



Ministerstwo Nauki
i Szkolnictwa Wyższego

**Regionalna Inicjatywa Doskonałości w Dyscyplinach
Informatyki, Elektrotechniki, Elektroniki, Automatyki i Robotyki
na Politechnice Częstochowskiej**



Projekt w ramach programu Regionalna Inicjatywa Doskonałości, decyzja nr 020/RID/2018/19

Architektura Systemów Komputerowych

Analiza architektur i ocena wydajności przykładowych systemów komputerowych

dr hab. inż. Łukasz Szustak, prof. PCz
lszustak@icis.pcz.pl

Katedra Informatyki
Wydział Inżynierii Mechanicznej i Informatyki
Politechnika Częstochowska



Wydział Inżynierii
Mechanicznej
i Informatyki
The Faculty of Mechanical Engineering
and Computer Science



Ministerstwo Nauki
i Szkolnictwa Wyższego

Regionalna Inicjatywa Doskonałości w Dyscyplinach Informatyki, Elektrotechniki, Elektroniki, Automatyki i Robotyki na Politechnice Częstochowskiej



Projekt w ramach programu Regionalna Inicjatywa Doskonałości, decyzja nr 020/RID/2018/19

**Zamieszczane materiały służą wyłącznie do celów
indywidualnego kształcenia. Nie wyrażam zgody na ich
utrwalanie przekazywanie osobom trzecim ani
rozpowszechnianie.**

dr hab. inż. Łukasz Szustak, prof. PCz

lszustak@icis.pcz.pl

Katedra Informatyki

Wydział Inżynierii Mechanicznej i Informatyki
Politechnika Częstochowska



Wydział Inżynierii
Mechanicznej
i Informatyki
The Faculty of Mechanical Engineering
and Computer Science

Ocena wydajności procesorów: Model Roofline

- Teoretyczna ocena wydajności procesorów może zostać przeprowadzona w oparciu o model „Roofline” (Roofline Model - Samuel Williams, Andrew Waterman, and David Patterson)
- Model Roofline bazuje na dwóch parametrach:
 - Intensywność obliczeń (operational intensity)
 - Teoretyczna osiągalna wydajność (attainable performance)

Ocena wydajności procesorów: intensywność obliczeń

- Model Roofline bazuje na parametrze, który „przekształca” wydajność przepustowości danych pomiędzy procesorem a pamięcią główną (GB/s) na intensywność obliczeń (operational intensity)
- Miarą intensywności obliczeń **IO** jest stosunek liczby operacji do ilości danych potrzebnych do wykonania tych operacji [flop/byte]
- Przykładowo, dla typu float (rozmiar 4 bajty):

```
for( int i=0; i<N; ++i)
  for( int j=0; j<M; ++j)
    for(int k=0; k<L; ++k)
      C[i][j] += A[i][k] * B[k][j]
```

- Dla każdej iteracji najbardziej zagnieżdżonej pętli realizowana jest **jedna** operacja dodawania oraz **jedna** operacje mnożenia:
zatem liczba operacji wynosi 2

Ocena wydajności procesorów: intensywność obliczeń

- Model Roofline bazuje na parametrze, który „przekształca” wydajność przepustowości danych pomiędzy procesorem a pamięcią główną (GB/s) na intensywność obliczeń (operational intensity)
- Miarą intensywności obliczeń **IO** jest stosunek liczby operacji do ilości danych potrzebnych do wykonania tych operacji [flop/byte]
- Przykładowo, dla typu float (rozmiar 4 bajty):

```
for( int i=0; i<N; ++i)
  for( int j=0; j<M; ++j)
    for(int k=0; k<L; ++k)
      C[i][j] += A[i][k] * B[k][j]
```

- **Liczba operacji wynosi 2**
- Dla każdej iteracji najbardziej zagnieżdżonej pętli zostaną pobrane dwa elementy z pamięci: A[i][k] oraz B[k][j]
 - Uwaga element C[i][j] nie zależy od indeksu k najbardziej zagnieżdżonej pętli, dlatego nie jest na tym etapie brany pod uwagę
- **Zatem ilość pobieranych danych wynosi 2*4B**

