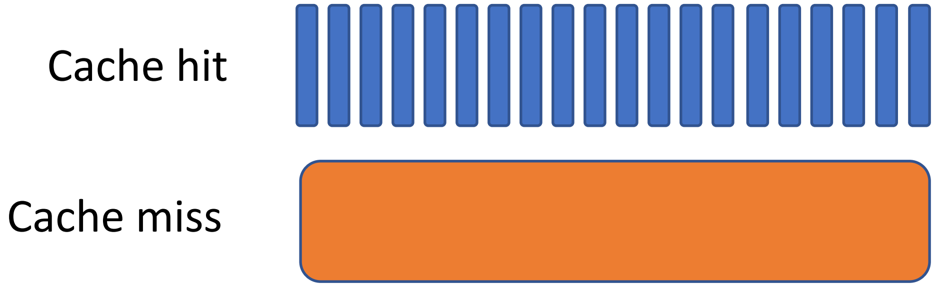
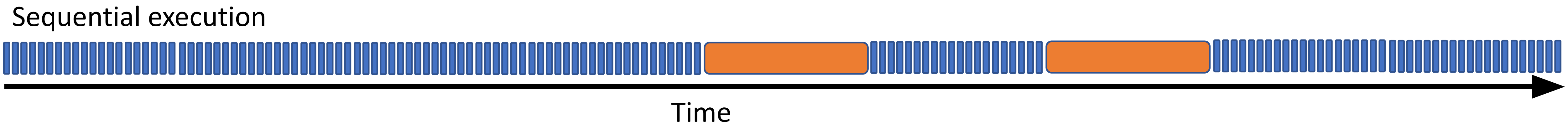
6: Heavy Tail Distributions

Caching effect on sequential computation

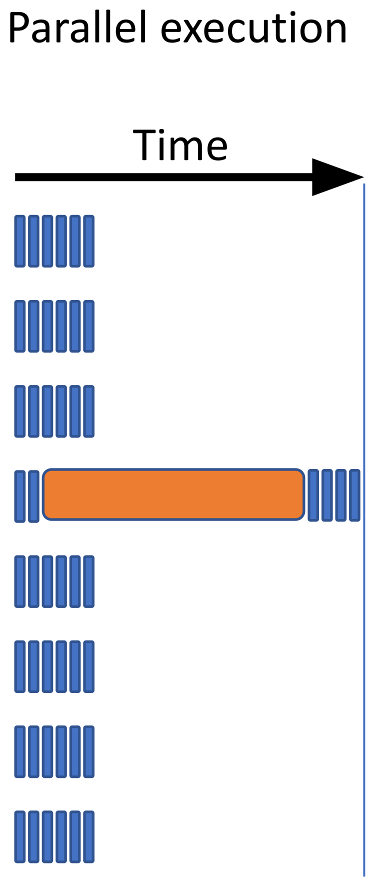


- Cache hit: 1ns
- Cache miss 100ns
- Cache miss occurs 1% of the time
- Sequential execution



- Expected time per access $\approx 2 nano\ second$
- Time to complete n accesses $\approx 2n \pm 6\sqrt{n}\sigma \quad \sigma \approx 10\ nano\ second$
- If $n = 1{,}000{,}000$ then total time $=2{,}000{,}000 \pm 60{,}000$ small variability

Caching effect on parallel computation



- Expected time per access $\approx 2nano\ second$
- Time to complete n accesses $\approx 2n \pm 6\sqrt{n}\sigma \quad \sigma \approx 10\ nano\ second$
- Suppose we have 1,000 CPUs, we have the n=1,000 per core
- We get compute time to be total time =$1{,}000 \pm 380$
- Is this right?
- Why not?
- Because normal approximation is a poor approximation
- Distribution where extreme values are much more likely than the normal are called "Heavy tail distribution"
- Very common in computing.

