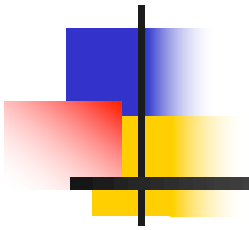
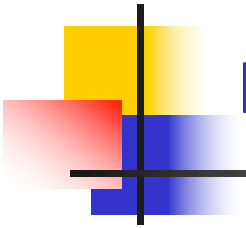


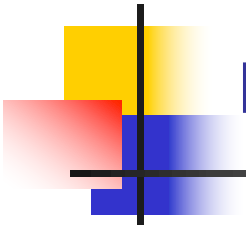
IN3030 L08v24 – About Moore's Law, the speed of light, latency

Eric Jul
Programming Technology Group
Department of Informatics
University of Oslo



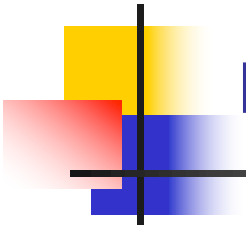
Review F7

I. Prime Numbers, review
II. Oblig Prime Numbers -- Review
III. Tidtagning
    I. JIT compilation
    II. Operativsystem?
    III. Søppel/Garbage Collection
IV. JMH
V. Amdahl
VI. Gustavson



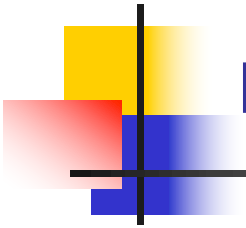
Plan for F8

I. Moore's Law
II. Performance improvements in processor power
III. Speed of light
IV. Why distribution?
V. How to present your timing results
VI. Hva er PRAM modellen - og hvorfor er den ubrukelig for oss



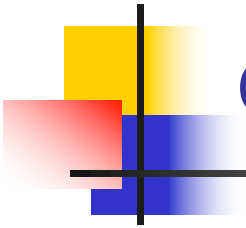
Moore's Law

- Used to be known as: «CPU speeds double every 2 years»



Moore’s Law

- Who knows this law?



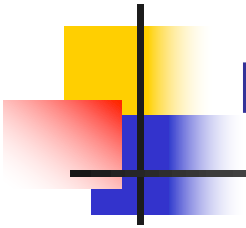
Original Moore's Law

- Transistors per square cm doubles:
- 1965 Article: Every year
- 1975 Article: Every two years

Perspektiv: Utvikling i CPU, minne, nettverk og mere om Moore's law

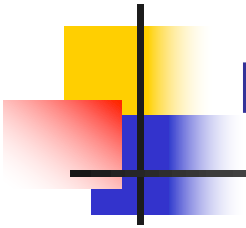
- 1980 CPU: Intel 8080 8-bit processor
  - «int»: one byte
  - «long»: two bytes
  - CPU clock frequency: 1 MHz
  - Memory: 64 kilobytes max, *i.e.*, 16 bit addresses (2 bytes)
  - 128 kB floppy disk
  - Data transmission 1.2 kbit/s == 150 bytes/s

- A look at Moore's law
  - Moore's original prediction 1965
  - Moore's revisted prediction 1975



Moore's Law Summary of Effects

- 1958-1975
  - Doubling of transistors every year
  - Derived effect: doubling of CPU speeds
- 1975-2005
  - Doubling of transistors every two years
  - Derived effect: doubling of CPU speeds – screaching halt at about 3 GHz
- 2005-2020
  - Doubling of transistors every two years
  - Derived effect: doubling of number of CORES *or* doubling of the amount of on-chip cache



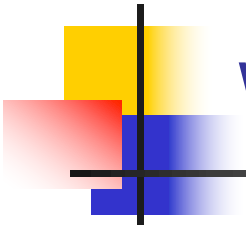
Newer Machines

- 2023: My Macbook M2 Pro w M2 procesor
  - 2.4 up-to 3.2 GHz – 19x cores, 12 GPU cores
  - 16 Gbyte Memory
  - 10 Gbit/s network
  - 1 GB Solid State Disk (SSD)
  - Liquied Retina XDR Display 14" 3024x1964 pixels

Performance 1980 vs 2023

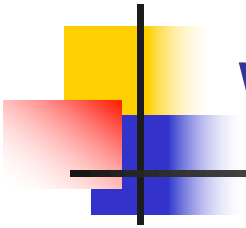
- 2023: Intel 8080 vs M2
  - 1 MHz CPU vs 8 x 3.2 GHz ~ factor 60,000
  - 16 Gbyte Memory vs. 64 kbytes ~ factor 250,000
  - 10 Gbit/s network vs 1 kbit/s ~ factor 10,000,000
  - 128 kB floppy disk vs 1 TB SSD ~ factor 8,000,000
  - 1200 baud serial line vs 10 Gbit/s Ether ~ factor 8,000,000

- ☺



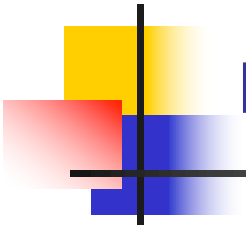
What is ping?

- `Ping`: low-level internet messaging: sends an empty message from one computer to another. The other computer returns it
- sending an empty message from Copenhagen/Oslo to Seattle & back
- In 1988, when the internet was first »opened» in Scandinavia: a ping took a little under 200 ms
- How long does this take 35 years later – in 2023?



What about ping time?

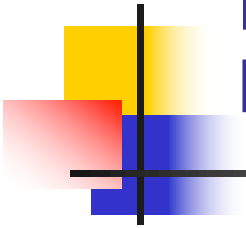
- Ping: sending an empty message from Oslo to Seattle & back
- How long does this take 1988 vs 2022?



