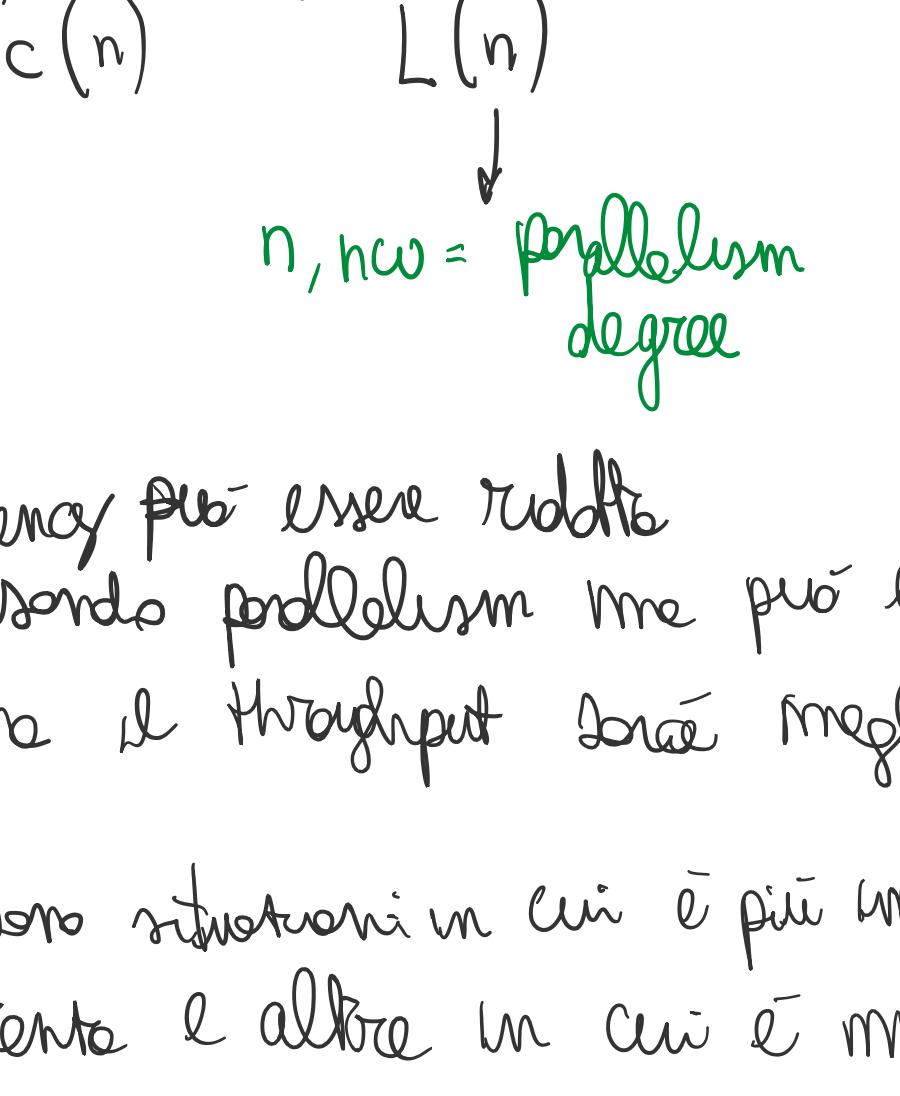


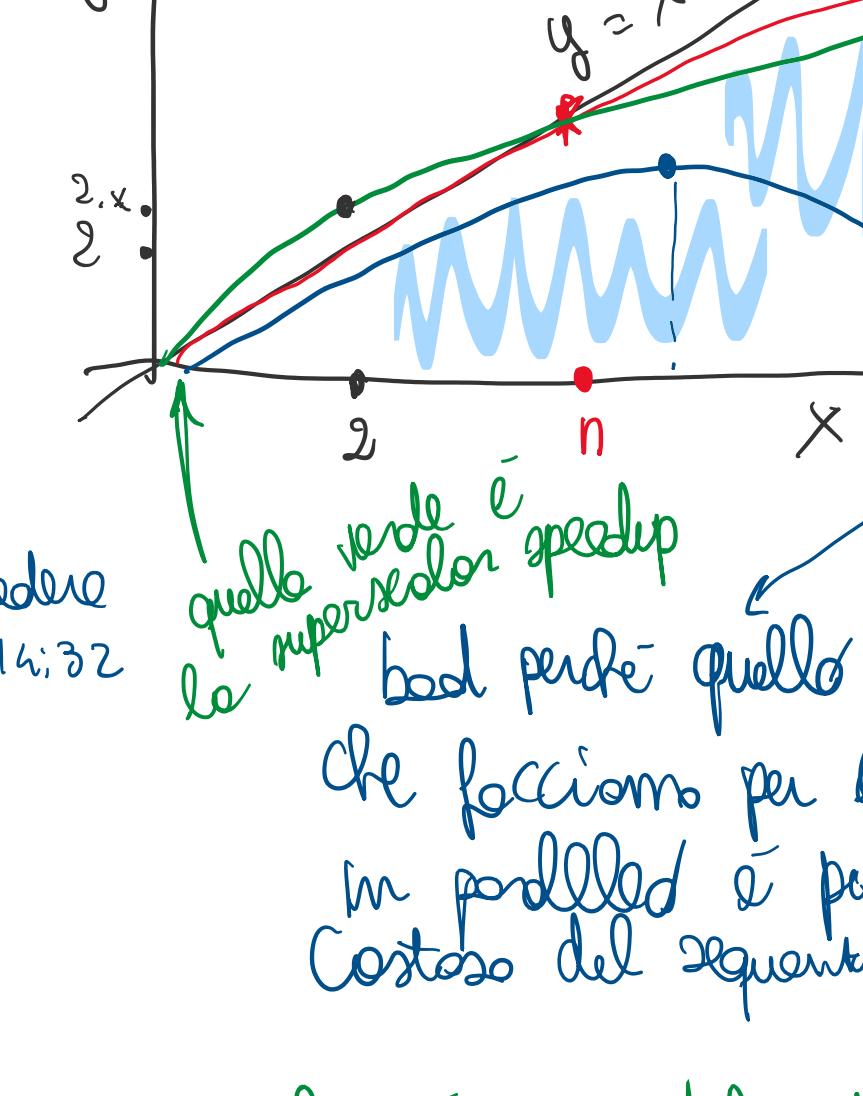
Performance Measure:

Latency

È più legato al tempo che serve per eseguire un certo task.



Se ho n task da eseguire poniamo di completion task T_c



Se però si parla di: $T_c(n) = L(n)$

$$n, \text{new} = \text{parallelism degree}$$

Latency può essere ridotta usando parallelismo ma può essere migliore nel sequenziale ma il throughput sarà migliore nel parallelo.