Ocena wydajności procesorów: intensywność obliczeń

- Model Roofline bazuje na parametrze, który „przekształca” wydajność przepustowości danych pomiędzy procesorem a pamięcią główną (GB/s) na intensywność obliczeń (operational intensity)
- Miarą intensywności obliczeń **IO** jest stosunek liczby operacji do ilości danych potrzebnych do wykonania tych operacji [flop/byte]
- Przykładowo, dla typu float (rozmiar 4 bajty):

```
for( int i=0; i<N; ++i)
  for( int j=0; j<M; ++j)
    for(int k=0; k<L; ++k)
      C[i][j] += A[i][k] * B[k][j]
```

- **Liczba operacji wynosi 2**
- **Ilość wymaganych danych = 2*4B = 8 bajtów**
- $IO = \frac{2 \text{ operacje}}{8 \text{ bajtów}} \left[\frac{\text{flop}}{\text{byte}} \right]$

Ocena wydajności procesorów: teoretyczna osiągalna wydajność

- Teoretyczna osiągalna wydajność określa maksymalną wydajność jaką dana aplikacja może uzyskać na konkretnej architekturze
- Teoretyczna osiągalna wydajność jest wyznaczana na podstawie równania:

$$\textit{Teoretyczna osiągalna wydajność} = \min \left\{ \begin{matrix} R_p \\ MB_p * IO \end{matrix} \right.$$

- Gdzie,

R_p - oznacza maksymalną teoretyczną wydajność

MB_p - oznacza maksymalną przepustowość danych do pamięci

IO – oznacza Intensywność obliczeń

Ocena wydajności procesorów: teoretyczna osiągalna wydajność

Przykładowo:

- Teoretyczna wydajność procesora Intel Xeon Platinum 8280 wynosi (opcja Turbo, wszystkie rdzenie, operacja dodawania (bez FMA), wektorowo AVX-512, float):

$$R_p = 2,4 \text{ GHz} * 28 \text{ rdzenie} * 2 \text{ jednostki FPU} * 16 \text{ SIMD float} * 1 \text{ operacja} = 2150.4 \left[\frac{\text{Gflop}}{\text{s}} \right]$$

- Teoretyczna maksymalna przepustowość MB_P danych do pamięci dla tego procesora wynosi ok $MB_P = 140 \left[\frac{\text{GB}}{\text{s}} \right]$

$$\text{Teoretyczna osiągalna wydajność} = \min \left\{ \begin{array}{l} 2150.4 * 10^9 \left[\frac{\text{flop}}{\text{s}} \right] \\ 140 * 10^9 \left[\frac{\text{byte}}{\text{s}} \right] * IO \left[\frac{\text{flop}}{\text{byte}} \right] \end{array} \right.$$

Ocena wydajności procesorów: teoretyczna osiągalna wydajność

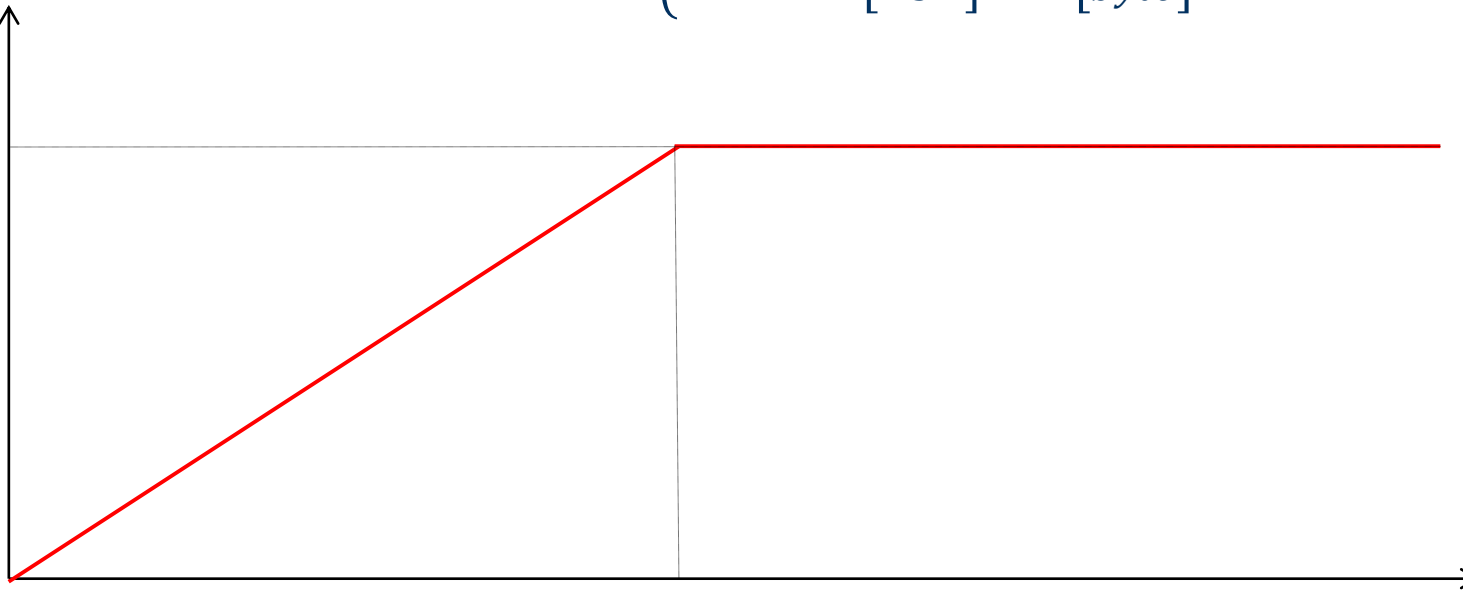
$$f(IO) = \min \left\{ \begin{array}{l} 2150.4 * 10^9 \left[\frac{flop}{s} \right] \\ 140 * 10^9 \left[\frac{byte}{s} \right] * IO \left[\frac{flop}{byte} \right] \end{array} \right.$$

