



# Hardware and system for Big Data

Interesting papers about hardware architecture and big data

# Review of articles about computer architecture

# A New Golden Age for Computer Architecture

Article written by J.L.Hennesy et DA.Patterson in ACM communication (2019)

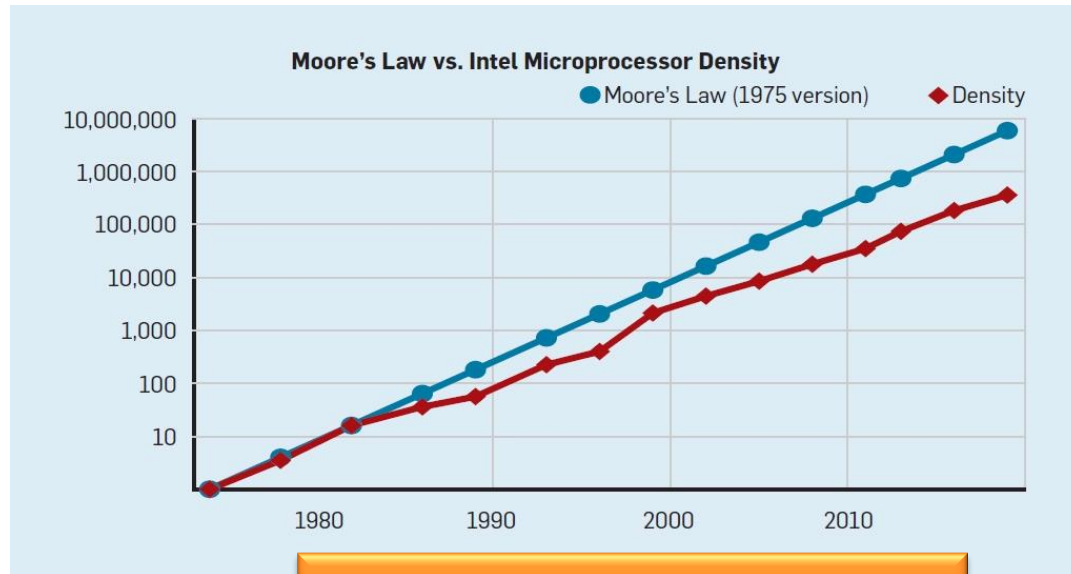
<https://cacm.acm.org/magazines/2019/2/234352-a-new-golden-age-for-computer-architecture/fulltext>