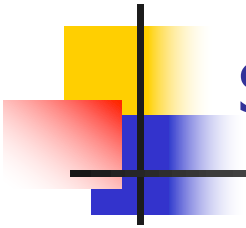
Round-trip ping Time

- 1988: Approximately 180-200 ms
- 2022: How much faster?
  - 1,000,000x faster?
  - 100,000x faster?
  - 10,000x faster?
  - 1,000x faster?
  - 100x faster?
  - 10x faster?
  - Same time?

Latency has ***not*** changed: dominated and limited by the *incredibly* slow speed of light!

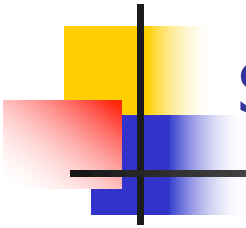


- How fast is the speed of light approximately?
- Great circle route over and back 16,000 km
- Speed of light in fiber 11/15, copper 3/5
- Great Circle round trip time: minimum of about 100 ms
- So ping time will ***NEVER*** go below 100 ms!



Speed of Light

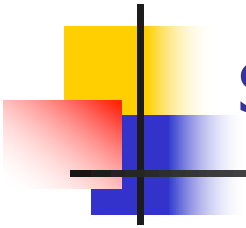
- How fast is the speed of light approximately?
  - 300,000 km/s
- Speed of light
  - in fiber 11/15: about 220,000 km/s
  - in copper 3/5: about 180,000 km/s



Speed of light revisited

- How fast is the speed of light approximately?
  - 300,000 km/s

- How far does light travel in 1 ns?
- How far in copper in 1 ns?
- When a 3 GHz core executes one cycle, how far does light travel?



Speed of light: answers to questions

- How fast is the speed of light?
  - About 300,000 km/s
  - 299,792,458 m/s EXACTLY – by definition
- How far does light travel in 1 ns?
  - About 30 cm – call it a *lightfoot*
- How far in copper in 1 ns?
  - About 18 cm
- When a 3 GHz core executes one cycle, how far does light travel in vacuum and in copper?
  - About 10 cm (vacuum) and 6 cm (copper)

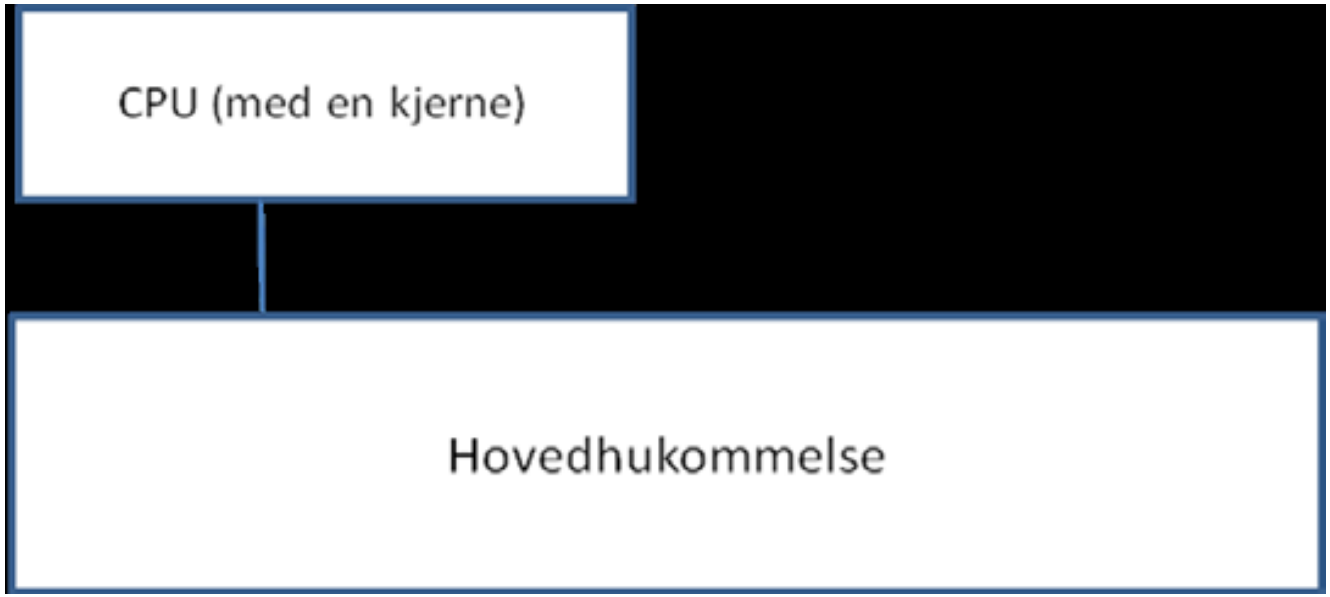
PRAM modellen for parallelle beregninger

- PRAM (Parallel Random Access Memory) antar to ting:
  - Du har uendelig mange kjerner til beregningene
  - Enhver aksess i lageret tar like lag tid,
    - ignorerer f.eks fordelen med cache-hukommelsen
- Da blir mange algoritmer lette å beregne og programmere
- Problemet er at begge forutsetningene er helt gale.
- Det PRAM gjør er å telle antall instruksjoner utført

  Det har vi sett er helt feilaktig (Radix og Matrise-mult)
- Svært mange kurs og lærebøker er basert på PRAM

