s}{\tau_c}}$$

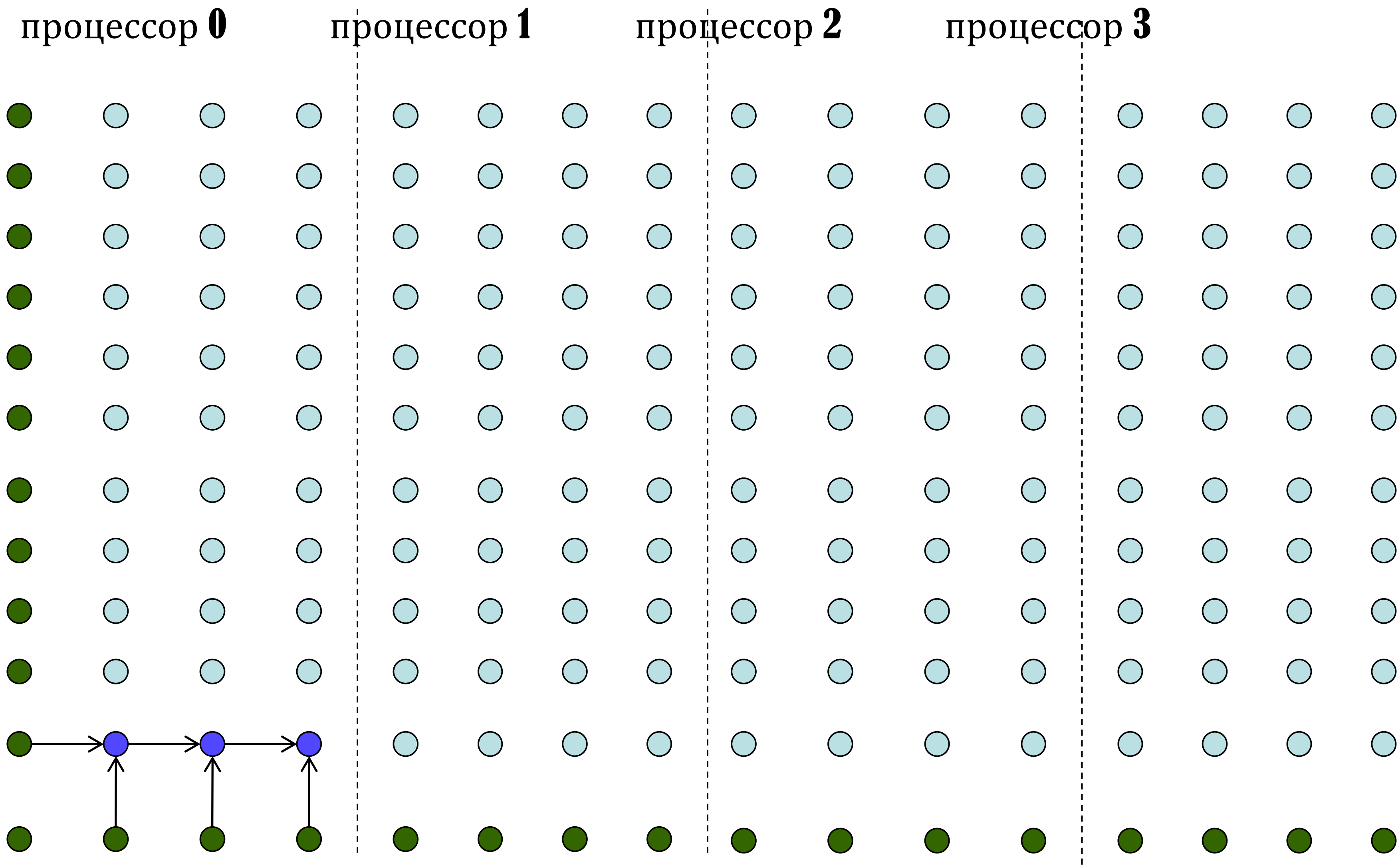
Метод конвейерного параллелизма



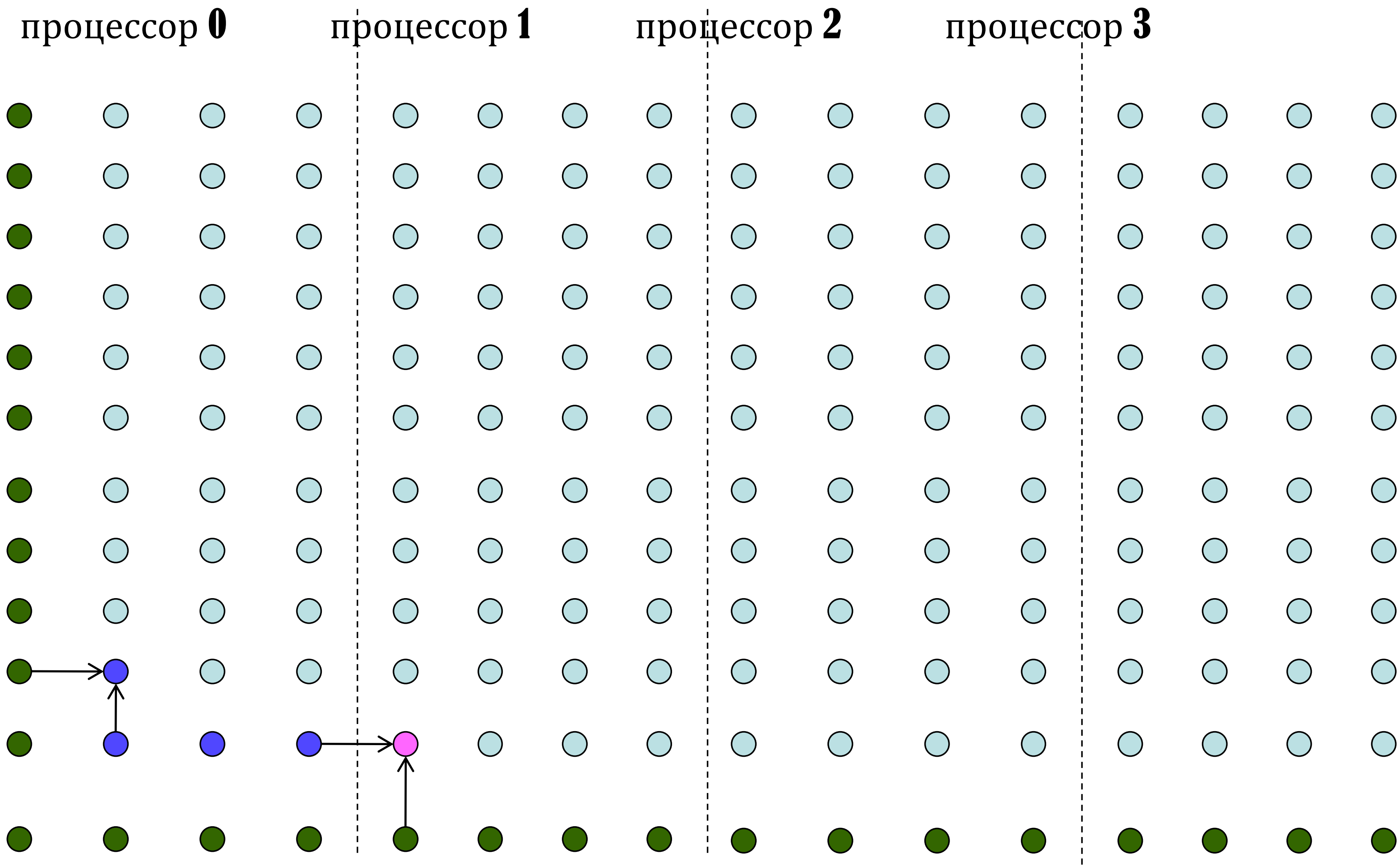
Метод конвейерного параллелизма



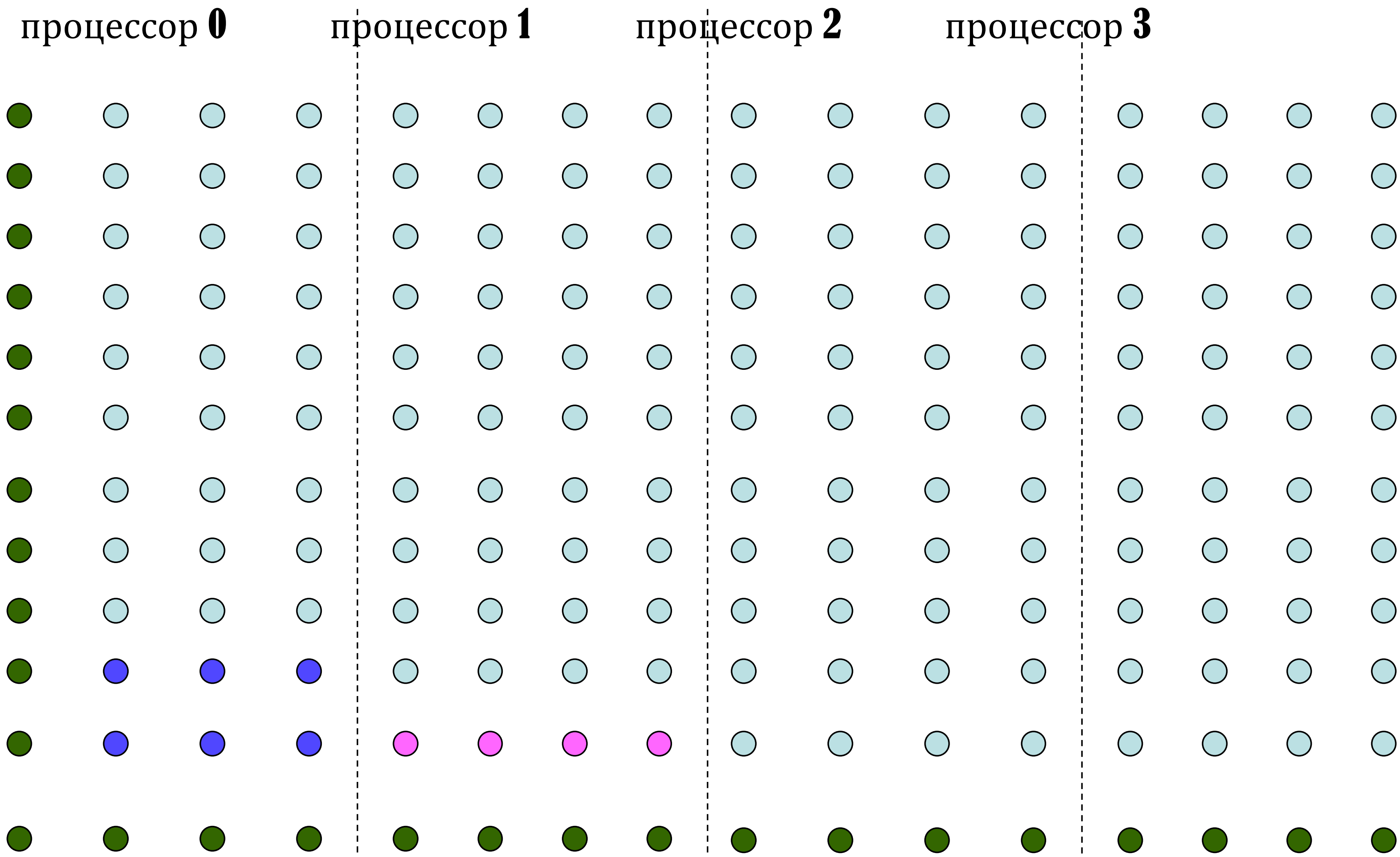
Метод конвейерного параллелизма



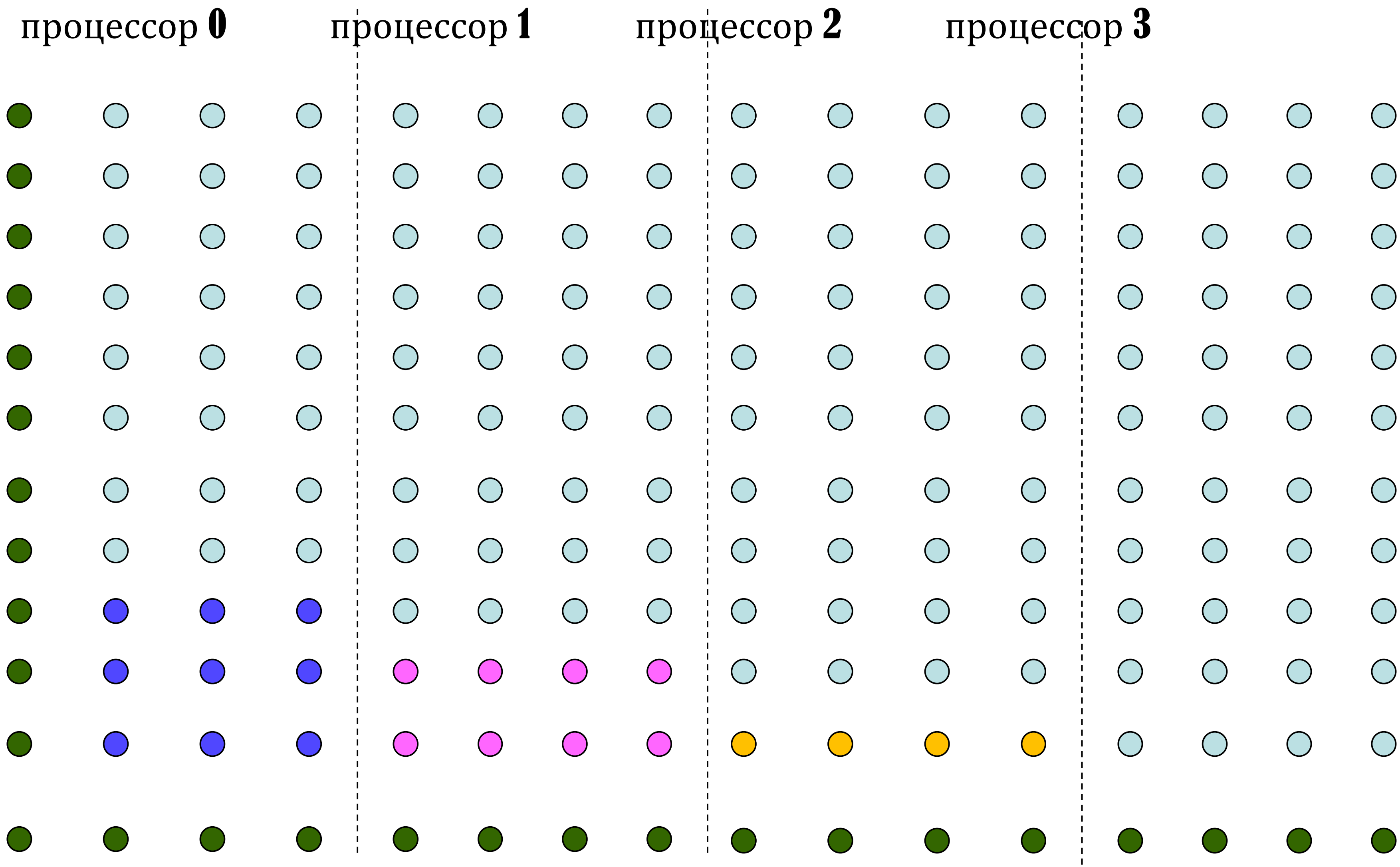