Hardware and system  
for big data

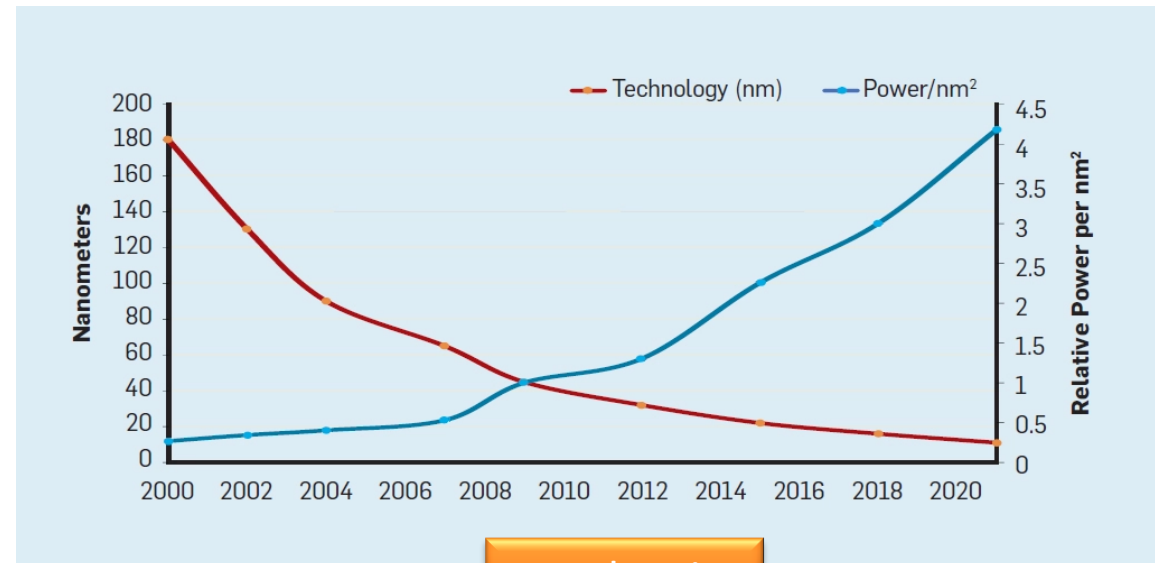
02/10/2019

## Moore's law



The end of sequential execution

## Dennard scaling

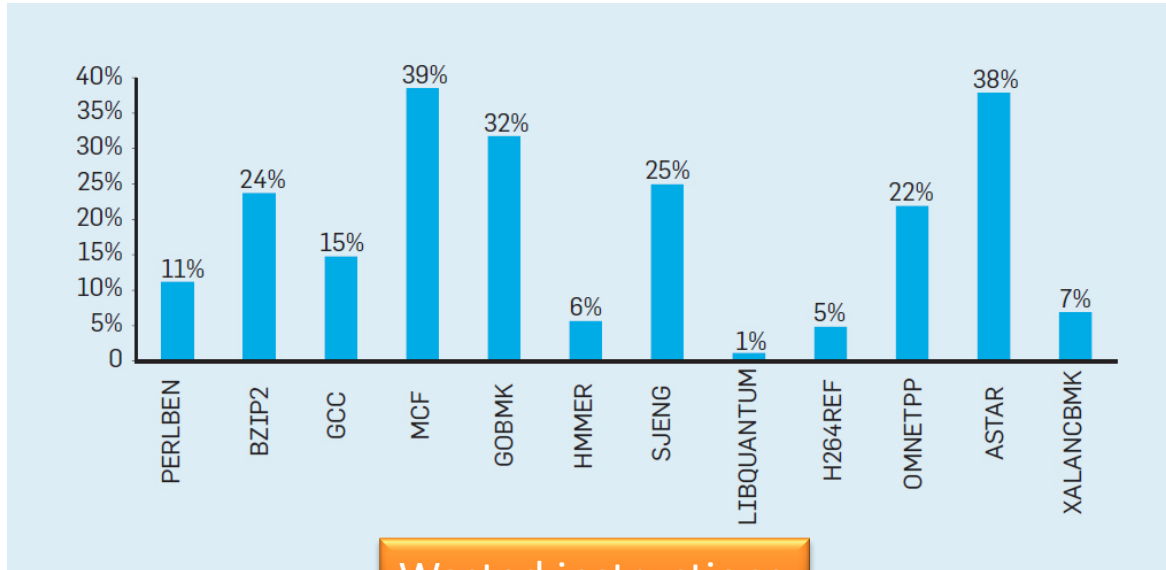


Heat barrier

# A New Golden Age for Computer Architecture (contd)



Processor level parallelism = speculative execution



Wasted instructions

Processor security breach

Side channel attack + memory mapping

- Meltdown
- Spectre
- ...

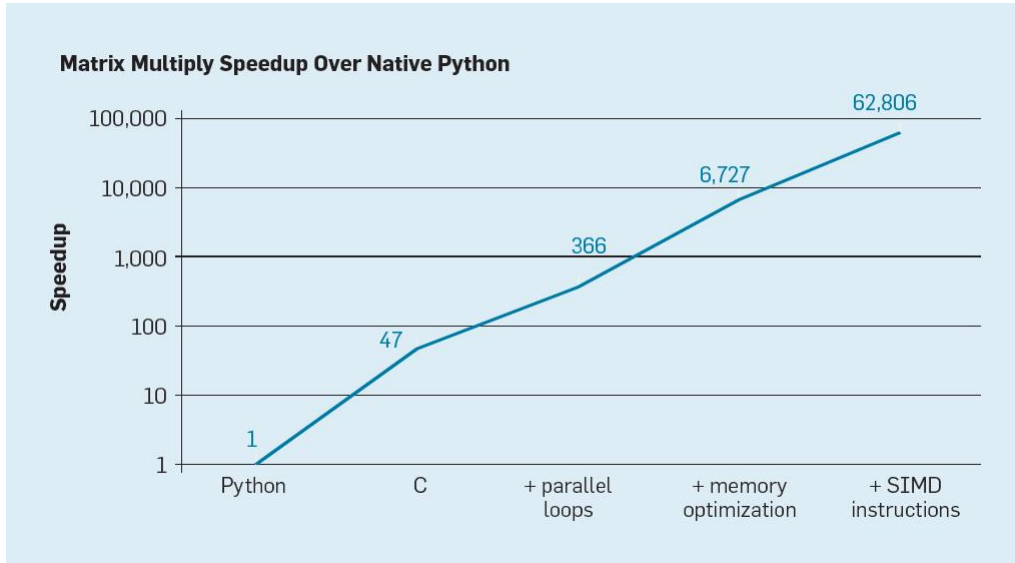
Hardware and system for big data

02/10/2019

# A New Golden Age for Computer Architecture (contd)



## Code optimization



## Domain specific architecture

- GPU
- TPU
- FPGA
- Edge (ARM) computing
- ...