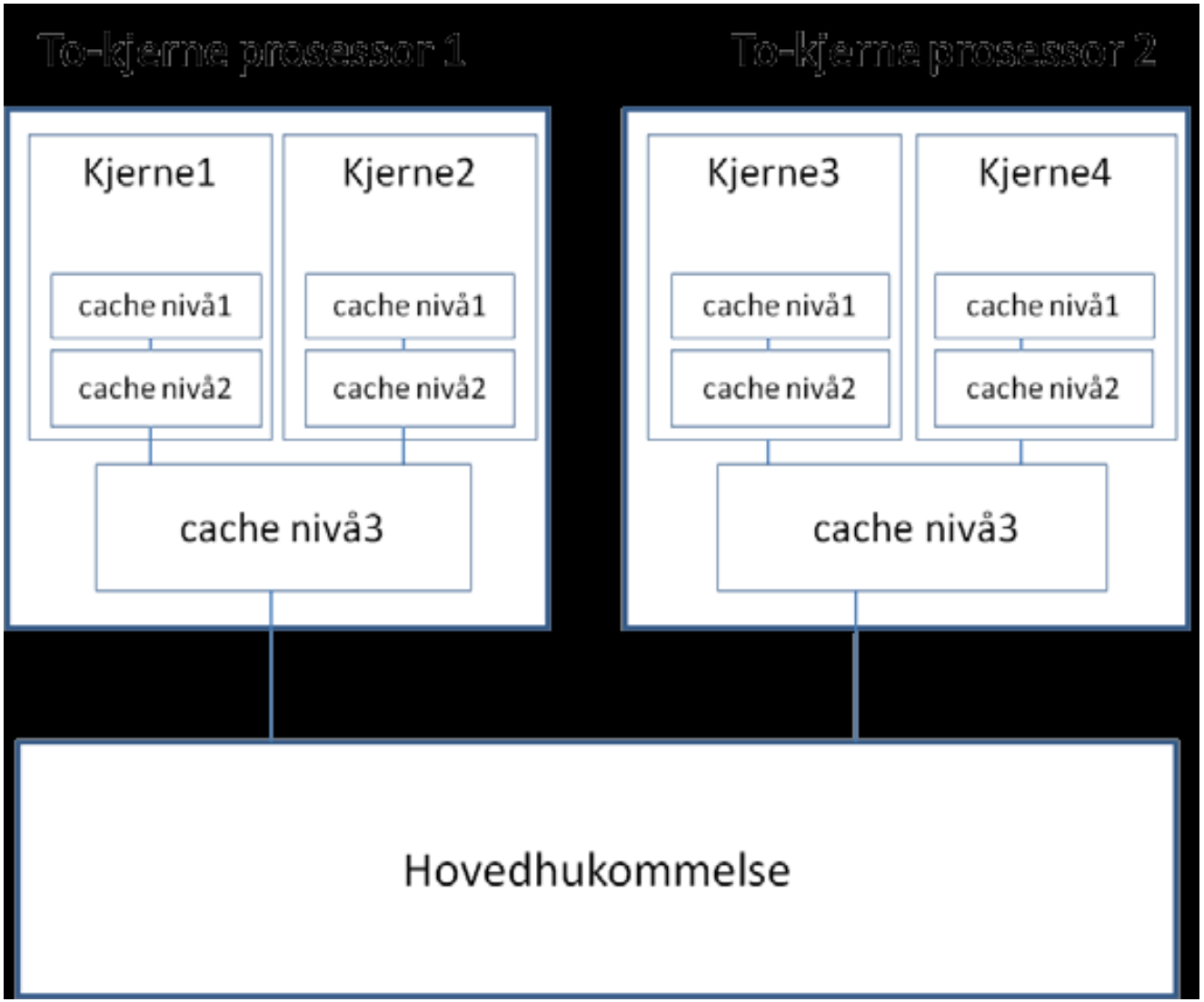
- PRAM vil si at de to sekvensielle algoritmene (med og uten transponering) gikk den utransponerte fortest!
- Dette kurset bruker **ikke** PRAM-modellen!

Maskin 1980 (uten cache)