Teoretyczna osiągalna
wydajność $\left[\frac{flop}{s} \right]$

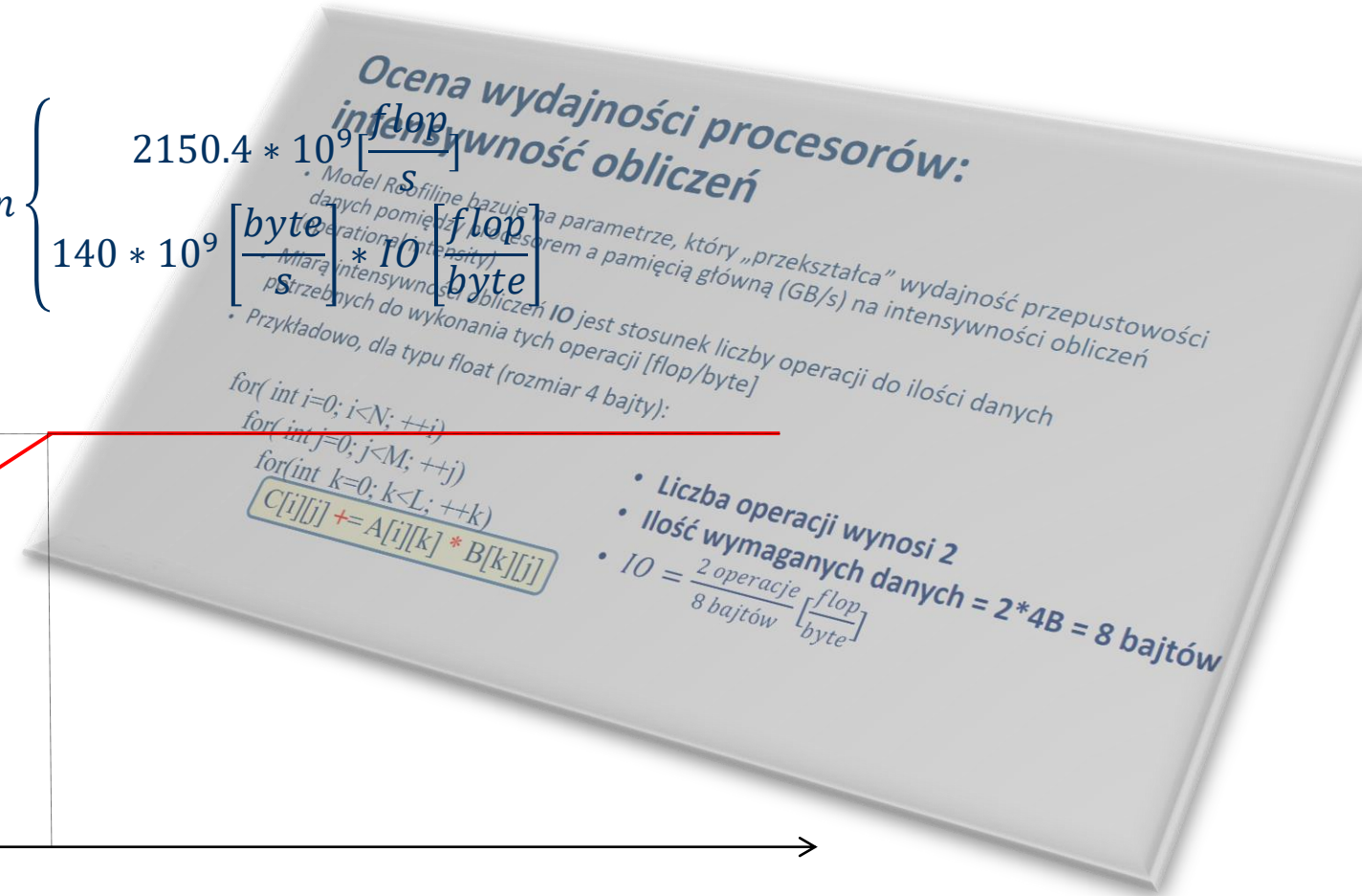
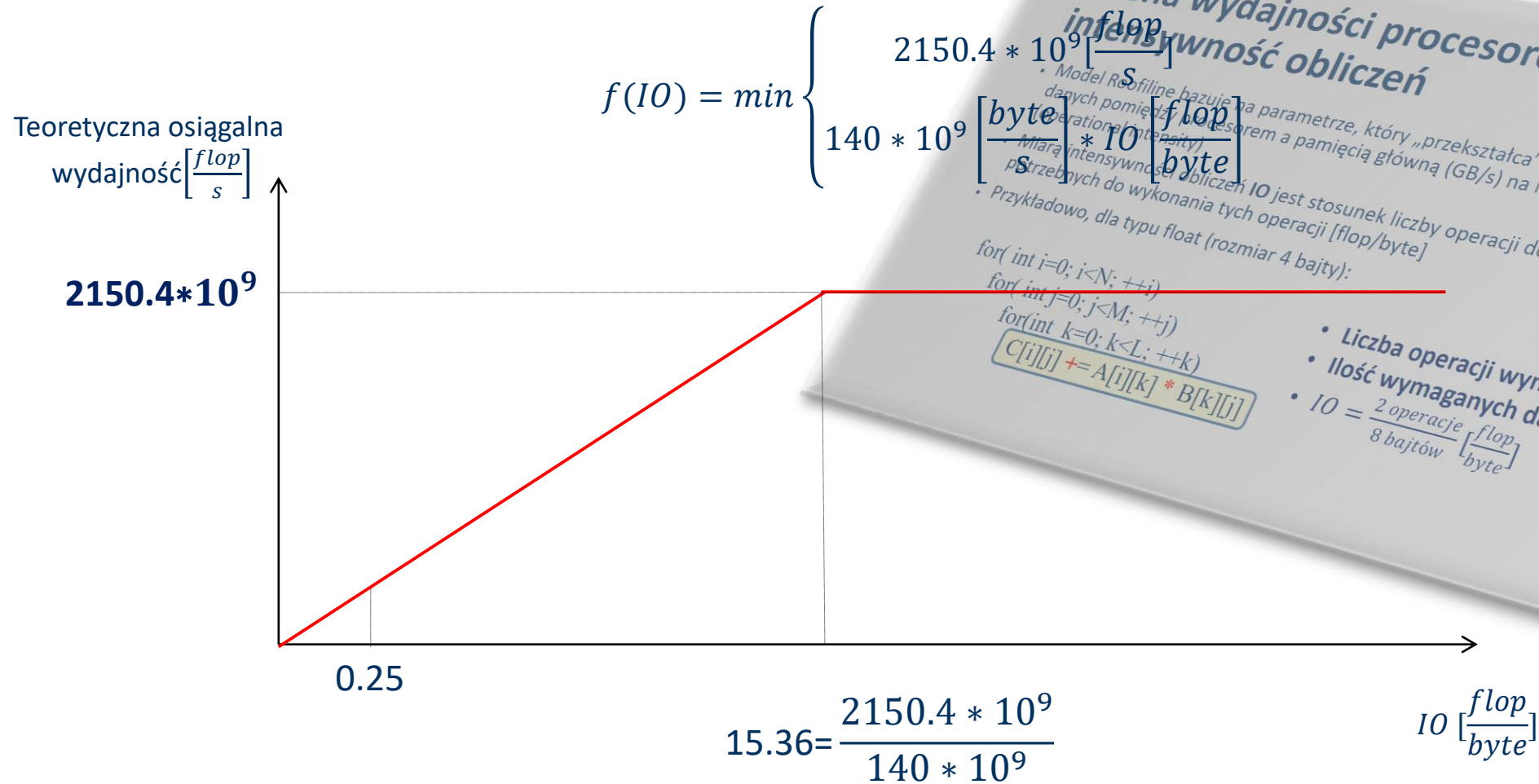
$2150.4 * 10^9$