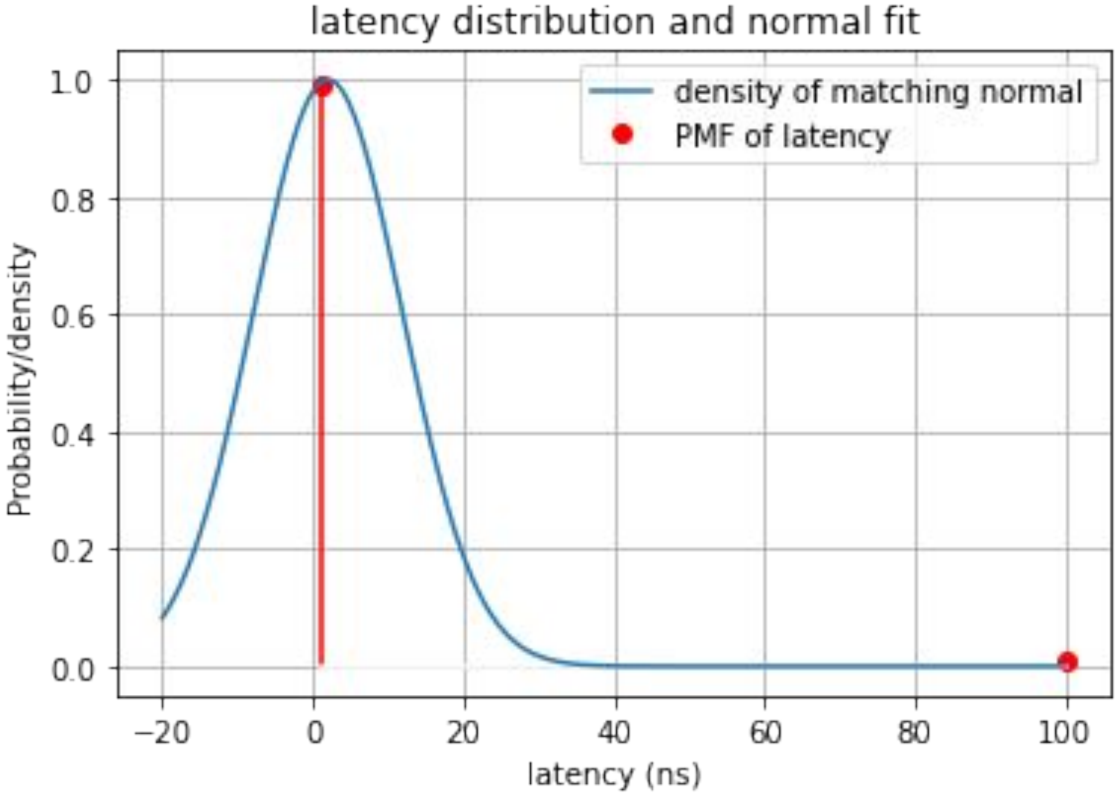
Distribution of access latencies

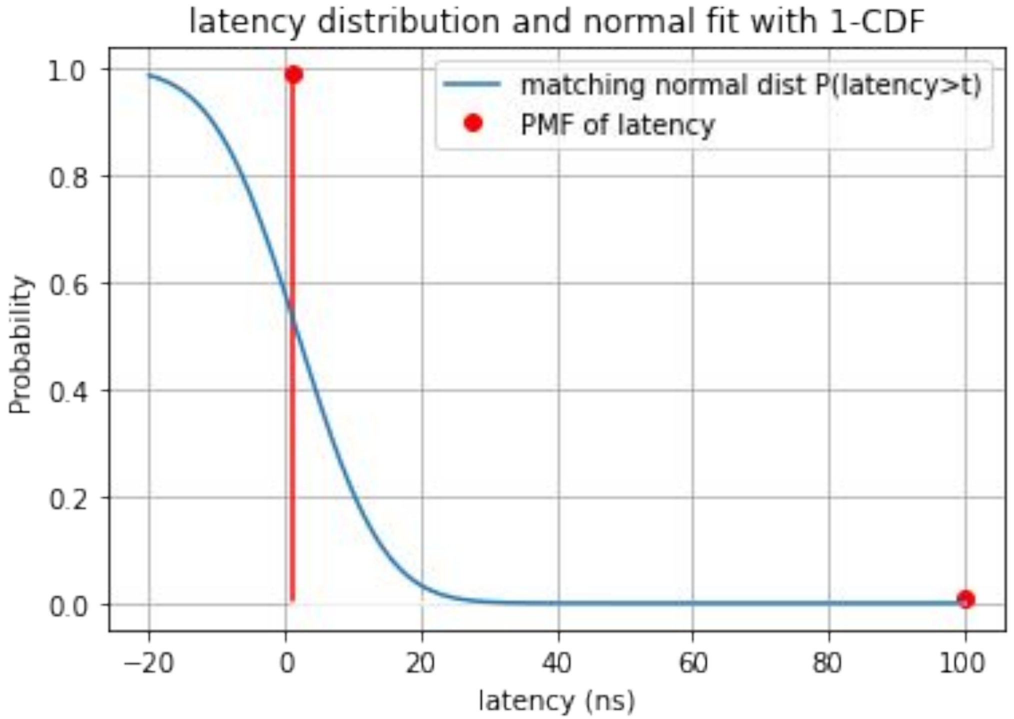
- Example: (t=latency)
  - Cache hit: 1 ns, occurs 99% of time
  - Cache miss: 100ns, occurs 1% of the time
- Expected latency: $\mathrm{E}[\mathrm{t}] = 1\times 0.99 + 100\times 0.01 = 1.99ns$
- Standard deviation:

$$\sigma(t) = \sqrt{E[t^2] - E[t]^2} \approx \sqrt{100.99 - 3.96} \approx 10$$

Heavy tails are hard to visualize

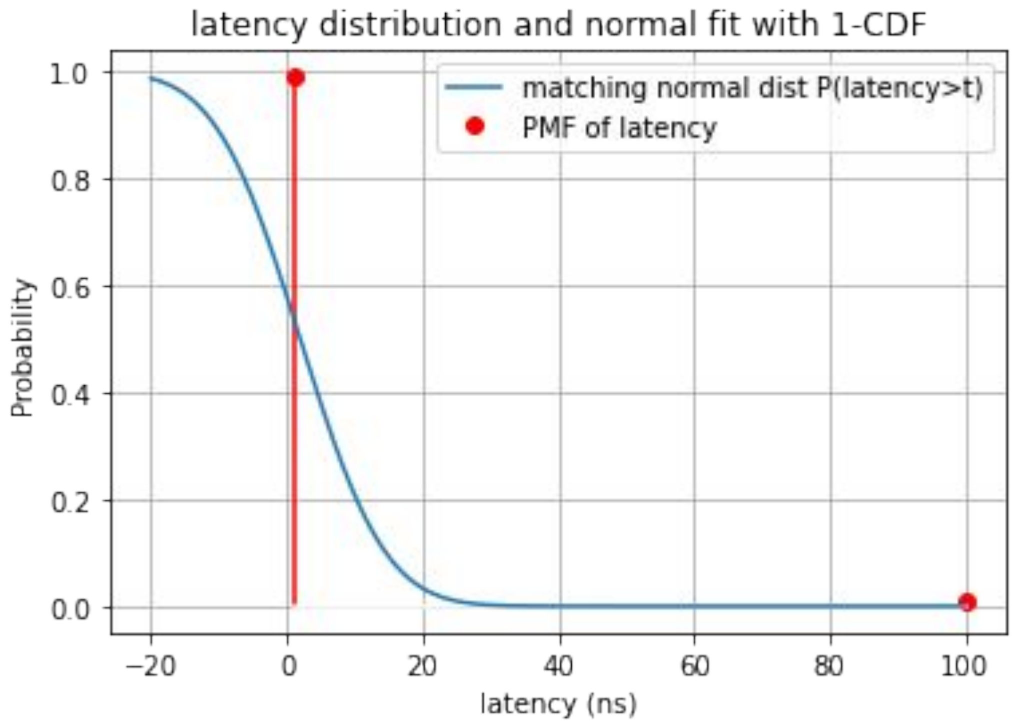
- Both the normal density and the probability look very close to zero for latency=100
- Problem 1:
  - The distribution of the latencies is a point-mass-function (PMF)
  - The distribution defined by the normal is a density (PDF)
  - The two are incomparable.