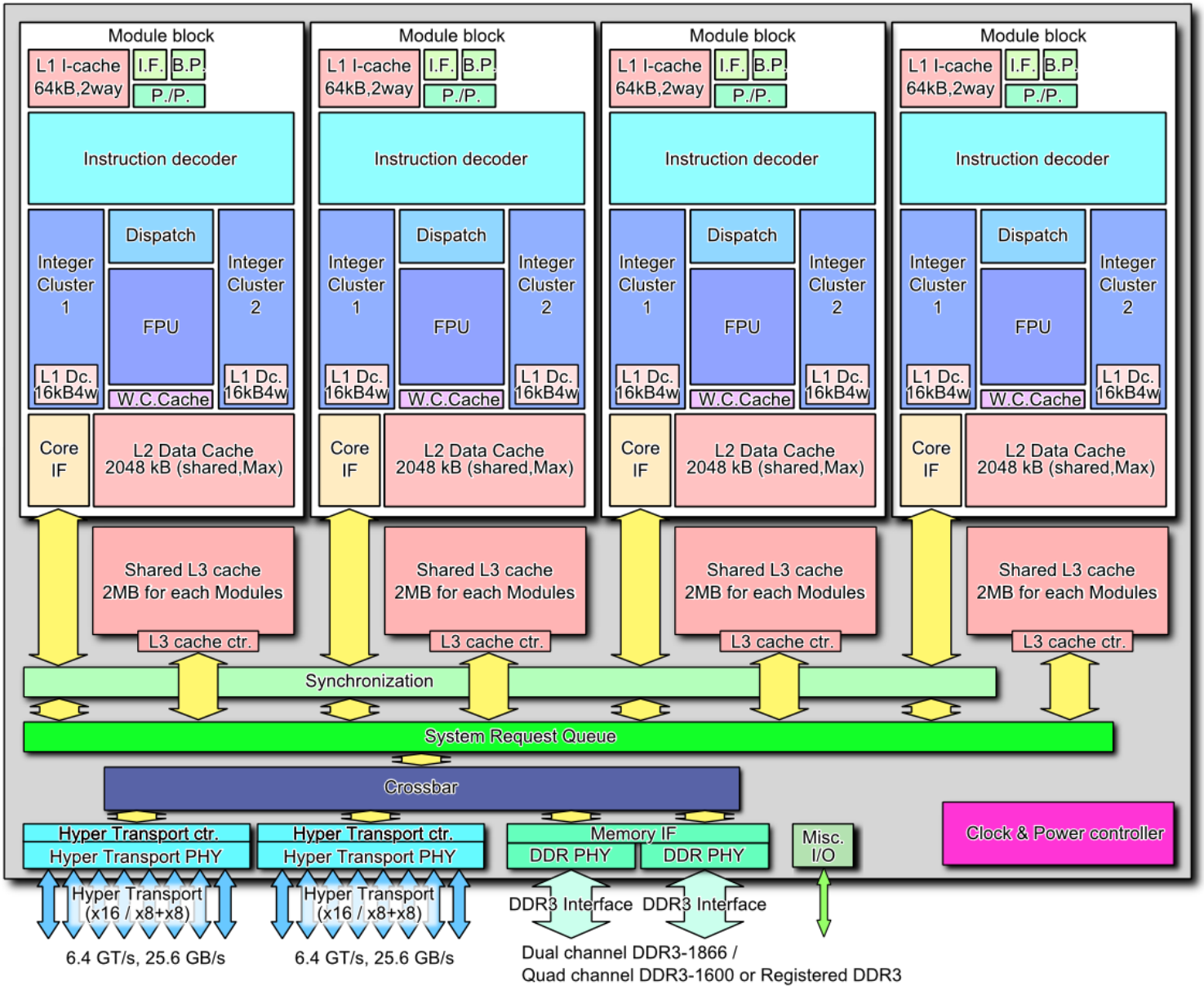


Figur 19.1 Skisse av en datamaskin i ca. 1980 hvor det bare var én beregningsenhet, en CPU, som leste sine instruksjoner og både skrev og leste data (variable) direkte i hovedhukommelsen. Intel 8080: 1 MHz CPU

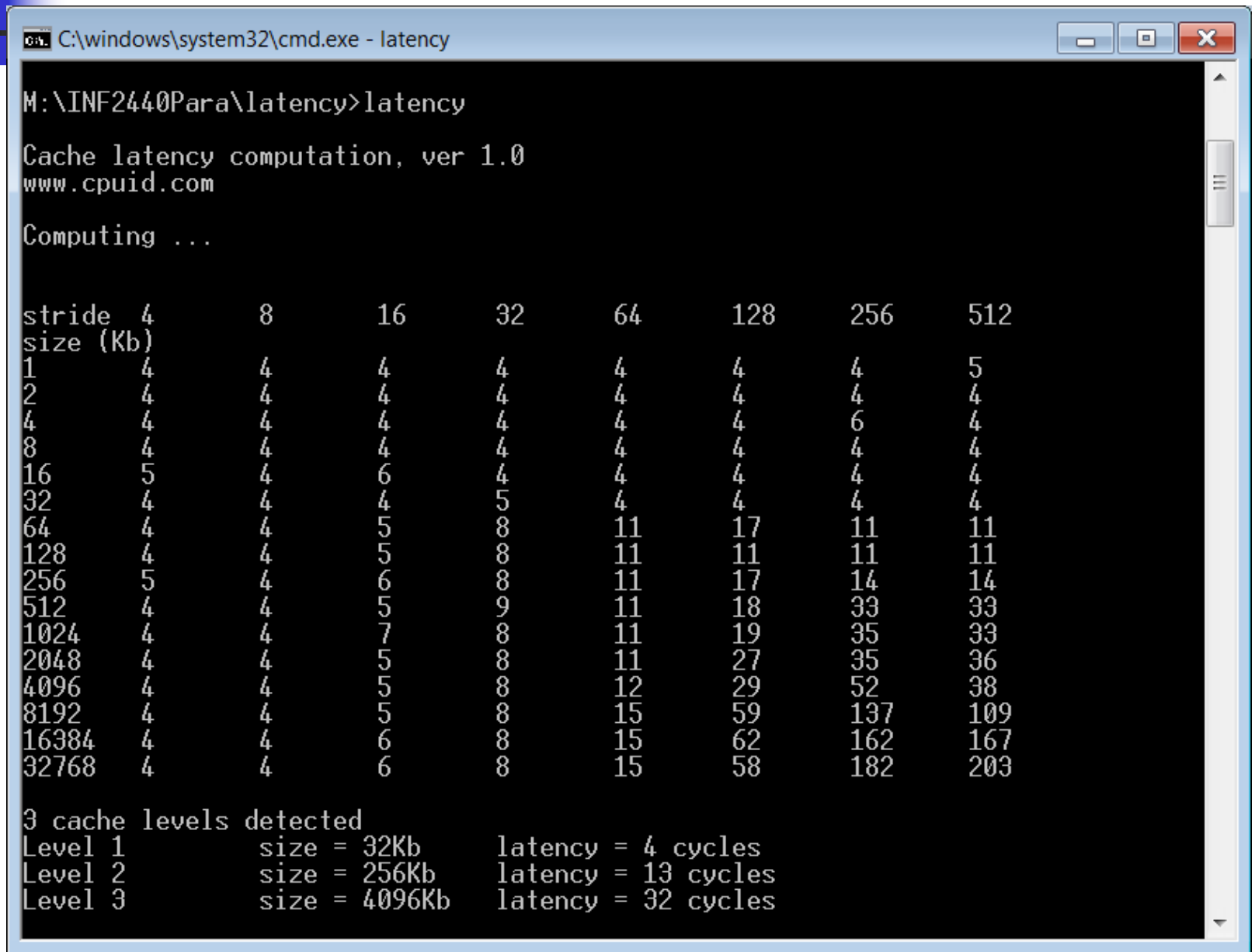
Maskin ca. 2010 med to dobbeltkjerne CPU-er