$$15.36 = \frac{2150.4 * 10^9}{140 * 10^9}$$

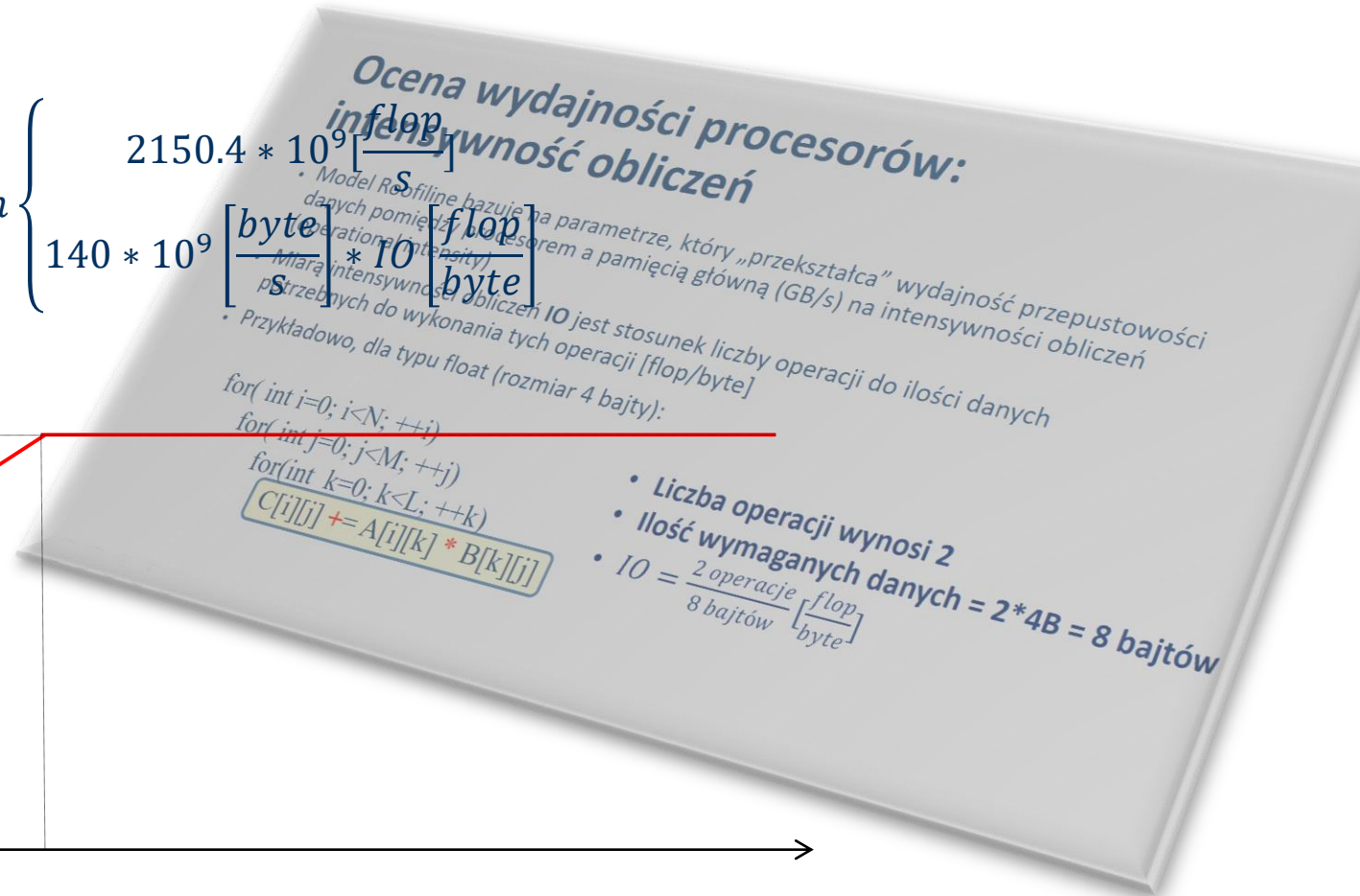