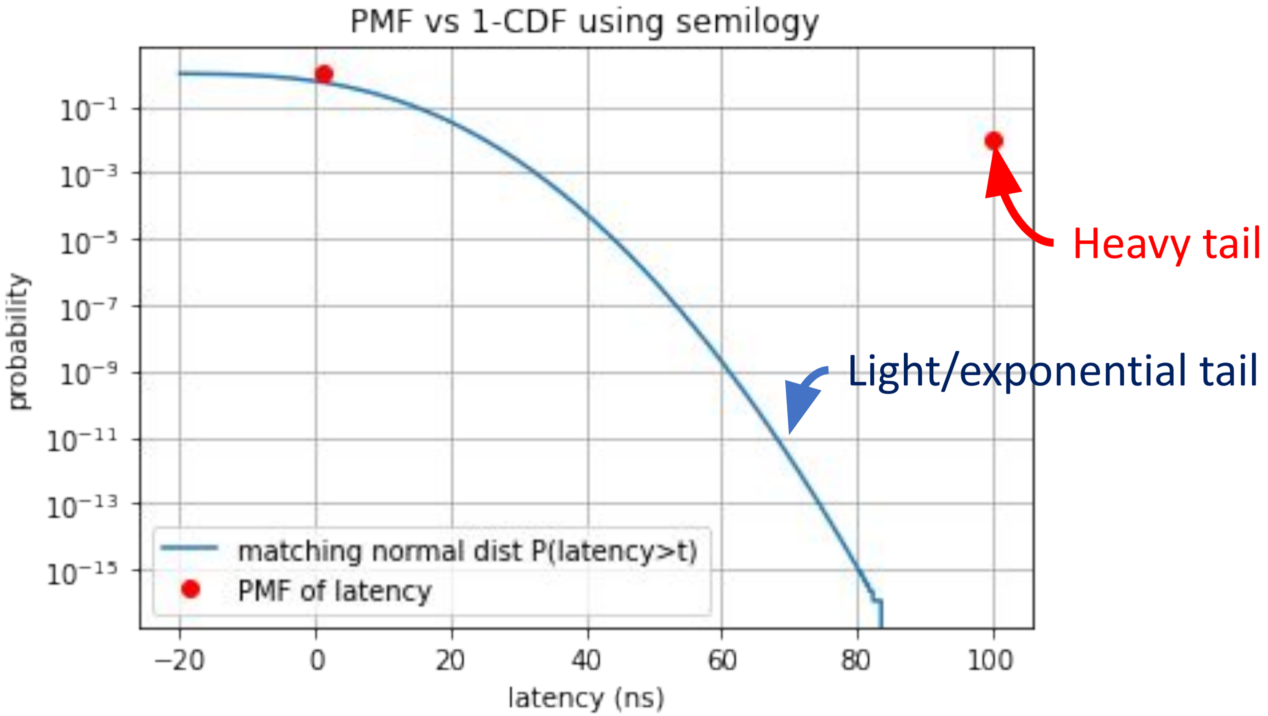




Using log scale to compare small probabilities

- $P_{normal}(latency \geq 100ns) \approx 10^{-20} \approx 0$
- The probability of a delay of 100ns is $0.01 = 10^{-2} \approx 0$
- A huge difference between these two approximate zeros:





Summary: heavy tail distribution

- A distribution has heavy tails if:
- The probability of outliers is much much higher than the probability given by the normal distribution with the same mean and variance.
- If the distribution is heavy tail, you can't use the estimate
$$E(t) \pm k\sigma(t)$$
- Heavy tail distributions are very common in the memory hierarchy.
- Why? Because a cache miss is a rare but expensive operation.

A heavy tails example

- Example: A program's run time is
  - 1 second with probability 99.9%
  - 300 sec (5 minutes) with probability 0.1%
  - Mean=1.3sec, std=9.5sec
- You run the program in parallel on 1000 dataset, 1000 computers
- You wait until all of the runs finish.
- <u>The laggard problem: </u>the slowest run determines the overall running time. More than half the time we need to wait 300 seconds.

The latency of the task is not normal

- If we assume that the distribution is normal, then the probability that even ONE of the runs takes more than 62 seconds is less than 10^{-9}
- Still, experiments show that the laggard takes more than 1000 seconds, causing the other computers to not be utilized.

The reason:

- With probability 99.9%, the run finishes in 0.5 second
- With probability 0.1%, the run finish in 1000 seconds
- When you have 1000 computers, there is a large probability that one of the machines will take 1000 seconds. Causing the fast machines to waste 999.5 seconds.

What about just one machine?

- Using just one machine.
- Run the jobs sequentially
- We are interested in the sum, not the max

```
#1000sec= 0 Probability=   0.368 ,time=  500 Cumulative mean=183.848
#1000sec= 1 Probability=   0.368 ,time= 1500 Cumulative mean=735.759
#1000sec= 2 Probability=   0.184 ,time= 2499 Cumulative mean=1195.654
#1000sec= 3 Probability=  0.0613 ,time= 3498 Cumulative mean=1410.051
#1000sec= 4 Probability=  0.0153 ,time= 4498 Cumulative mean=1478.825
#1000sec= 5 Probability= 0.00305 ,time= 5498 Cumulative mean=1495.586
#1000sec= 6 Probability=0.000506 ,time= 6497 Cumulative mean=1498.874
#1000sec= 7 Probability=7.19e-05 ,time= 7496 Cumulative mean=1499.414
#1000sec= 8 Probability=8.94e-06 ,time= 8496 Cumulative mean=1499.489
#1000sec= 9 Probability=9.86e-07 ,time= 9496 Cumulative mean=1499.499
```

- Very rarely more than 6500 sec
- Mean is 1500 Sec (it was about 1000 for 1000 machines!)
- The distribution has exponential (= light tails)

Problem in HW

- Estimate the fraction of 1’s in a large RDD
- Distribution of a single sample: 1 with prob p, 0 with probability 1-p
- Distribution of number of 1’s: binomial distribution
- Binomial approaches normal for number of samples ~>100
- We can bound the probability of a the tail using the normal distribution.