Hukommelses-systemet i en 4 kjerne CPU – mange lag og flere ulike beregningsmoduler i hver kjerne.:



Test av forsinkelse i data-cachene og hovedhukommelsen - latency.exe (fra CPUZ)

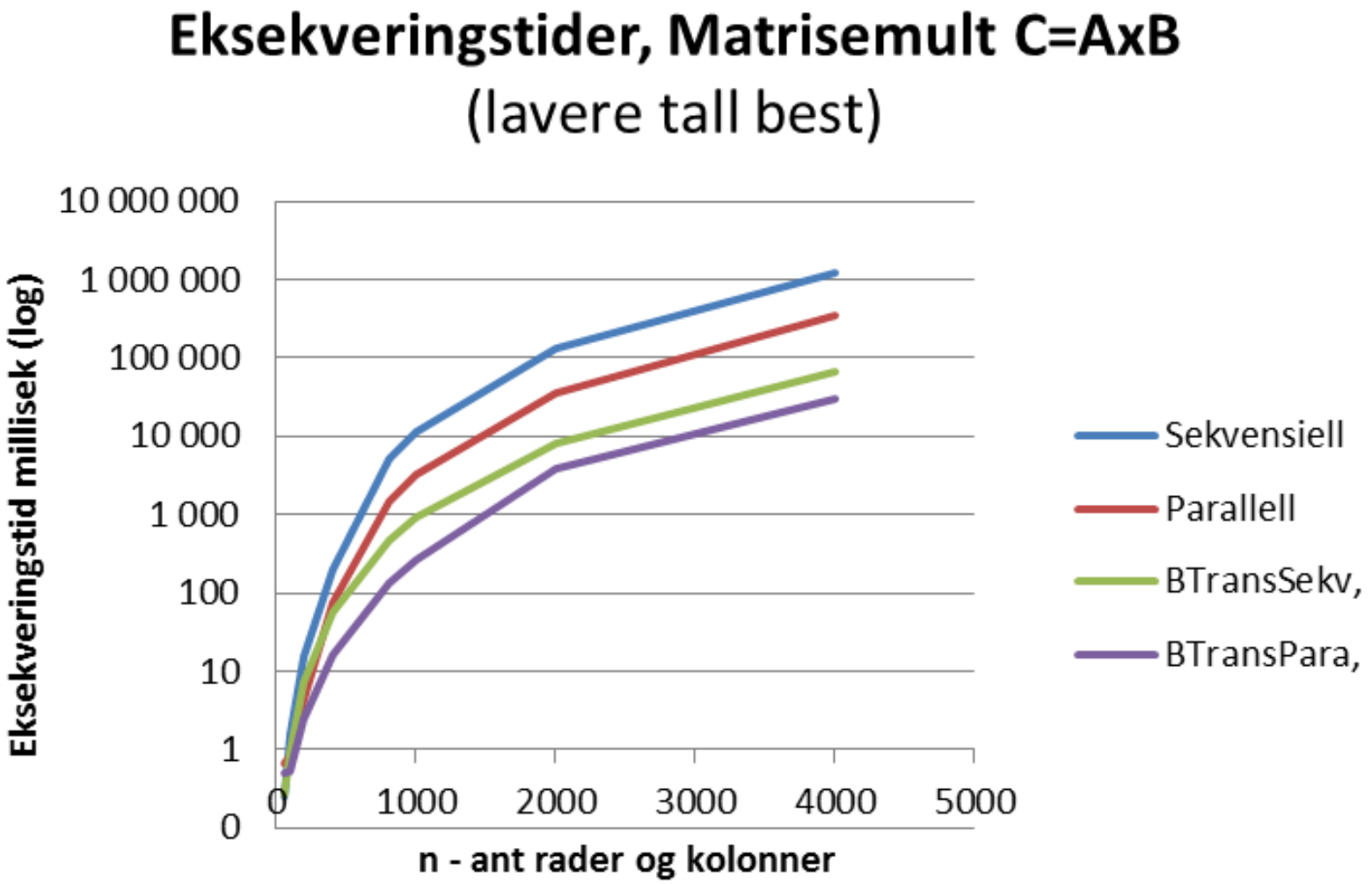


Oppsummering – ideen om at vi har *uniform* aksesstid i hukommelsen er helt galt

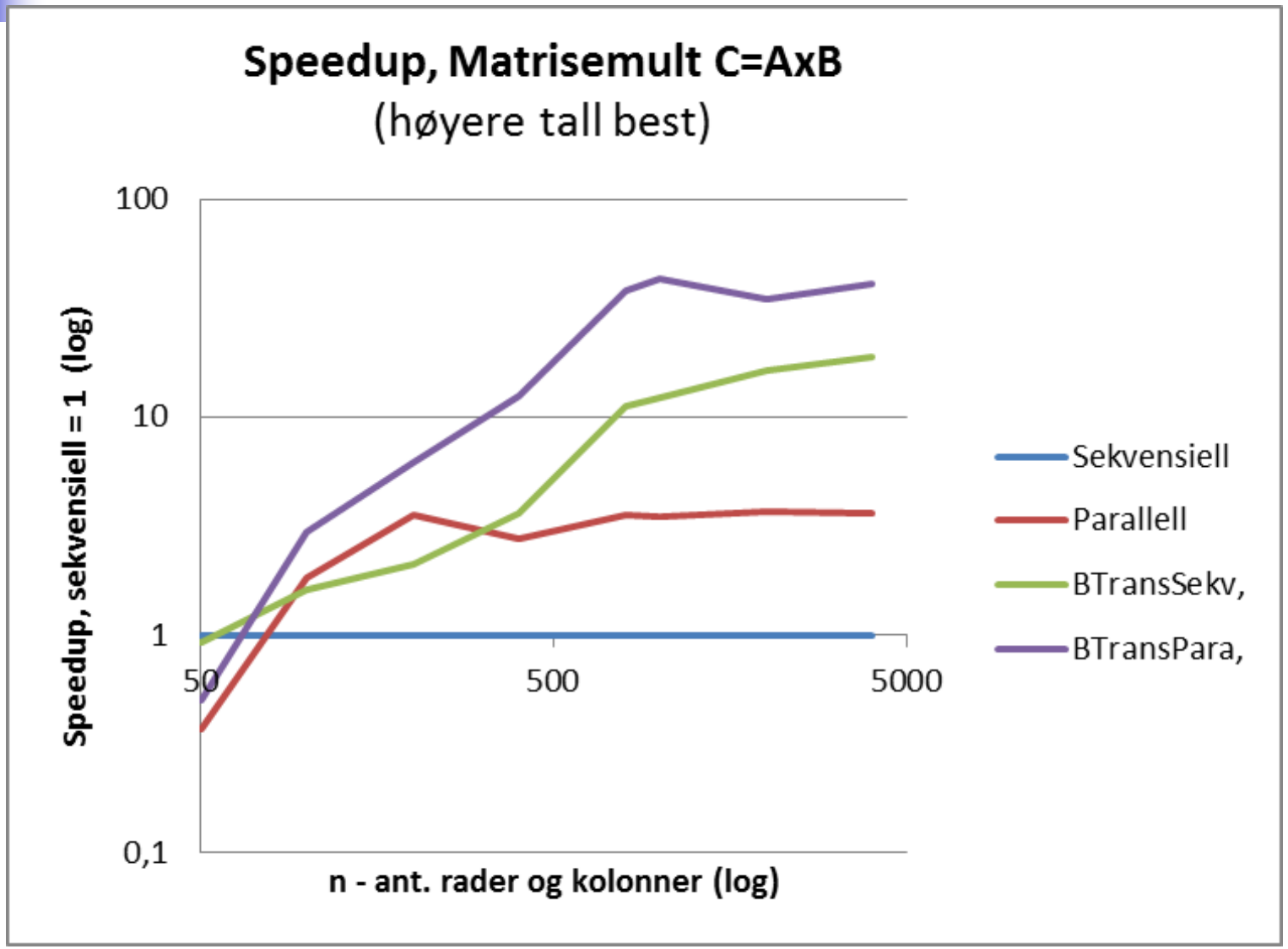
- Hukommelses-systemet i en multicore CPU ,Intel Core i5-459 3.3 GHz, – mange lag (typisk aksesstid i instruksjonssykler):

  1. Registre i kjernen (1) – 8/32 registre
  2. L1 cache (3-4) – 32 Kb
  3. L2 cache (13) – 256 kb
  4. L3 cache (32) – 8Mb
  5. Hovedhukommelsen (virtuell hukommelse) (ca. 200) – 8-64 GB
  6. Disken (15 000 000 roterende) = 5 ms – 1000 GB – 1-5 TB
     FlashDisk (ca 2 000 000 les, ca. 10 000 000 skriv) = ca. 1 ms