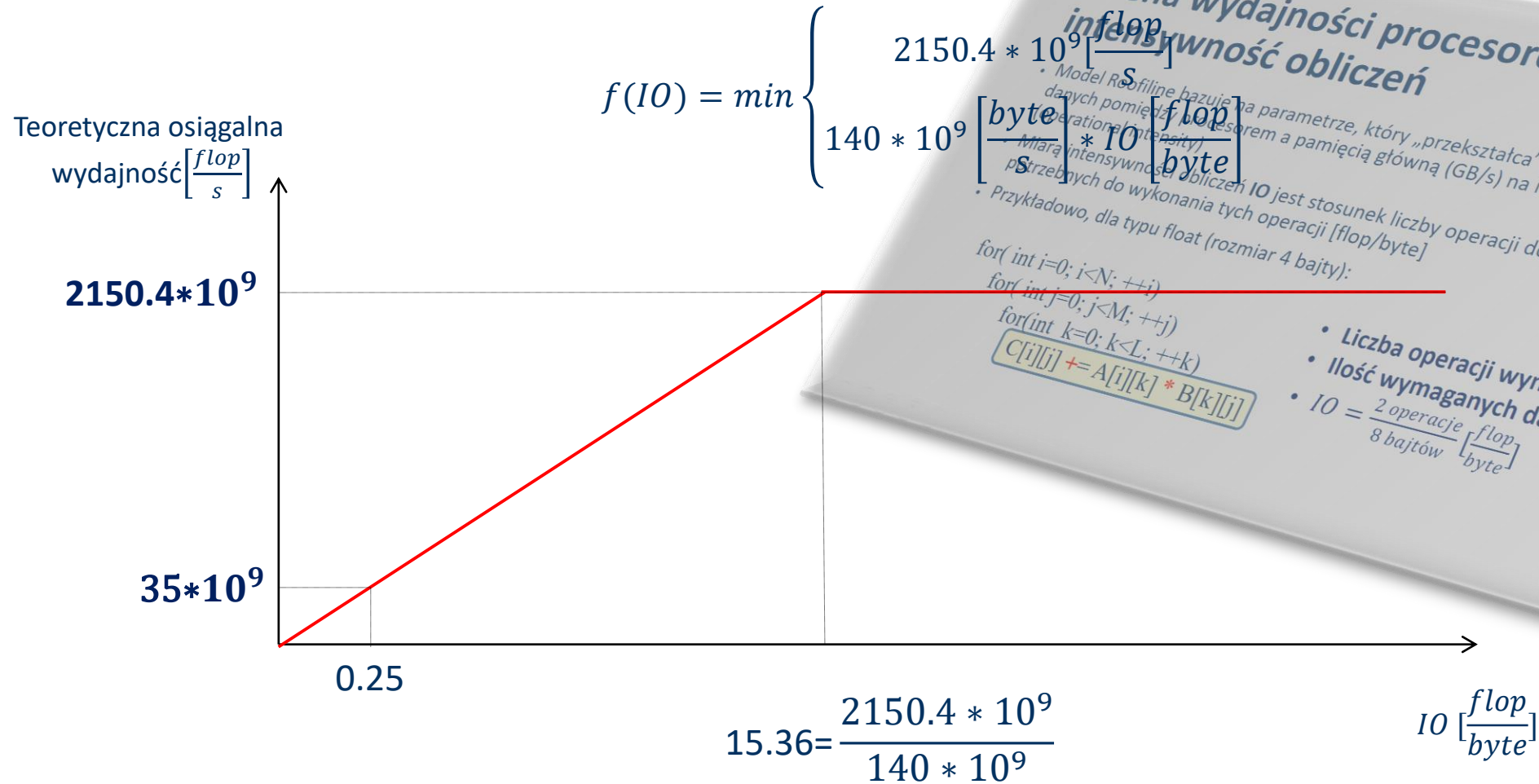
$IO \left[\frac{flop}{byte} \right]$