Метод конвейерного параллелизма



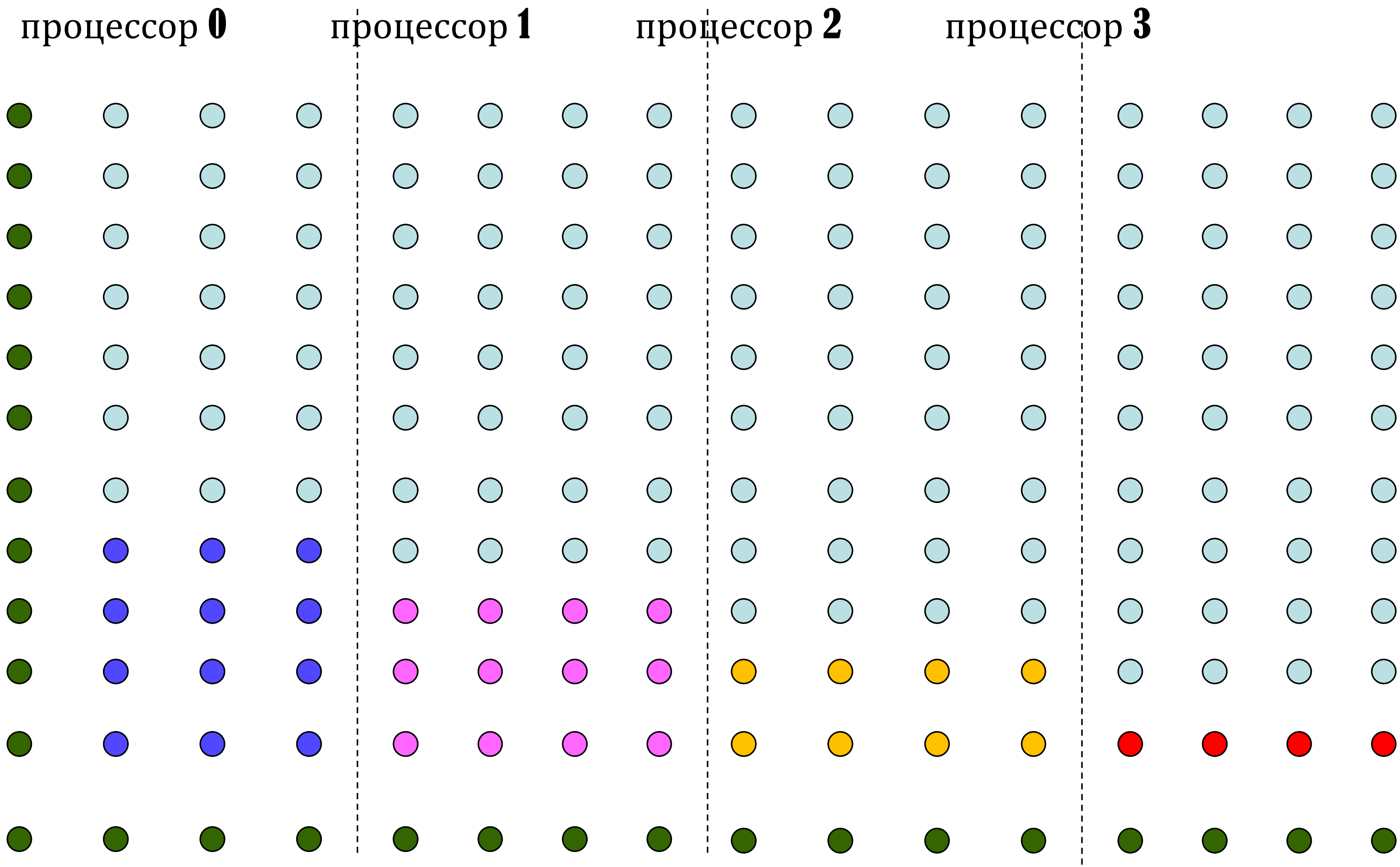
Метод конвейерного параллелизма



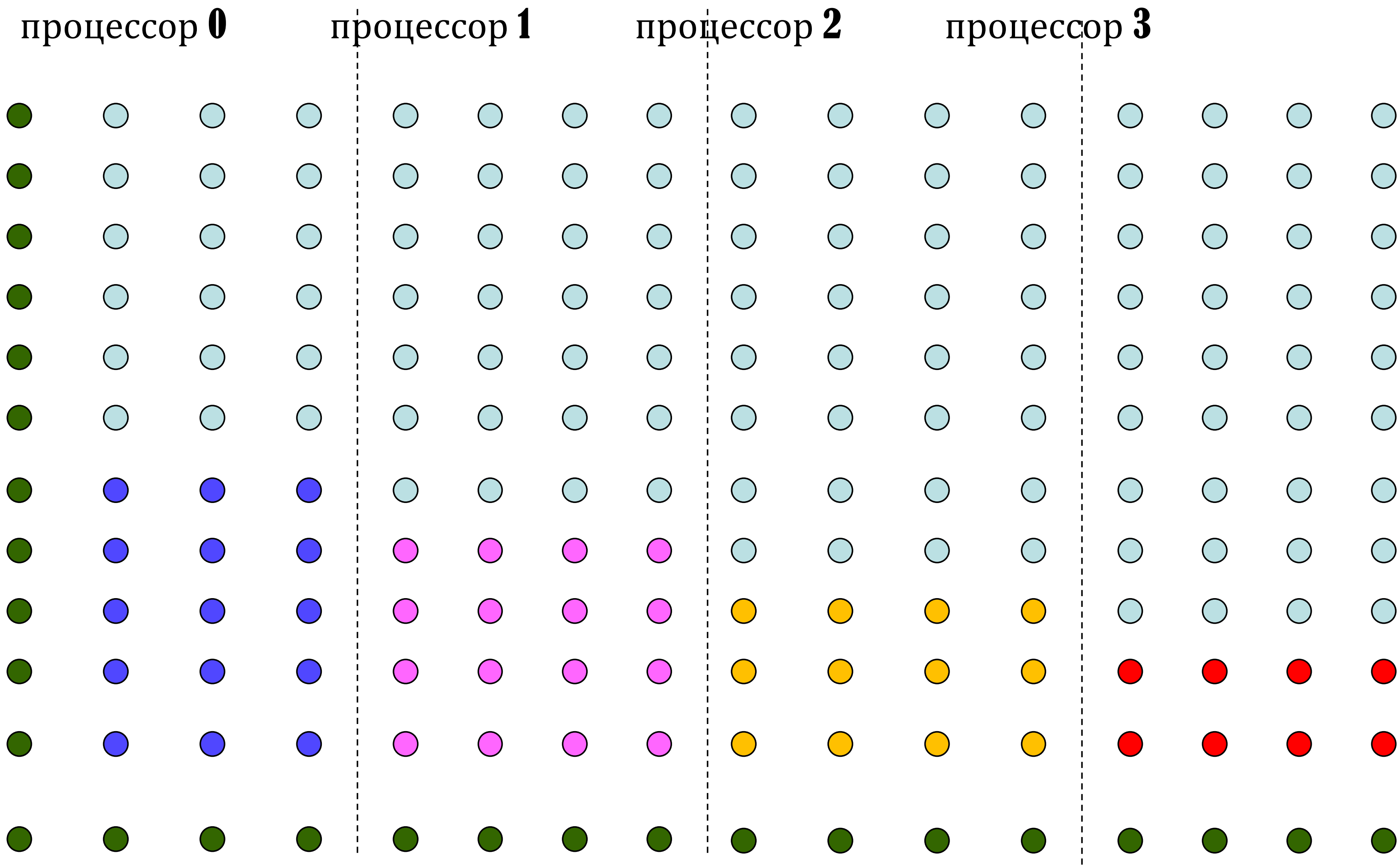
Метод конвейерного параллелизма



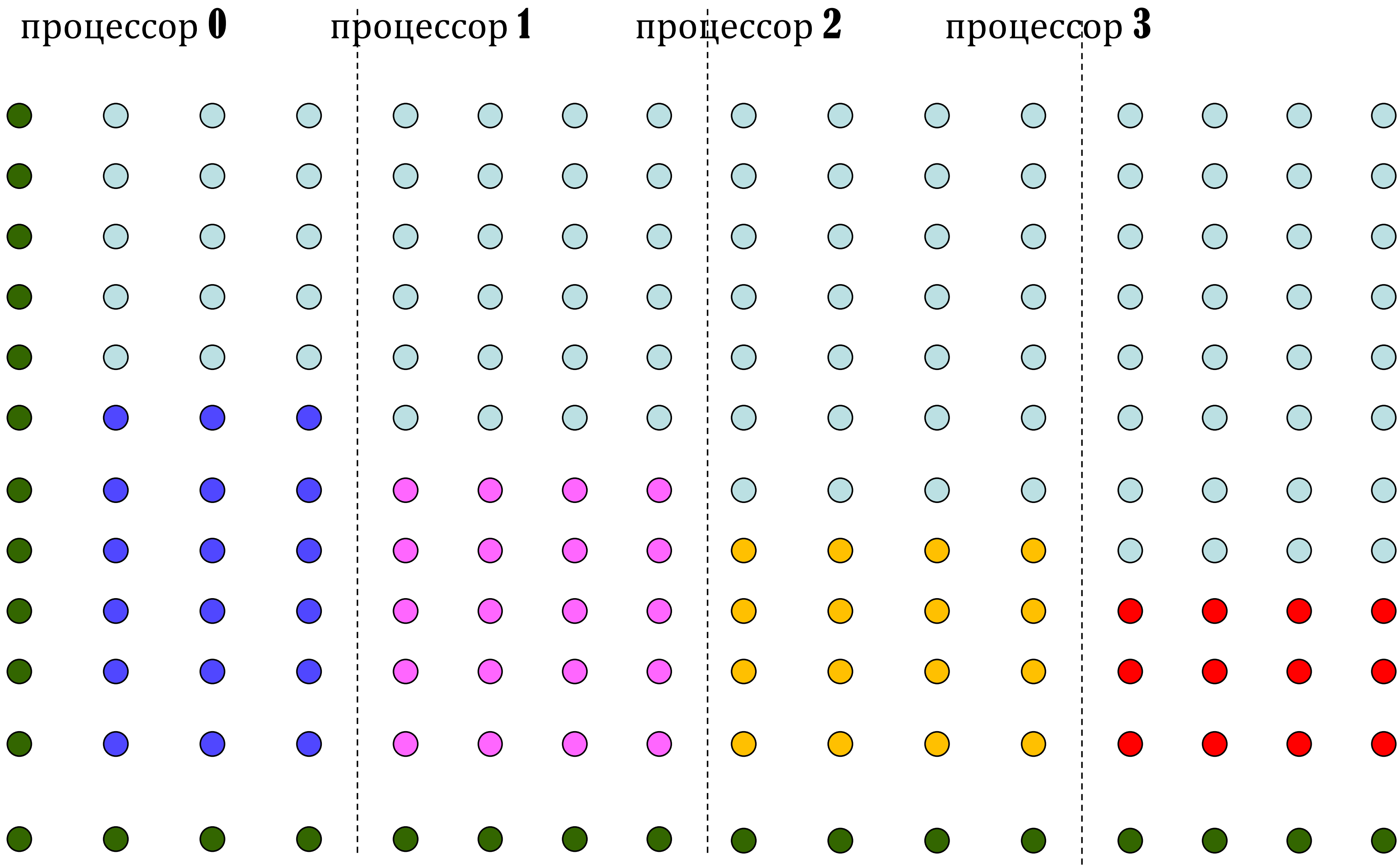
Метод конвейерного параллелизма