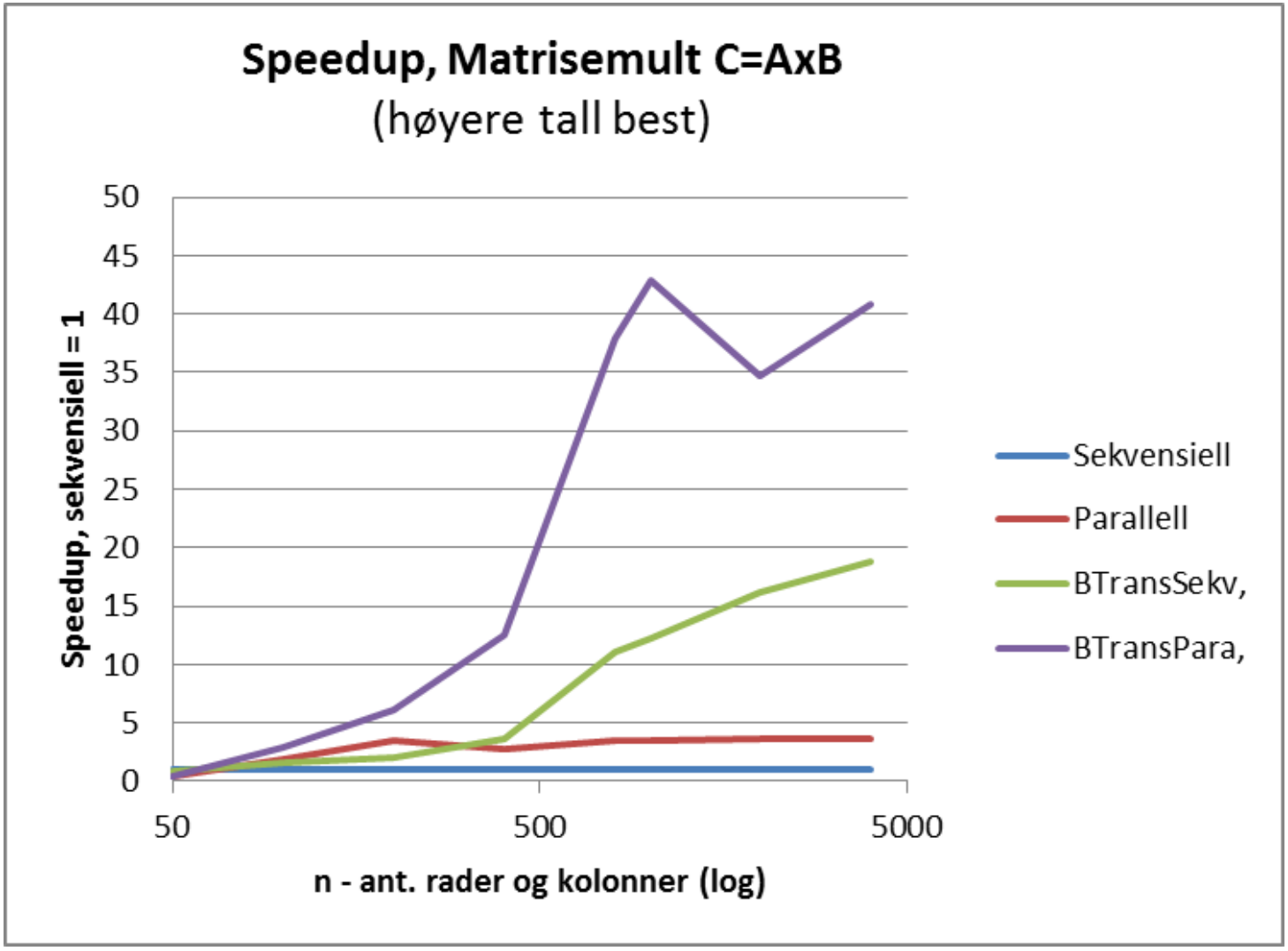
Kjøretider – i millisek. (y-aksen logaritmisk)