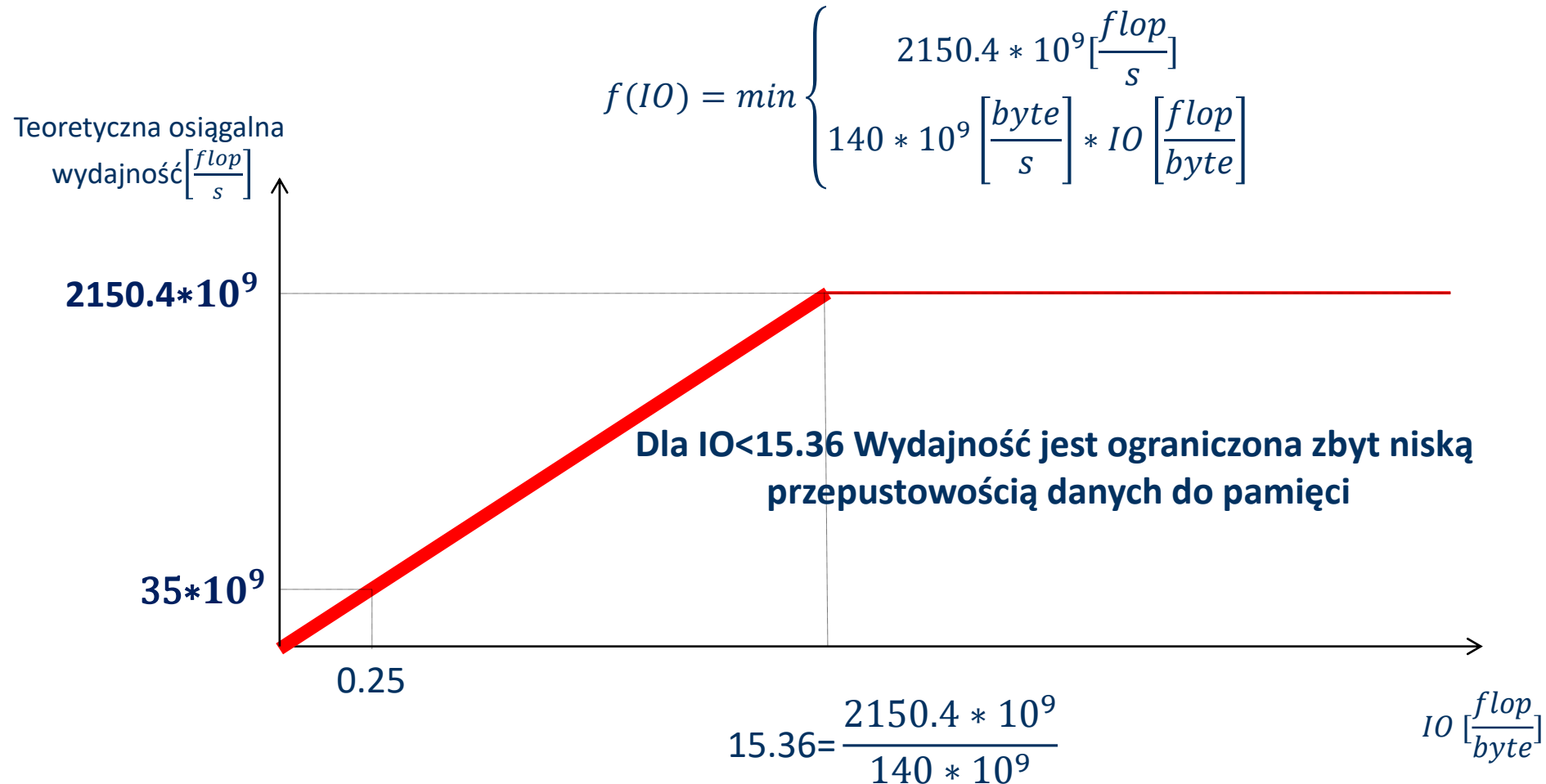
Ocena wydajności procesorów: teoretyczna osiągalna wydajność