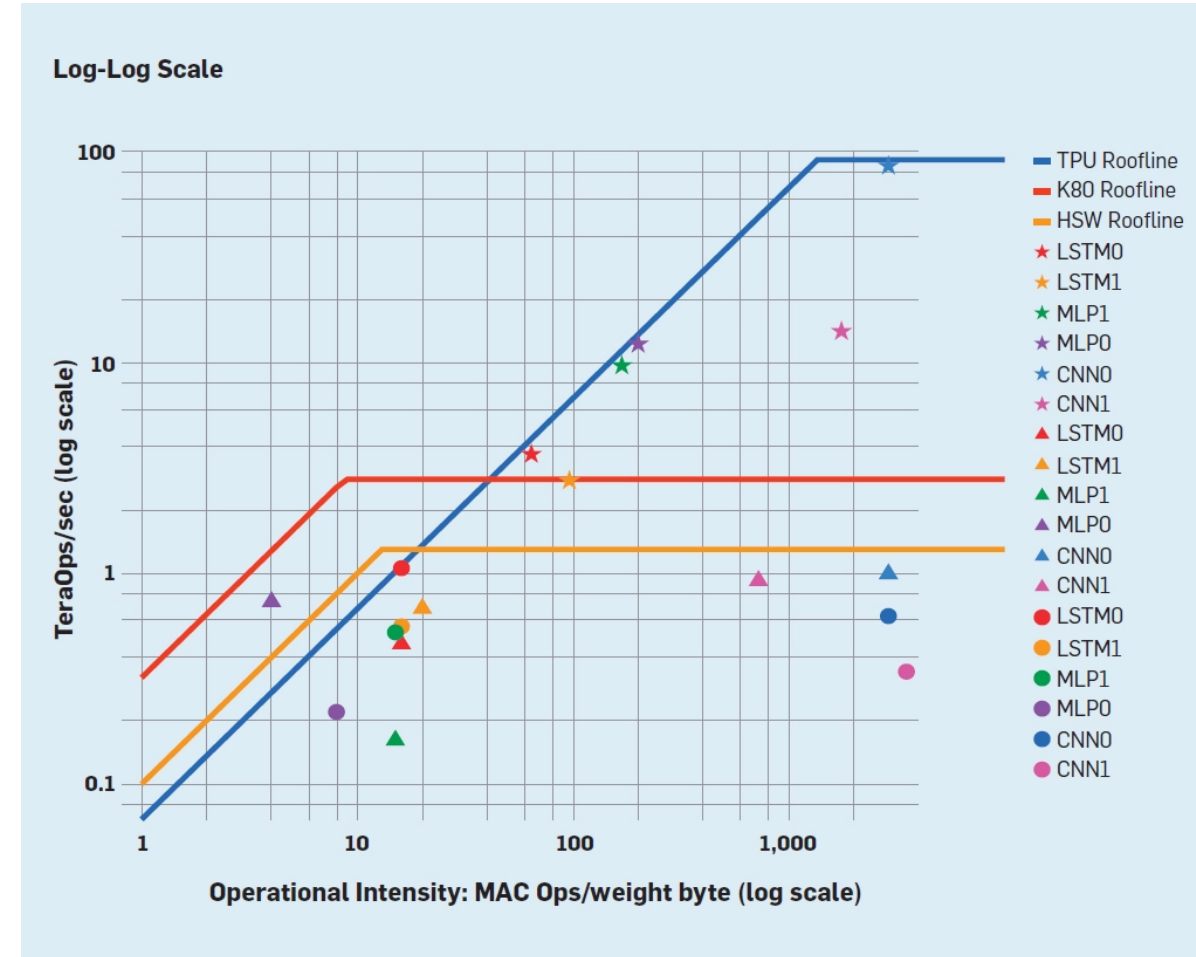
Hardware and system  
for big data

02/10/2019

# A Domain-Specific Architecture for Deep Neural Networks



- ACM communication, September 2018
- Article written by Google engineers to compare TPU and GPU
- Roofline model
- Stars are for the TPU, triangles for the K80, and circles for Haswell
- “Tensor processing units improve performance per watt of neural networks in Google datacenters by roughly 50×”