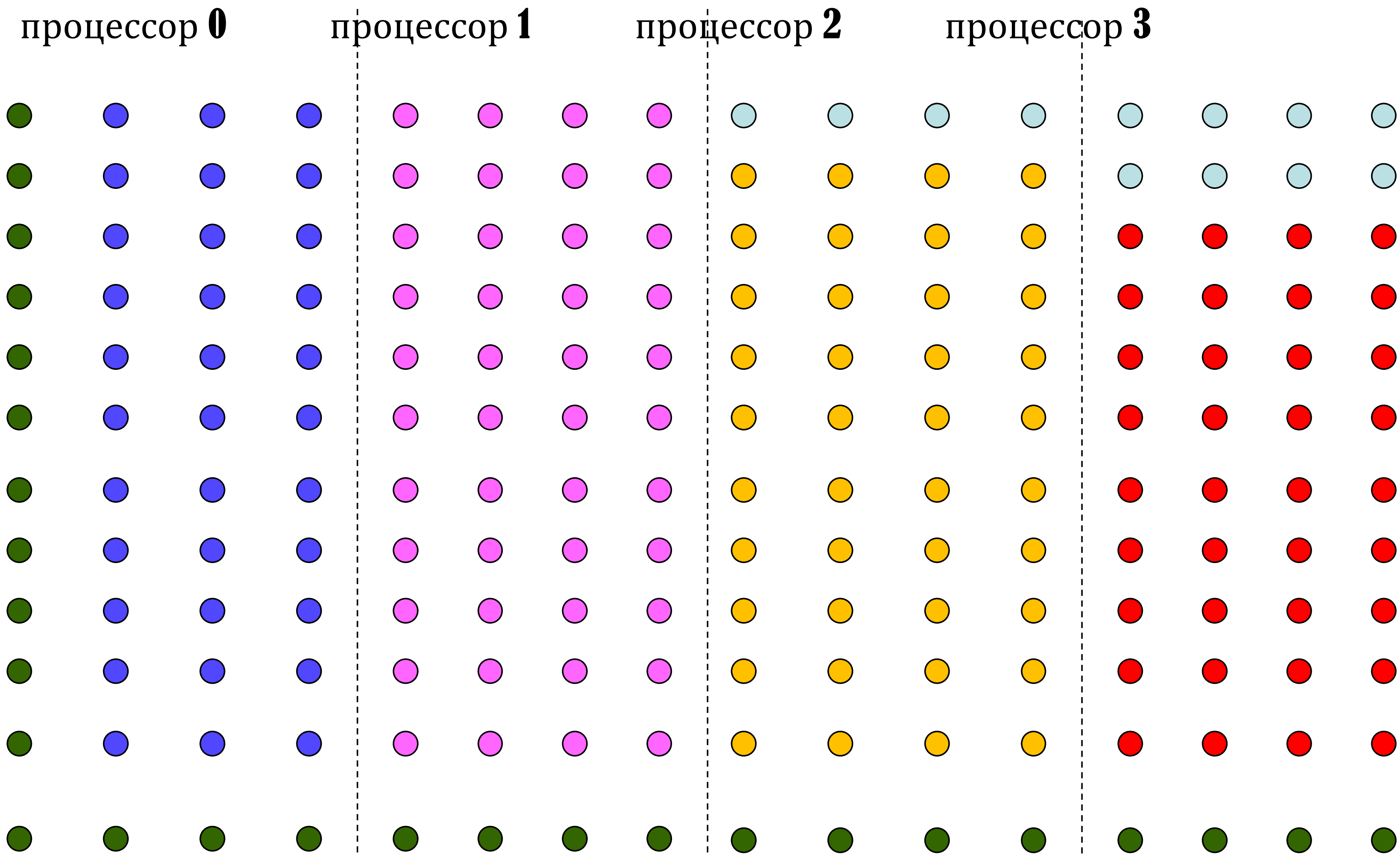
Метод конвейерного параллелизма



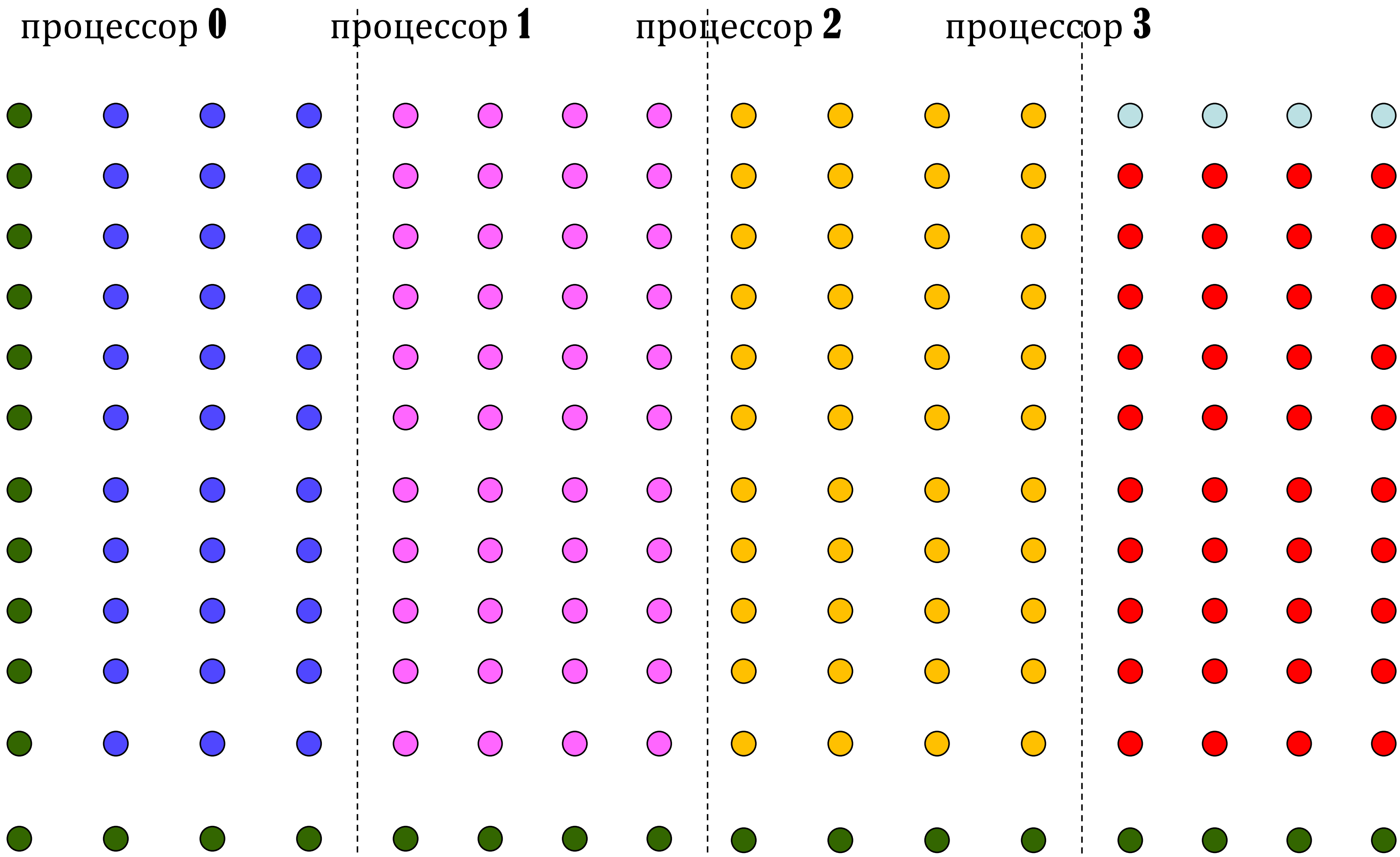
Метод конвейерного параллелизма



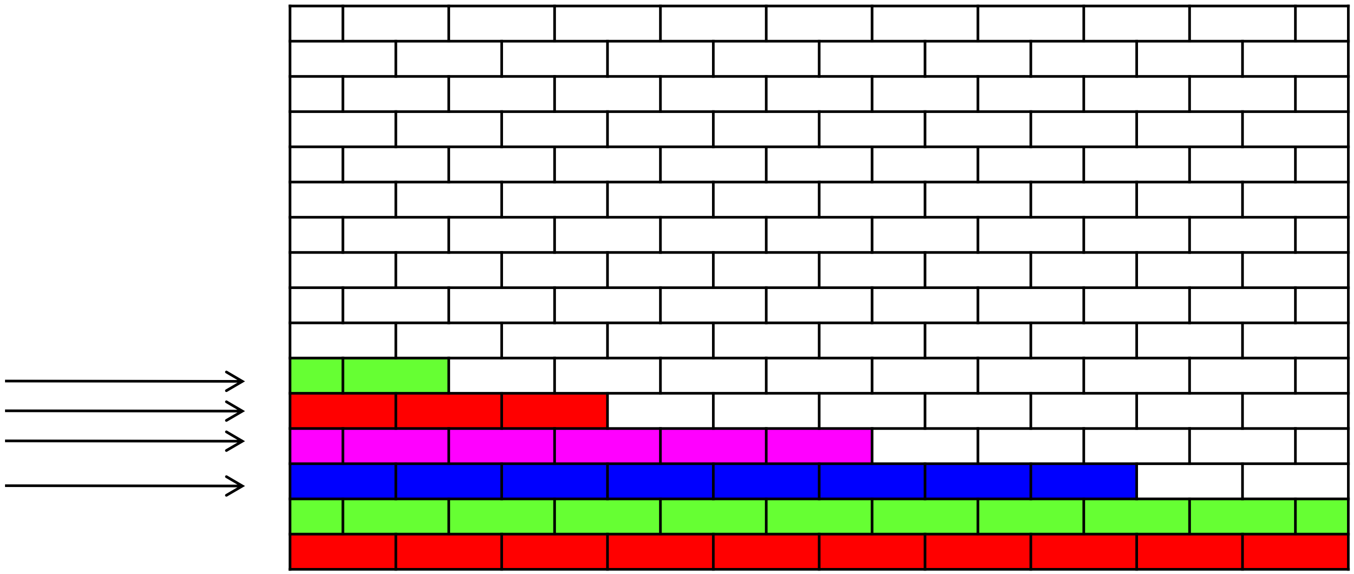
Метод конвейерного параллелизма



Метод конвейерного параллелизма



