




# Hyperthreading and "Almost Amdahl"



**Oregon State University**  
Mike Bailey  
mjb@cs.oregonstate.edu



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

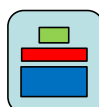


hyperthreadingandalmostamdahl.pptx

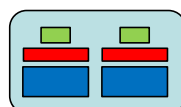
mp - March 15, 2020

1

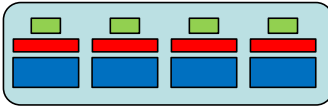
## Each of the Multiple Cores keeps its own State



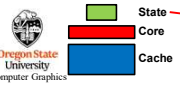
1 core, 1 state




2 cores, 2 states



4 cores, 4 states




- Registers
- Program Counter
- Stack Pointer




mp - March 15, 2020

2


## So, if that's what Multicore is about, what is *Hyperthreading*?



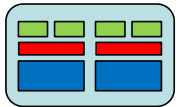
1 core, 1 state



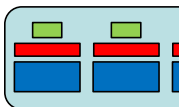
1 core, 2 states, with Hyperthreading



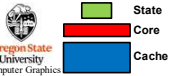
2 cores, 2 states





2 cores, 4 states, with Hyperthreading



4 cores, 4 states







mp - March 15, 2020

3

## What is Hyperthreading and what can it Do?

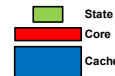
**Hyperthreading** is when a CPU chip has more states than cores.


In this case, if one thread of execution blocks (waiting for a memory fetch, for instance), then the other thread can resume execution with its state.


If we let  $H$  be the fraction of a CPU's capacity that one hyperthread can keep busy, then the remaining unused capacity is  $(1-H)$ . If another hyperthread can keep  $H\%$  of that capacity busy, then that leaves  $(1-H)(1-H)$  remaining unused capacity and so on.

If we have  $n$  hyperthreads, then the final remaining unused capacity is  $(1-H)^n$ . The capacity actually in use would then be  $1-(1-H)^n$ . If one thread can only keep the CPU  $H\%$  busy, then the speed-up is potentially:

$$SU = \frac{1-(1-H)^n}{H}$$







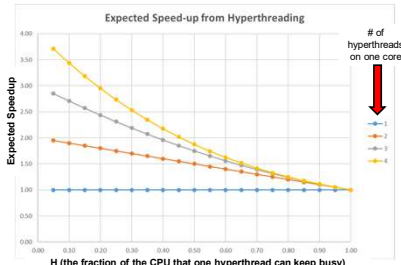
mp - March 15, 2020

4

## What is Hyperthreading and what can it Do?

If we have  $n$  hyperthreads, then the final remaining unused capacity is  $(1-H)^n$ . The capacity actually in use would then be  $1-(1-H)^n$ . If one thread can only keep the CPU  $H\%$  busy, then the speed-up is potentially:


$$SU = \frac{1-(1-H)^n}{H}$$



Expected Speed-up from Hyperthreading

# of hyperthreads on one core

H (the fraction of the CPU that one hyperthread can keep busy)

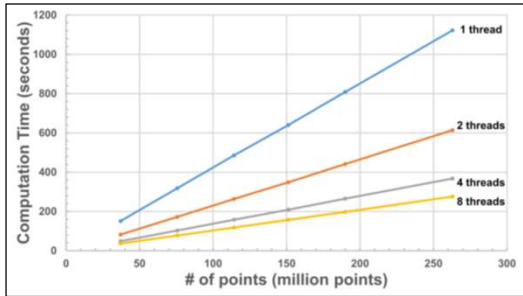


William Leslie


mp - March 15, 2020

5


## A Lidar Application: Four Cores with Two Hyperthreads per Core



Source: Erzhuo Che



Note that this is upside-down from our usual convention. Sorry. I got this from someone else.



mp - March 15, 2020

6