# Time is an illusion

## Lunchtime doubly so.

- ACMQueue January 2016
- Leslie Lamport's "Time, Clocks, and the Ordering of Events in a Distributed System" (1978),<sup>2</sup> and only a few more have come to appreciate the problems they face once they move into the world of distributed systems.
- The relative nature of time
- Synchronization vs syntonization
- Synchronization = exactly the same moment (whatever the way to measure it)
- Syntonization = exactly the same time tick (~frequency)

## C is not a low level language

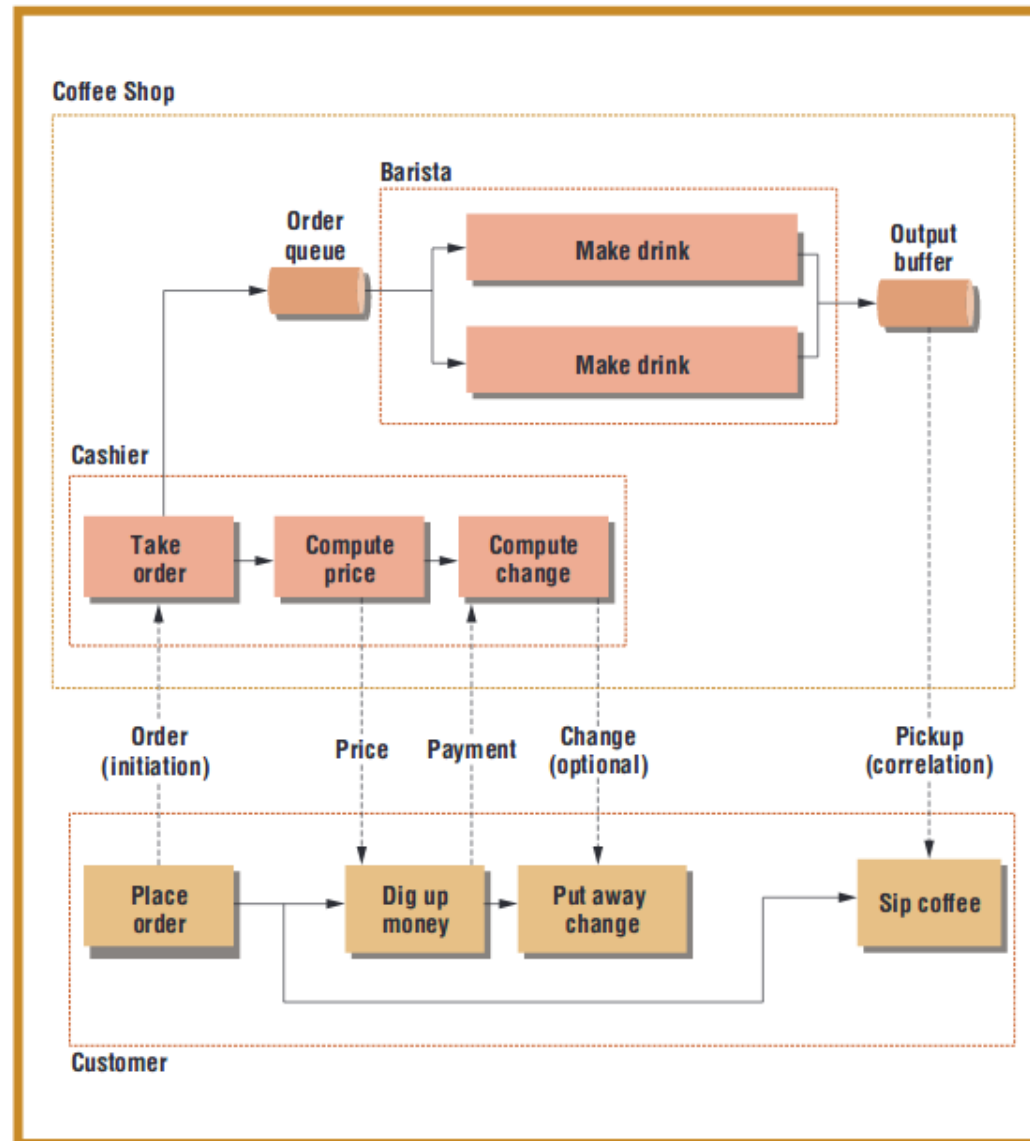
- ACM communication July 2018, D.Chisnall
- C n'est pas un langage pour le parallélisme
- Les processeurs utilisent tous une forme de parallelisme
- L'architecture autour de C a conduit à maintenir une vision système type PDP11 qui ne correspond plus à la réalité des processeurs modernes