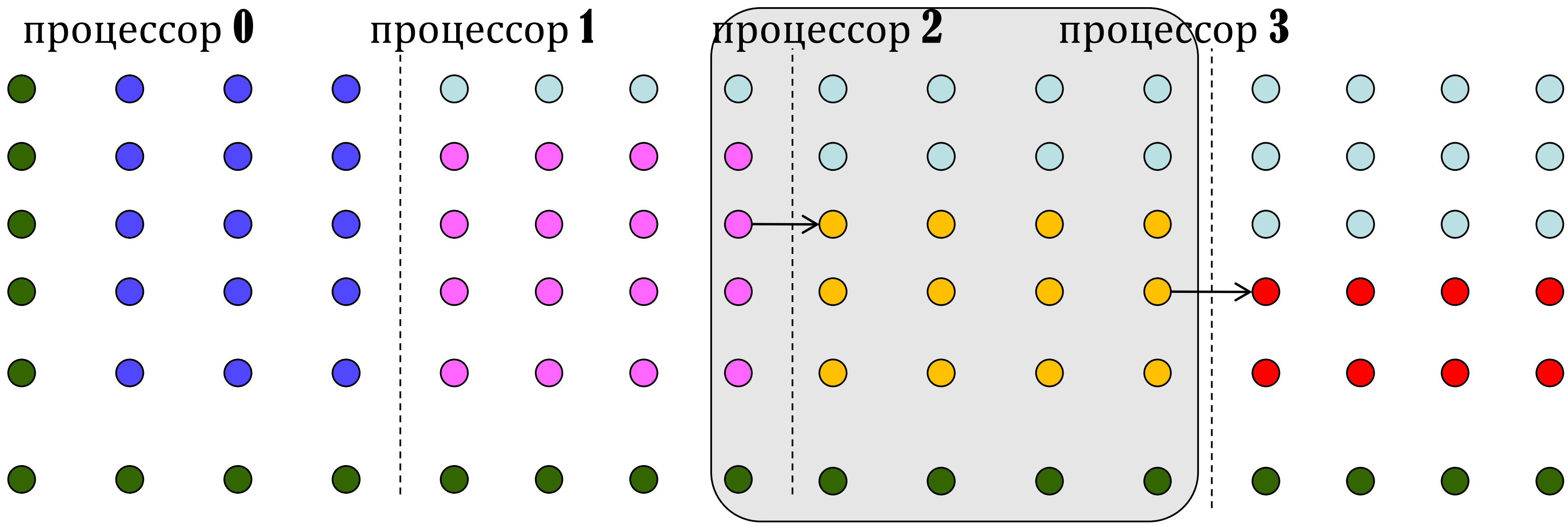
Метод конвейерного параллелизма



$$T_1(kn) = \tau_c kn \qquad T_p(kn) = \tau_c \frac{kn}{p} + k\frac{n}{p}\tau_s$$

$$S_p(kn) = p\frac{1}{1+\dfrac{\tau_s}{\tau_c}} \qquad E_p(kn) = \frac{1}{1+\dfrac{\tau_s}{\tau_c}}$$

