CS305: Computer Architecture

Empirical Evaluation

https://www.cse.iitb.ac.in/~biswa/courses/CS305/main.html

https://www.cse.iitb.ac.in/~biswa/

Empirical Evaluation
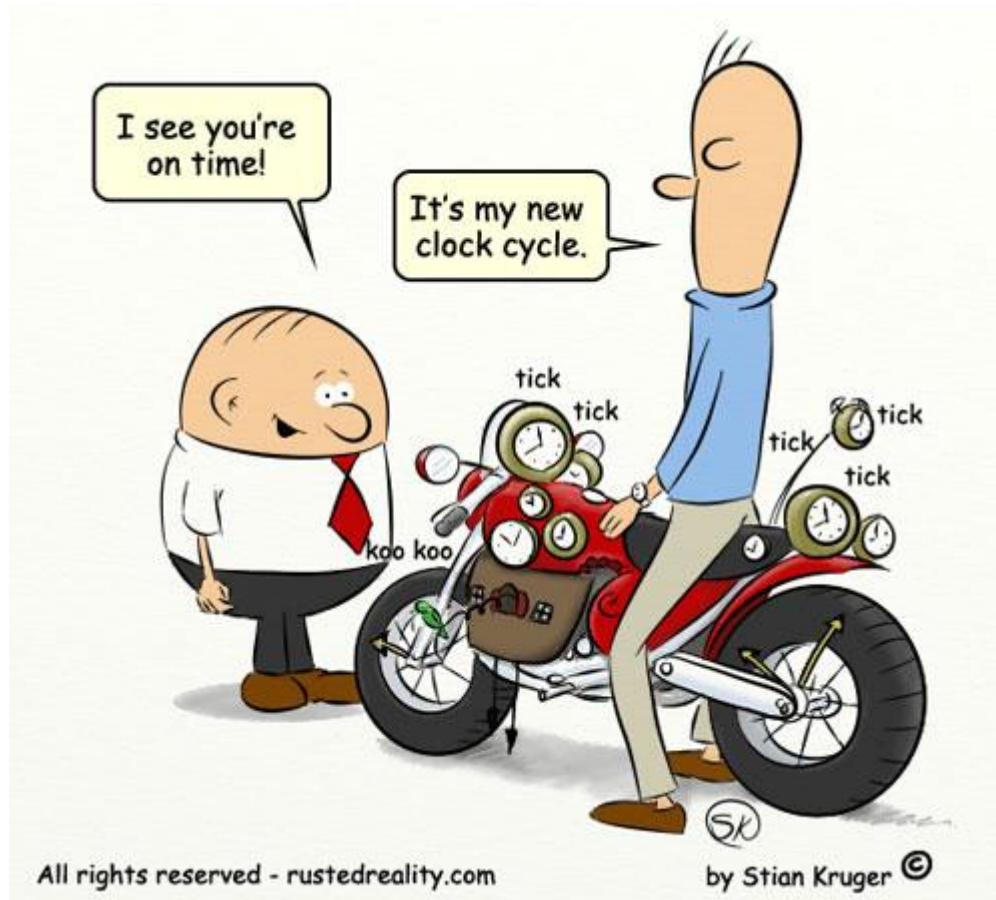
Single core

Multi core

Evaluation

Benchmarks

Metrics

Simulators

Latency and bandwidth

Performance

- _Latency_ (execution/response time): time to finish one task. It is additive (Performance = 1/latency)

- _Throughput_ (bandwidth): number of tasks/unit time. It is not additive

- _Latency lags bandwidth, Bandwidth hurts latency, Read: https://cacm.acm.org/magazines/2004/10/6401-latency-lags-bandwith/fulltext_

Computer Architecture

Latency vs bandwidth

Latency vs Bandwidth, How they affect each other?

Latency helps bandwidth but not vice versa.

DRAM latency ↓ More # Accesses ~DRAM Bandwidth ↑

Bandwidth usually hurts latency

Queues - Bandwidth ↑ Increases latency ↓

For the curious ones

Bandwidth problems can be cured with money.
Latency problems are harder because the speed of light is fixed – you can't bribe God

https://www.youtube.com/watch?list=PL2LuePcZTMh_MzNHqZWNdvWdAnAThHCKK&v=lfqgpuH10uc&feature=emb_logo
https://www.youtube.com/watch?v=GNK-67JUH7M&list=PL2LuePcZTMh_MzNHqZWNdvWdAnAThHCKK&index=2
https://www.youtube.com/watch?v=5CxpoGwCxKU&list=PL2LuePcZTMh_MzNHqZWNdvWdAnAThHCKK&index=3

Some of the major bottlenecks: latency and bandwidth bottlenecks

Amdahl's Law (common case fast)

$$\text{Speedup}_{overall} = \frac{\text{Execution Time}_{old}}{\text{Execution Time}_{new}}$$

$$= \frac{1}{(1 - \text{Fraction}_{enhanced}) + \dfrac{\text{Fraction}_{enhanced}}{\text{Speedup}_{enhanced}}}$$

Computer Architecture

Amdahl's Law

Which one will provide better overall speedup?
A. Small speedup on the large fraction of execution time.
B. Large speedup on the small fraction of execution time.
C. Does not matter.

Depends on the difference between small and large. Mostly it is A.

Thanks