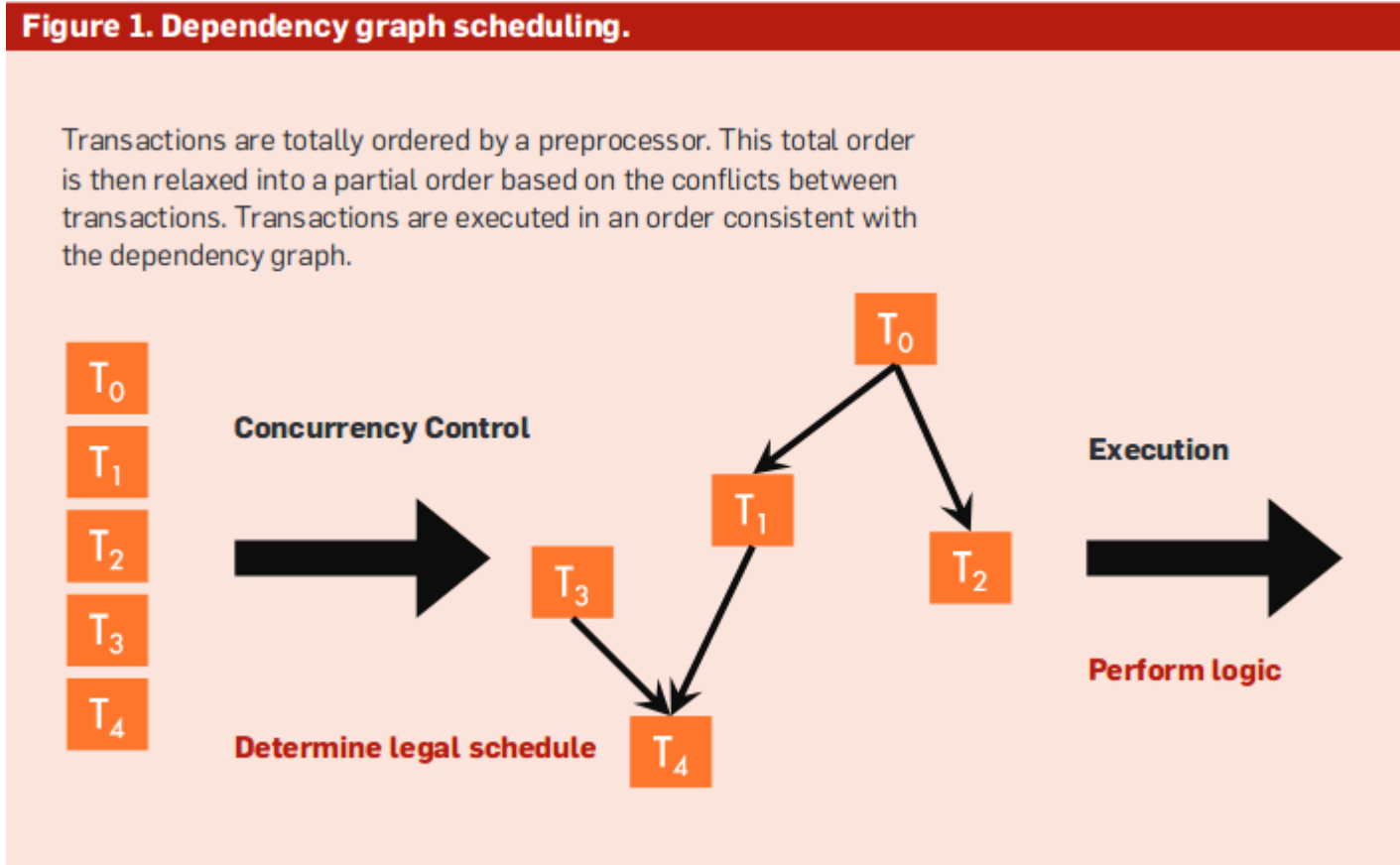
# Your coffee shop doesn't use two-phase commit