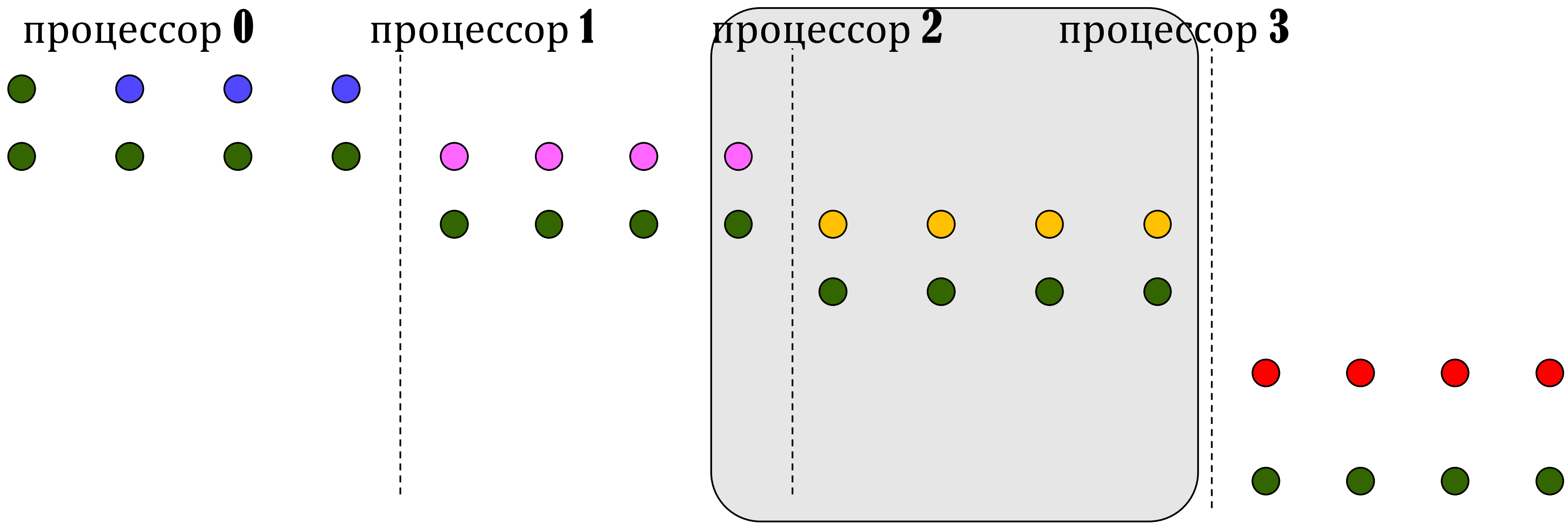
Метод конвейерного параллелизма



$$T_1(kn) = \tau_c kn \qquad T_p(kn) = \tau_c \frac{kn}{p} + 2k\tau_s$$

$$S_p(kn) = p\frac{1}{1 + 2\frac{p}{n}\frac{\tau_s}{\tau_c}} \qquad E_p(kn) = \frac{1}{1 + 2\frac{p}{n}\frac{\tau_s}{\tau_c}}$$