Kjøretidsresultater – Speedup , y-aksen logaritmisk

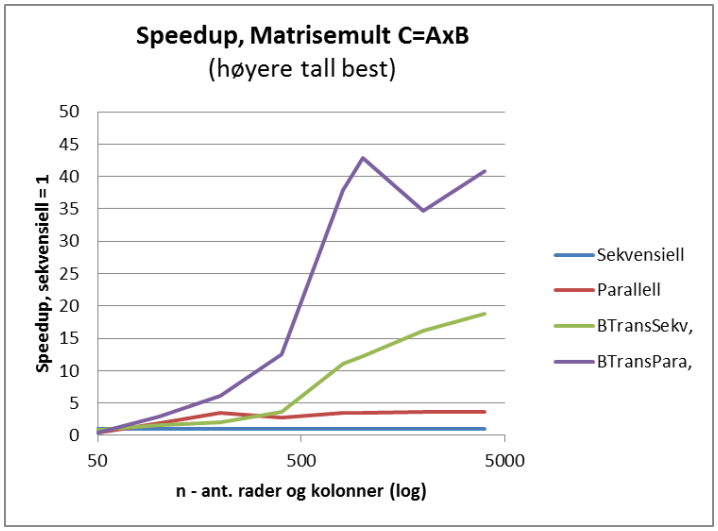


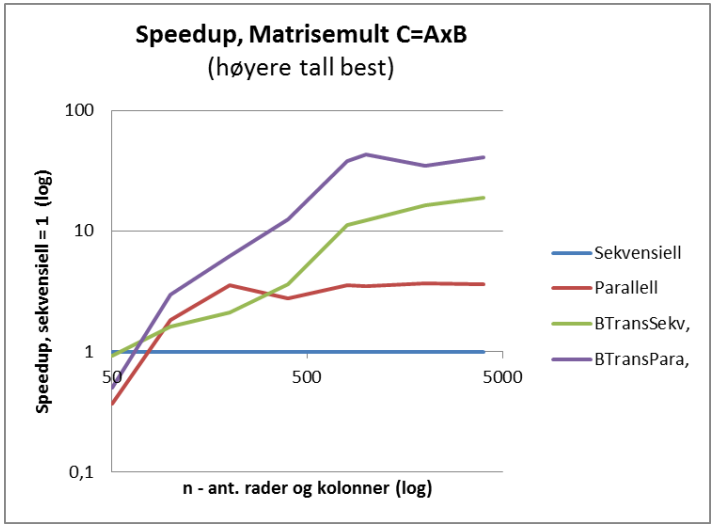
Speedup – med lineær y-akse

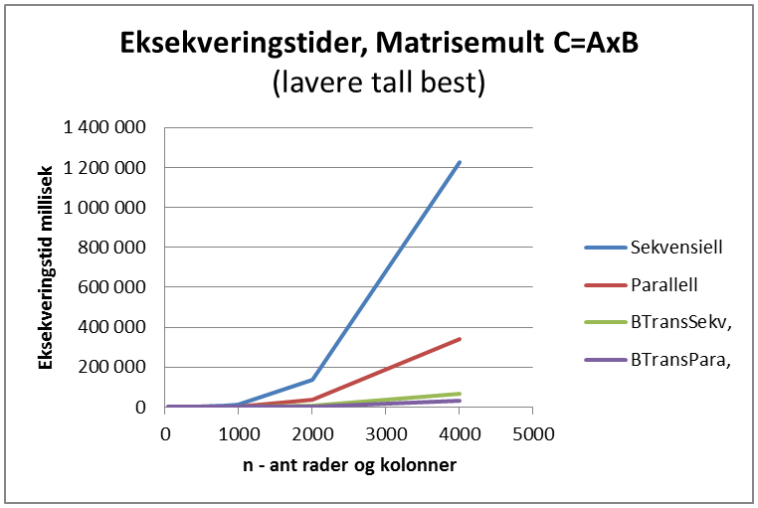


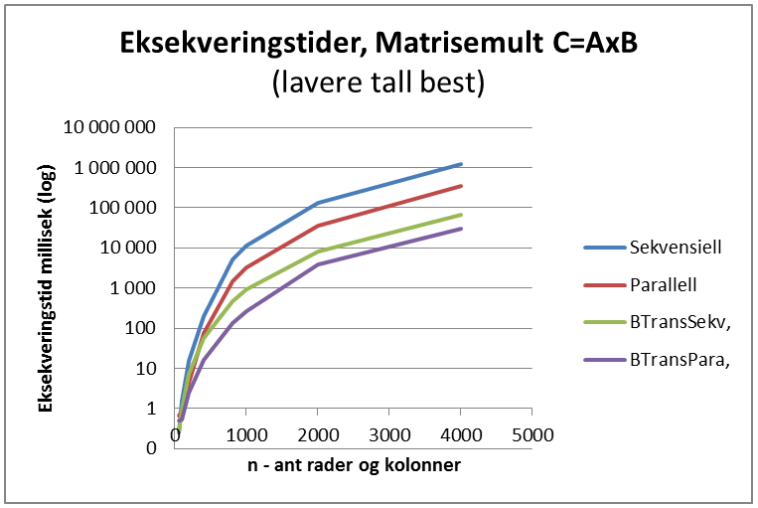
Hvordan framstille ytelse I