- IEEE Software March/April 20005



# An Overview of Deterministic Database Systems

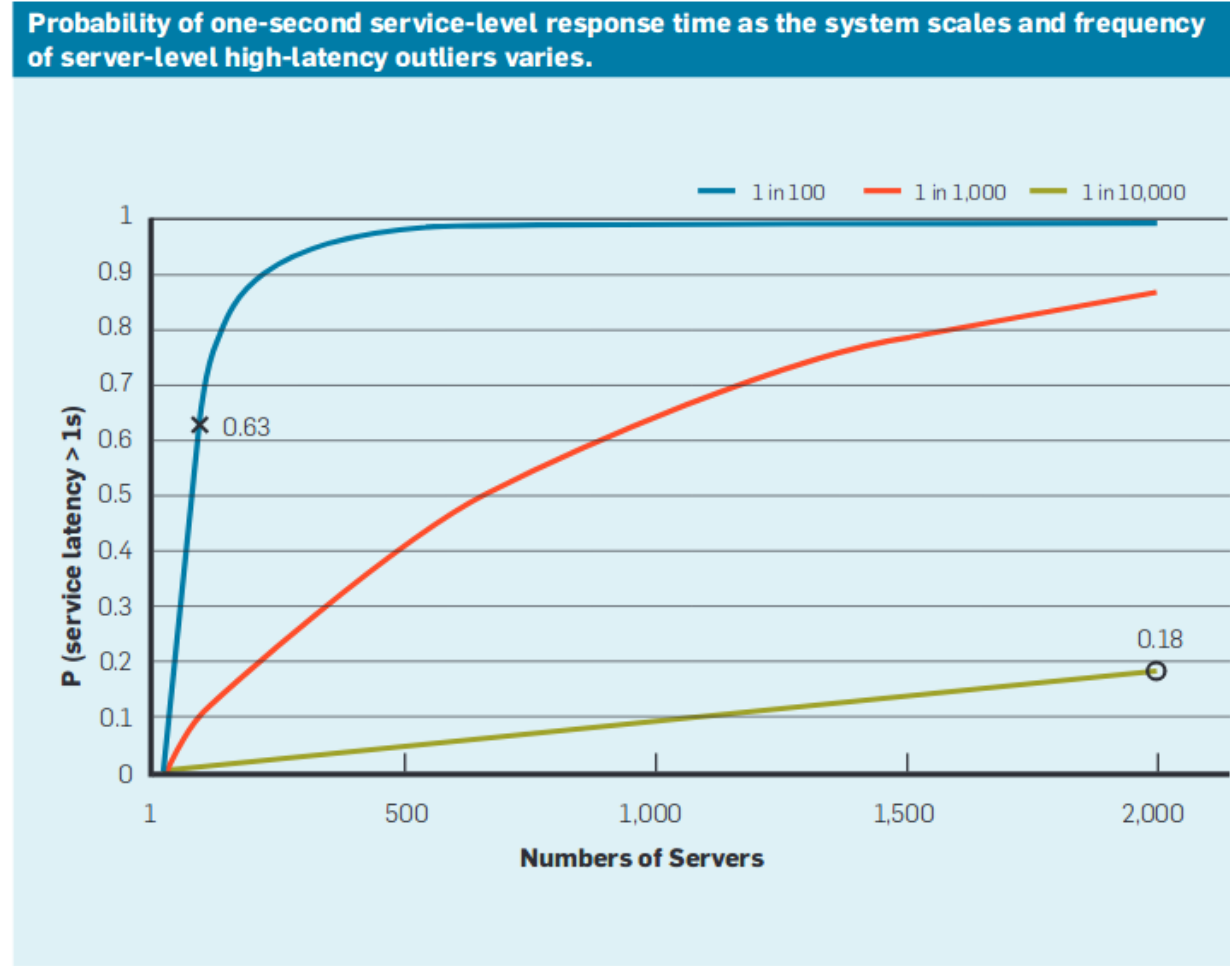
- ACM Communications september 2018
- A renewed architecture for databases



# The tail at scale



- ACM Communication February 2013
- Article written by Google engineers about how to manage response time variability
- “ Even for services with only one in 10,000 requests experiencing more than one-second latencies at the single-server level, a service with 2,000 such servers will see almost one in five user requests taking more than one second (marked “o” in the figure).”



nd system

02/10/2019