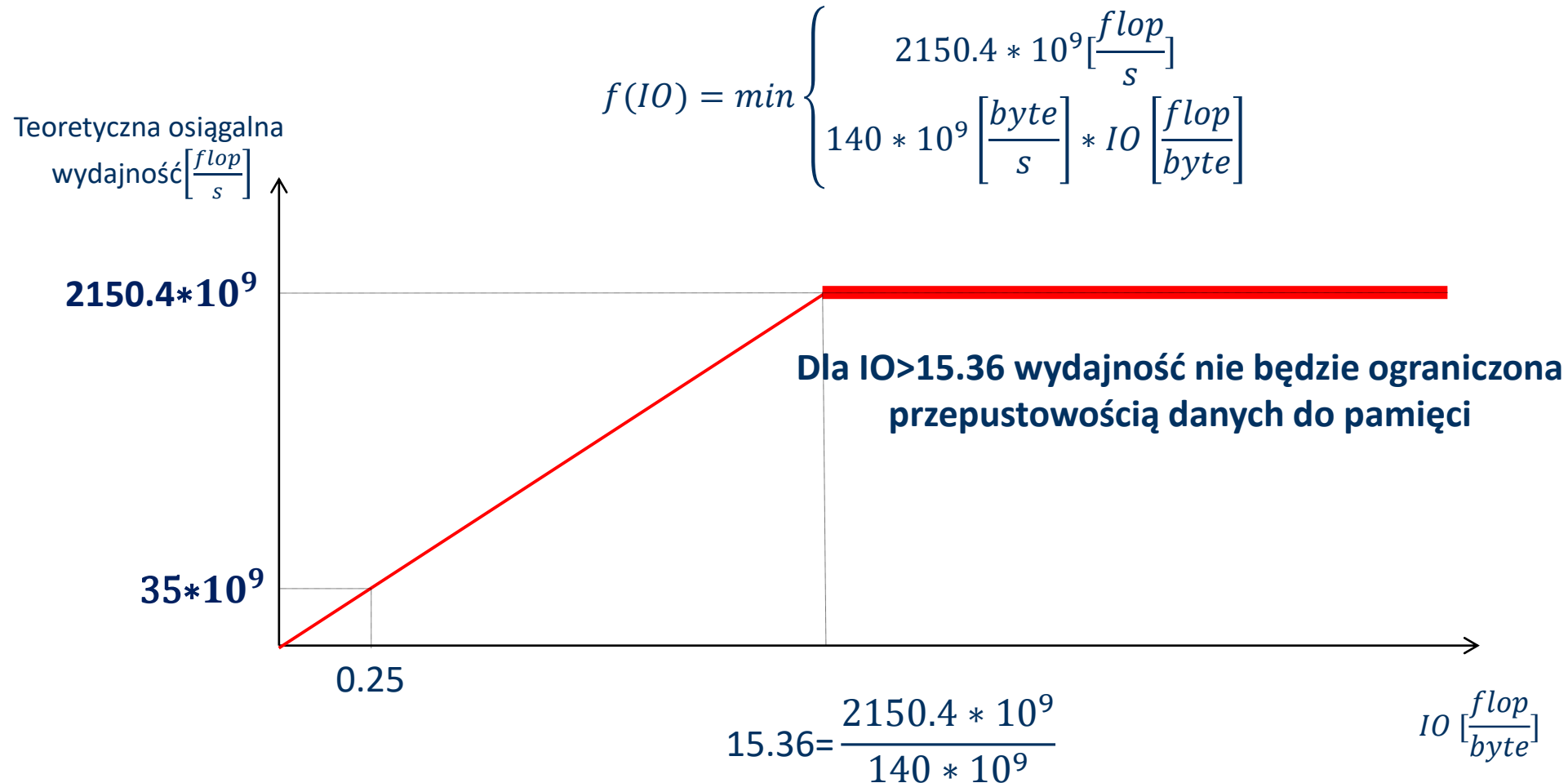
Ocena wydajności procesorów: teoretyczna osiągalna wydajność



Ocena wydajności procesorów: teoretyczna osiągalna wydajność



Ocena wydajności procesorów: teoretyczna osiągalna wydajność



Dziękuję za uwagę



Kontakt:

dr hab. inż. Łukasz Szustak, prof. PCz

lszustak@icis.pcz.pl

Katedra Informatyki

Wydział Inżynierii Mechanicznej i Informatyki

Politechnika Częstochowska