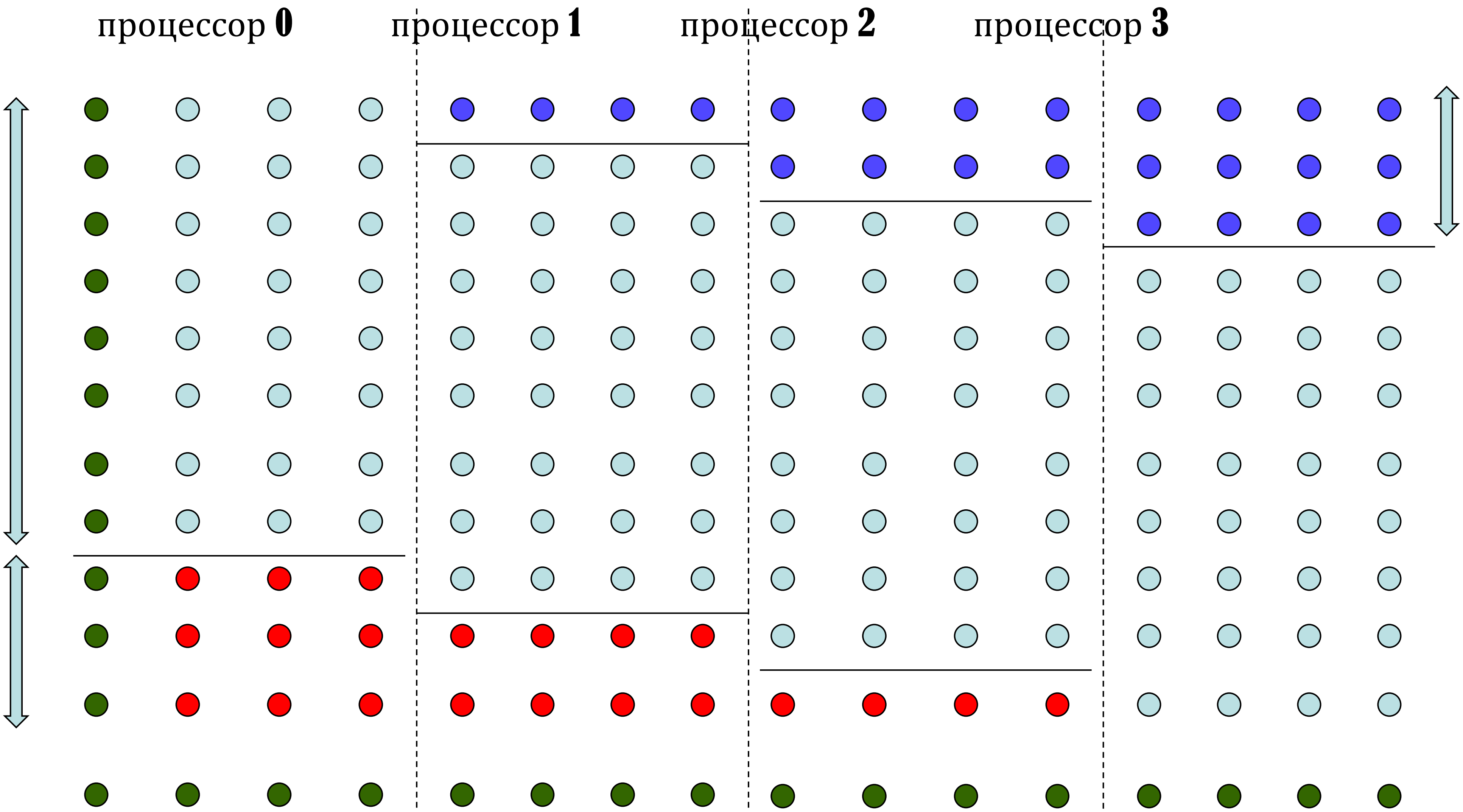
Объём хранимых данных



$$T_1(kn) = \tau_c kn \qquad T_p(kn) = \tau_c \frac{kn}{p} + 2k\tau_s$$

$$S_p(kn) = p \frac{1}{1 + 2\frac{p}{n}\frac{\tau_s}{\tau_c}} \qquad E_p(kn) = \frac{1}{1 + 2\frac{p}{n}\frac{\tau_s}{\tau_c}}$$

Учет стартовых и финальных затрат



Учет стартовых и финальных затрат

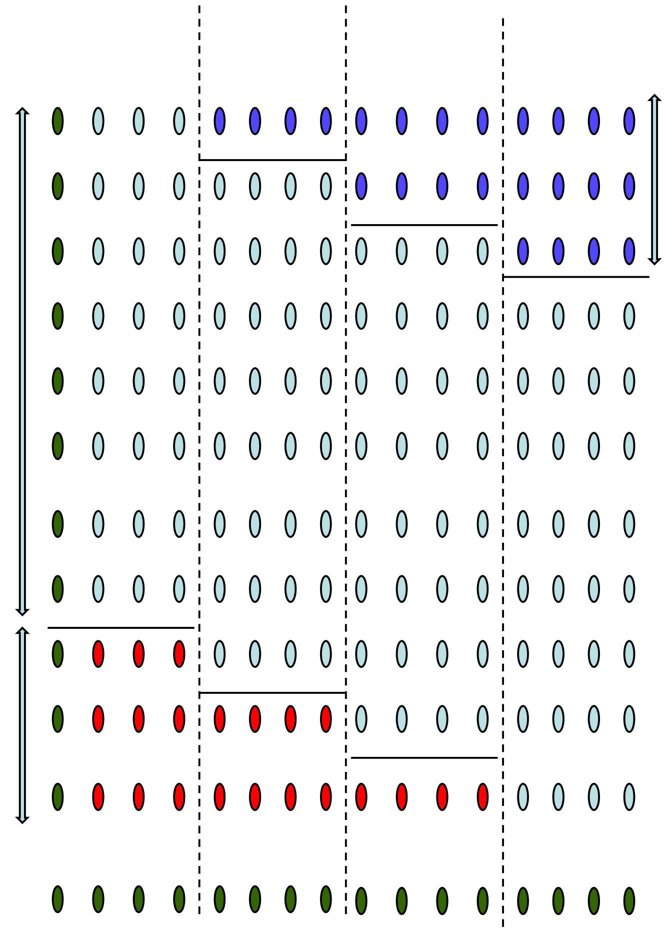
$k = 11$

$p = 4$

$$T_{end} = (p-1)\left(\frac{n}{p}\,\tau_c + 2\tau_s\right)$$

$$T_{middle} = (k+1-p)\left(\frac{n}{p}\,\tau_c + 2\tau_s\right)$$

$$T_{start} = (p-1)\left(\frac{n}{p}\,\tau_c + 2\tau_s\right)$$



$$T_{all} = T_{start} + T_{middle} + T\_end$$

Учет стартовых и финальных затрат

$$T^{all} = (p+k-1)\left(\frac{n}{p}\,\tau_c + 2\tau_s\right)$$

$$p + k \gg 1$$

$$S_p^{all} = p\,\frac{1}{\left(1+\frac{\mathrm{p}}{\mathrm{k}}\right)\left(1+2\frac{p}{n}\frac{\tau_s}{\tau_c}\right)}$$

$$E_p^{all} = \frac{1}{\left(1+\frac{\mathrm{p}}{\mathrm{k}}\right)\left(1+2\frac{p\tau_s}{n\tau_c}\right)}$$

Метод эффективен при $p \ll k$

Максимальная степень параллелизма: $\min(n, k)$

Максимальное ускорение: $\dfrac{pk}{p+k} \le \dfrac{k}{2}$

Простые алгоритмы

- Статическая и динамическая балансировка загрузки процессоров
  - Статическая балансировка
    - метод сдваивания
    - геометрический параллелизм
    - конвейерный параллелизм
  - Динамическая балансировка
    - коллективное решение
    - диффузная балансировка

Контакты

Якобовский М.В., *чл.-корр. РАН, проф., д.ф.-м.н.,*

Заместитель директора по научной работе Института прикладной математики им. М.В.Келдыша Российской академии наук