- Disse fire kurvene fremviser samme tall! Hvordan ?

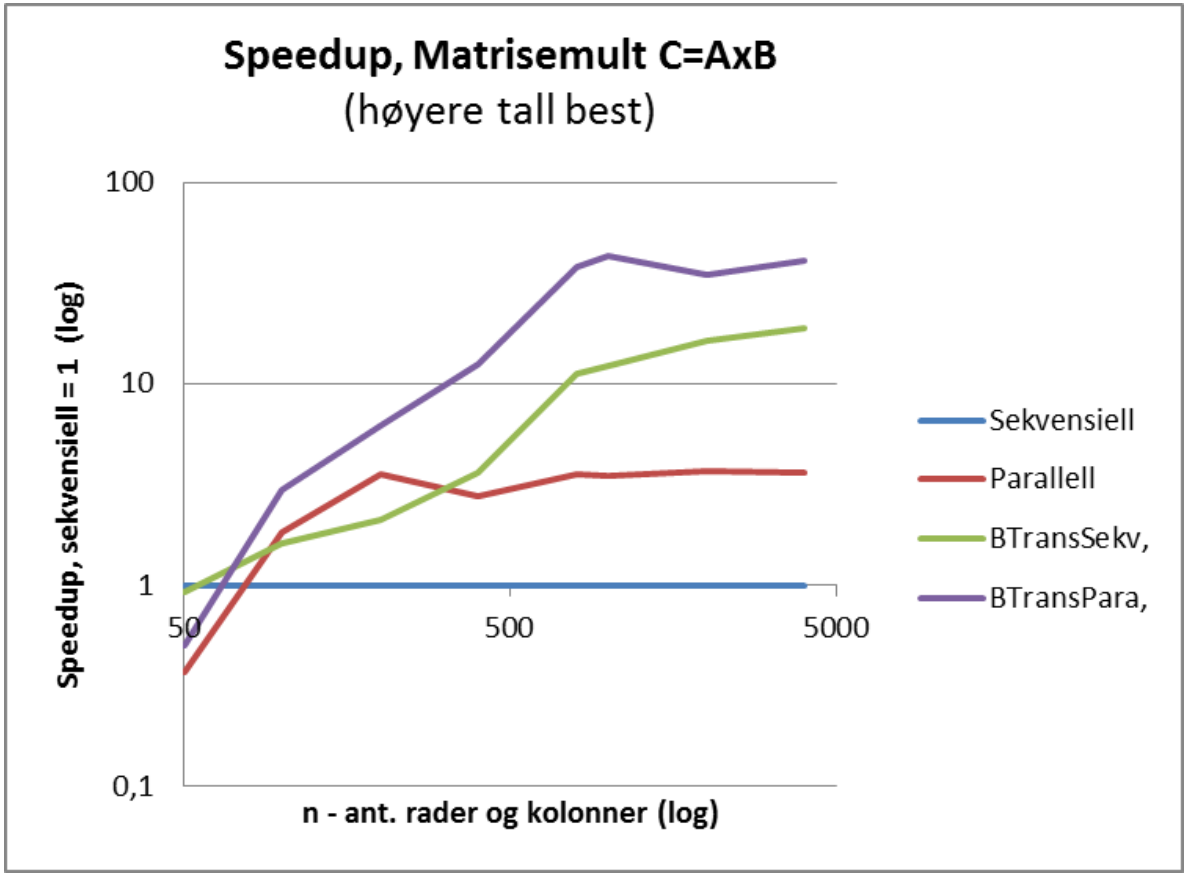








Både logaritmisk x- og y-akse



Fordel med log-akse er at den viser fram nøyaktigere små verdier, men vanskelig å lese nøyaktig mellom to merker på aksene.