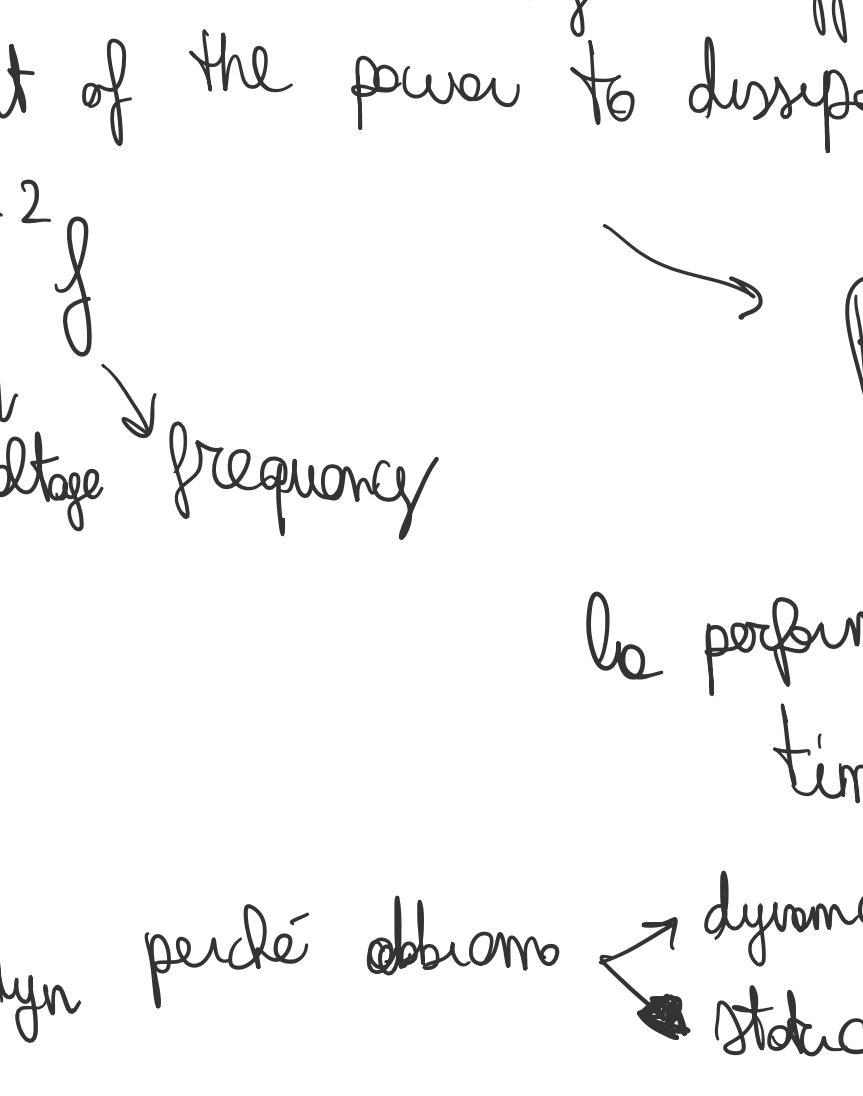
Ci sono situazioni in cui è più importante abbassare la latenza e altre in cui è meglio abbassare il throughput.

Speedup

Indicata con S:

$$S(n) = \frac{\text{best sequential time}}{\text{parallel time of execution}} = \frac{T_{\text{seq}}}{T_p(n)}$$

Speedup ha dei limiti:



Se super scalar è impossibile dal punto di vista teorico

$t = 8 \text{ s}$ è il best time per il sequenziale

T_{req} :

Per avere in uso dei punti verdi de stanno sopra all'asse se uso solamente 2 worker.

w_1
 w_2

che fanno per eseguire in parallelo è più costoso del sequenziale.

I didn't respect the assumption de il processore de uso per lo sequenziale è lo stesso della parallelo.

The amount of improvement with superscalar is limited.

Scalability

$$Sc(n) = \frac{T(1)}{T(n)} \rightarrow \text{time of parallel con parallelism=1}$$

std::vector::<pointer<int, int>>

Inserito nel vettore 1 item

Sometimes the things to do the library thread safe introduce overhead.

ci sono delle situazioni in cui $T(1) > T_{\text{req}}$

the overhead of creating threads.

100 μsec Comp 1 sec

10 1 sec

Supponiamo che l'esecuzione sequenziale ha 64K di dati nella cache.

$64K$

w_1
 w_2

I didn't respect the assumption de il processore de uso per lo sequenziale è lo stesso della parallelo.

All the rest can be made in parallel.

total amount of work to compute

$$\text{Work} = t_{\text{split}} + t_{w_1} + t_{w_2} + \dots + t_{w_n} + t_{\text{merge}}$$

the span is the longest critical path

all the rest can be made in parallel.

Se il sequenziale time è T_{seq}

the span is the best time to achieve as T_{par} .

$Sp \approx \text{best time}$

$Sp \approx \frac{\text{Work}}{\text{Span}}$

Consideriamo il grafo:

Il critico path da INIT a END è 1-3-5-8

the total work = somma di tutti i 10t

Span = 80t

Better speedup I can achieve is $Sp = \frac{10}{80}$

$t_1 = 10t$

$t_2 = 10t$

$t_3 = 10t$

$t_4 = 10t$

$t_5 = 50t$

$t_6 = 5t$

$t_7 = 10t$

$t_8 = 10t$