mail: lira@imamod.ru

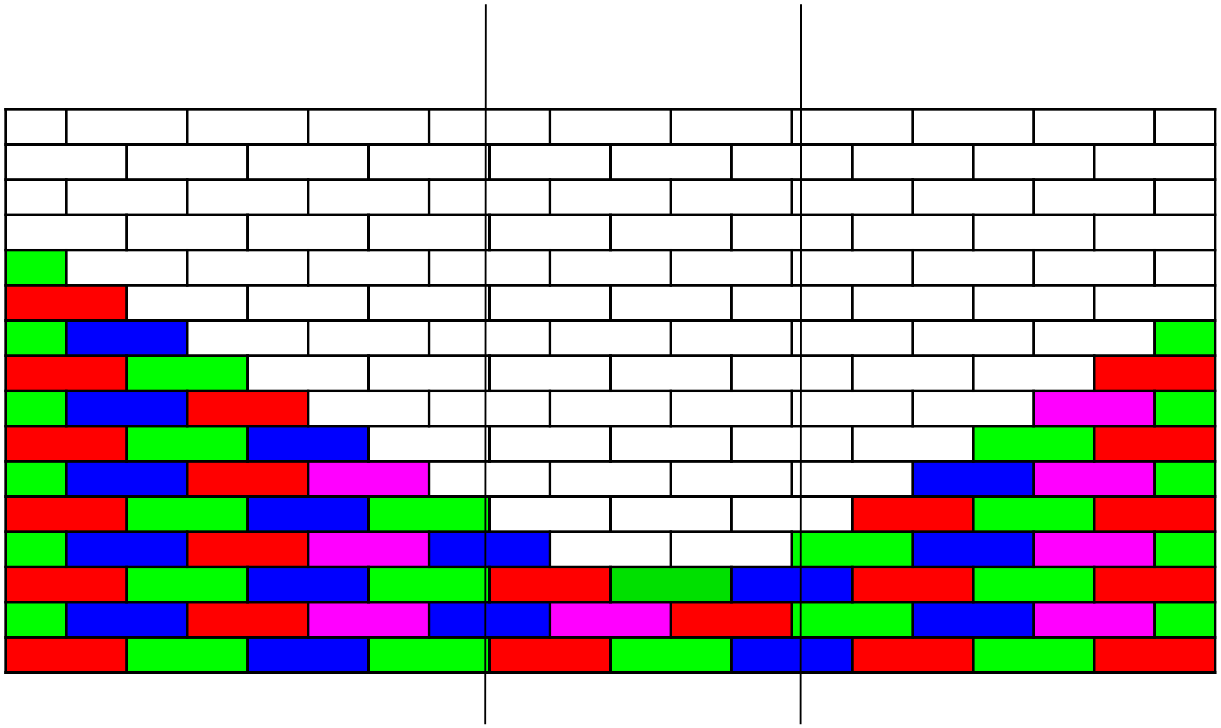
web: http://lira.imamod.ru

Диффузная балансировка

- Причины дисбаланса вычислительной нагрузки
  - Разные процессоры
  - Внешнее воздействие
  - Разная вычислительная сложность заданий

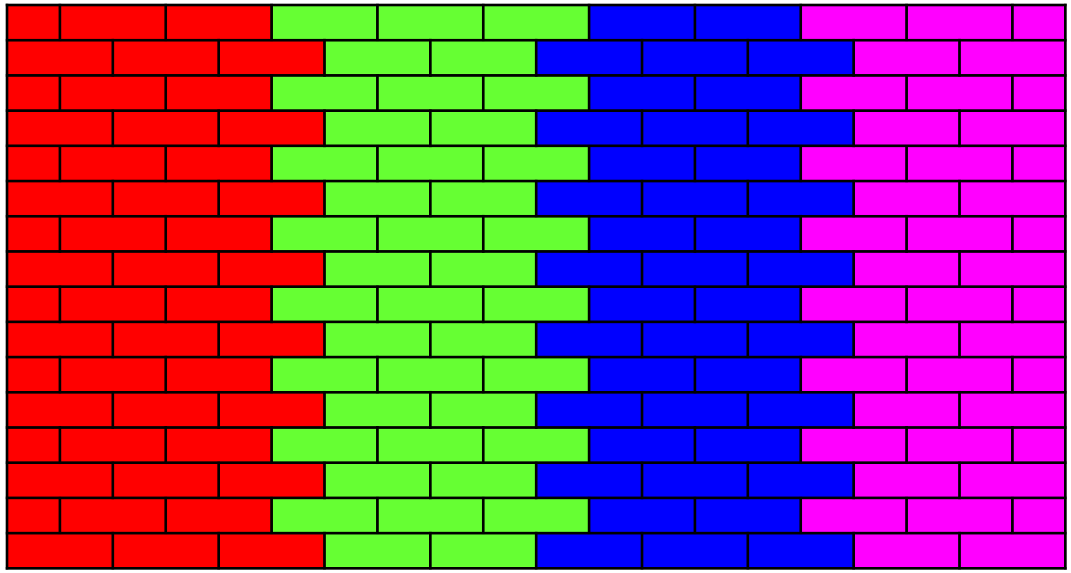
- Результат дисбаланса
  - Эффективная производительность определяется самым медленным процессором

Медленный процессор

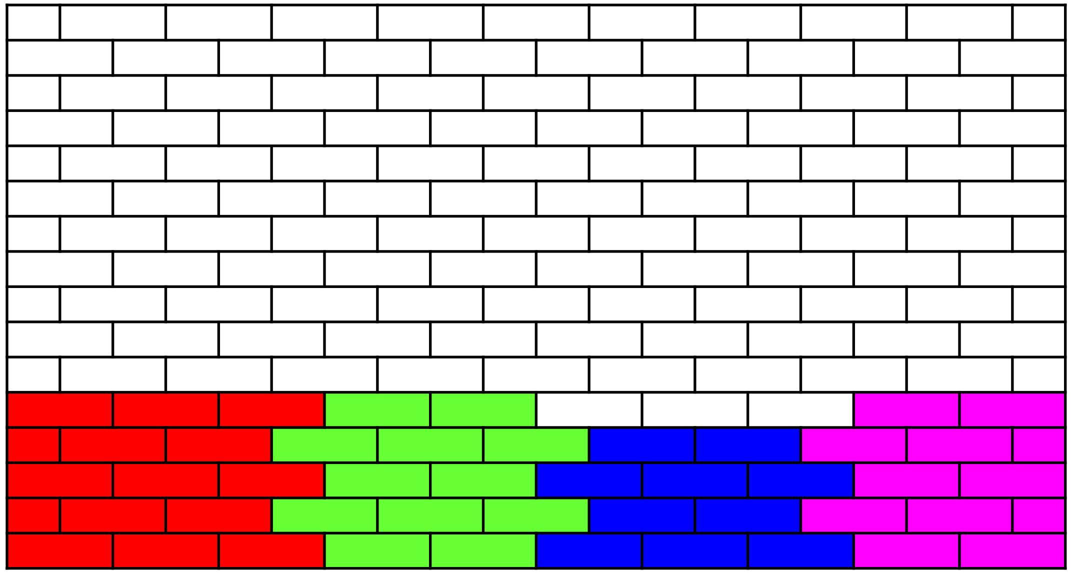


Какой объем работ забрать у среднего процессора и кому его передать?

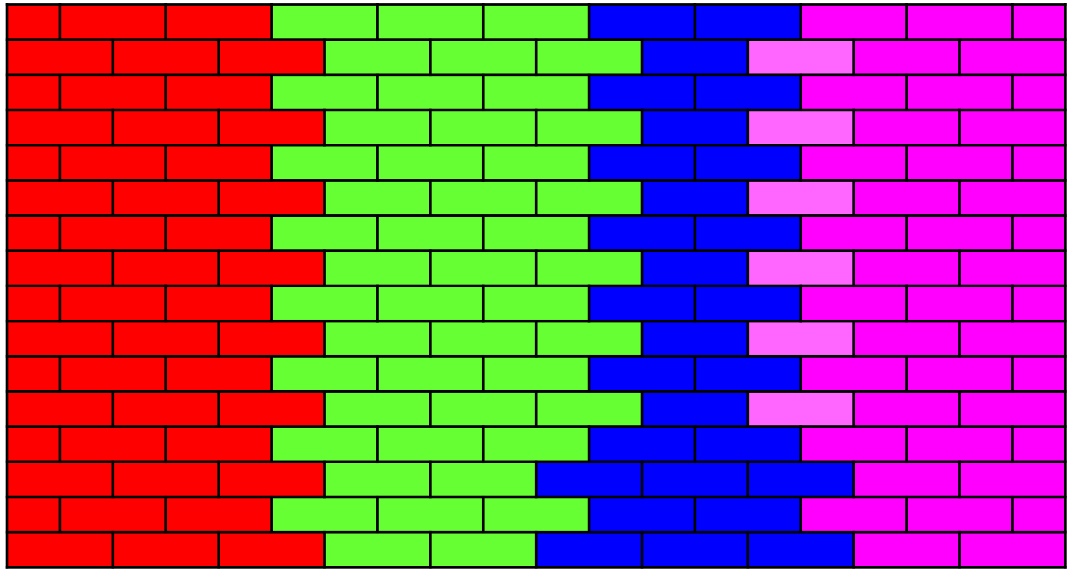
Метод геометрического параллелизма



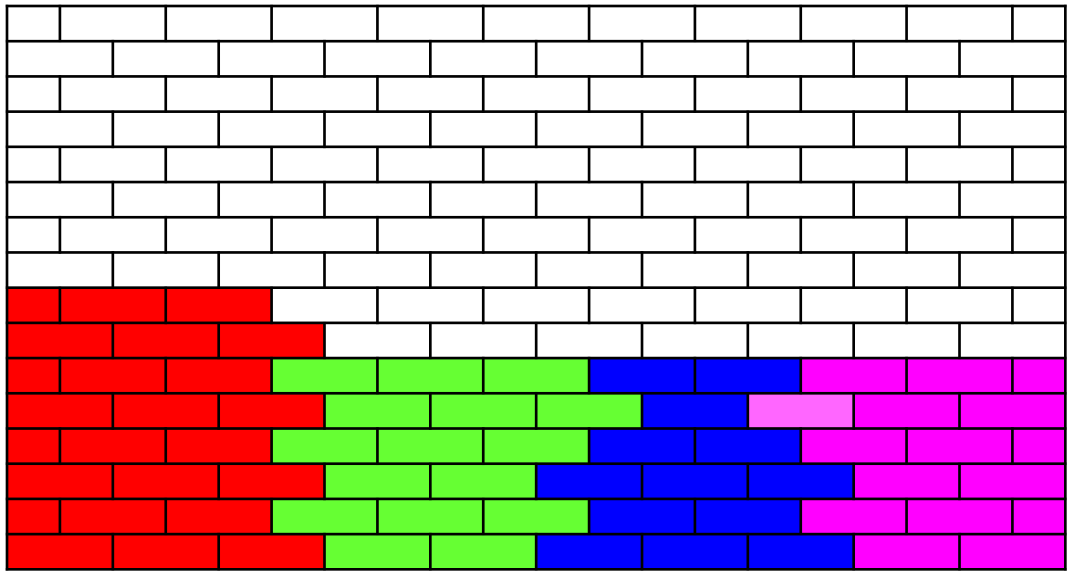
Метод геометрического параллелизма



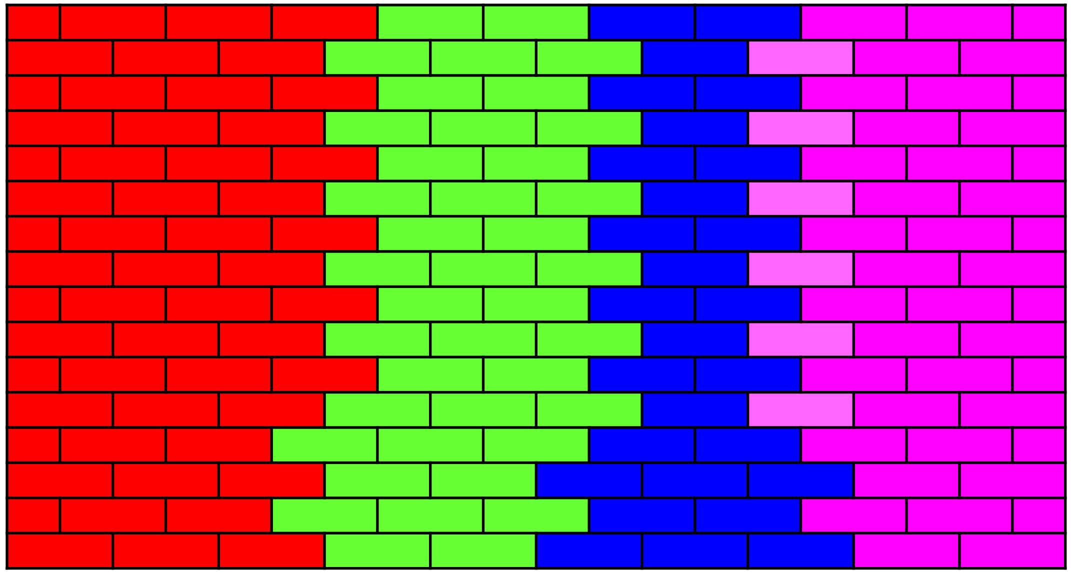
Диффузная балансировка загрузки



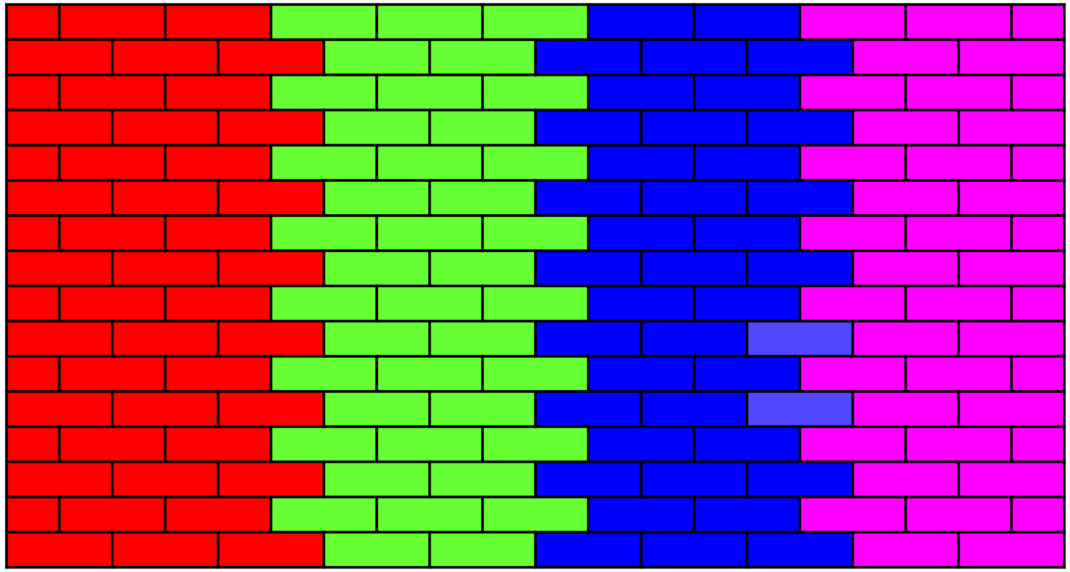
Диффузная балансировка загрузки



Диффузная балансировка загрузки



Статическое распределение



Постановка задачи диффузной балансировки

Дано:

- Количество точек – N
- Количество процессоров – p
- Процессор i обработал n_i точек за время t_i
- Для обработки любой точки требуется одинаковое число операций

Требуется:

- Найти количества точек n'_i , которое следует обработать процессорам на следующем шаге
- Определить сколько точек каждый из процессоров должен передать соседним процессорам

Диффузная балансировка

$$n_i' = N \frac{\frac{n_i}{t_i}}{\sum\limits_{j=0}^{p-1} \frac{n_j}{t_j}}$$

В предположении одинаковой трудоёмкости обработки каждой из точек

если Дисбаланс вычислительной нагрузки обусловлен разными производительностями процессоров

$$n_i' = n \frac{\dfrac{n_i}{t_i}}{\sum\limits_{j=1}^{p} \dfrac{n_j}{t_j}}$$

• если Дисбаланс обусловлен различной трудоемкостью обработки узлов сетки

процессоры обладают одинаковой производительностью

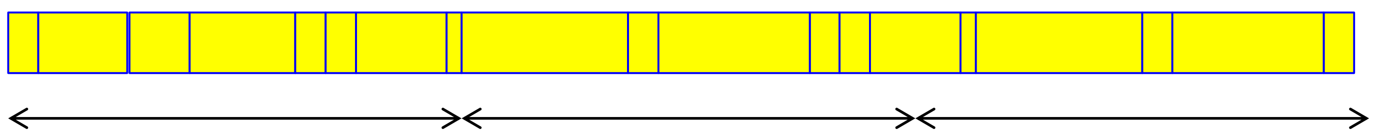
Общая вычислительная нагрузка определяется как

$$T = \sum_{k=1}^{p} t_i$$

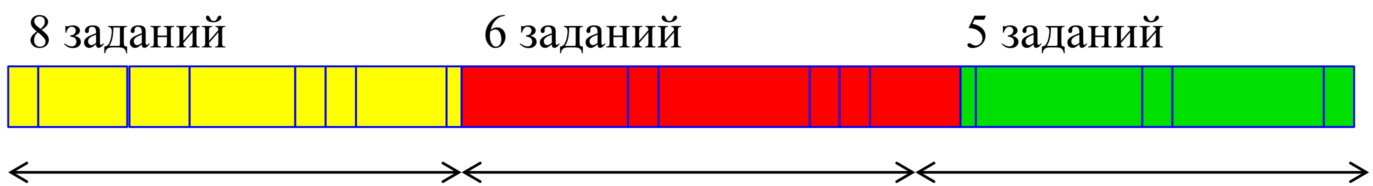
$$T_{mid} = \frac{T}{p}$$

Алгоритм распределения заданий по одинаковым процессорам

```
m=0;   t=0;
for(j=0;j<n;j++)
{
   t=t+σ_j
   if(t ≥ T_mid)   {          n''_i=j-m;   m=j;   t=0;   }
}
```



Каждому процессору назначается множество очередных заданий, суммарная трудоёмкость которых равна средней по всем процессорам



Перераспределение узлов сетки

номер i процессора	1	2	3	4
n_i	5	10	14	7
t_i	1	2	1	3
n_i'	7	7	19	3
n_i''	10	10	11	5

Два разных решения

Какое верно?

Почему одно из них верно?

Контакты

Якобовский М.В.

чл.-корр. РАН, проф., д.ф.-м.н.,

заместитель директора по научной работе Института прикладной математики им. М.В. Келдыша Российской академии наук

mail: lira@imamod.ru

web: http://lira.imamod.ru