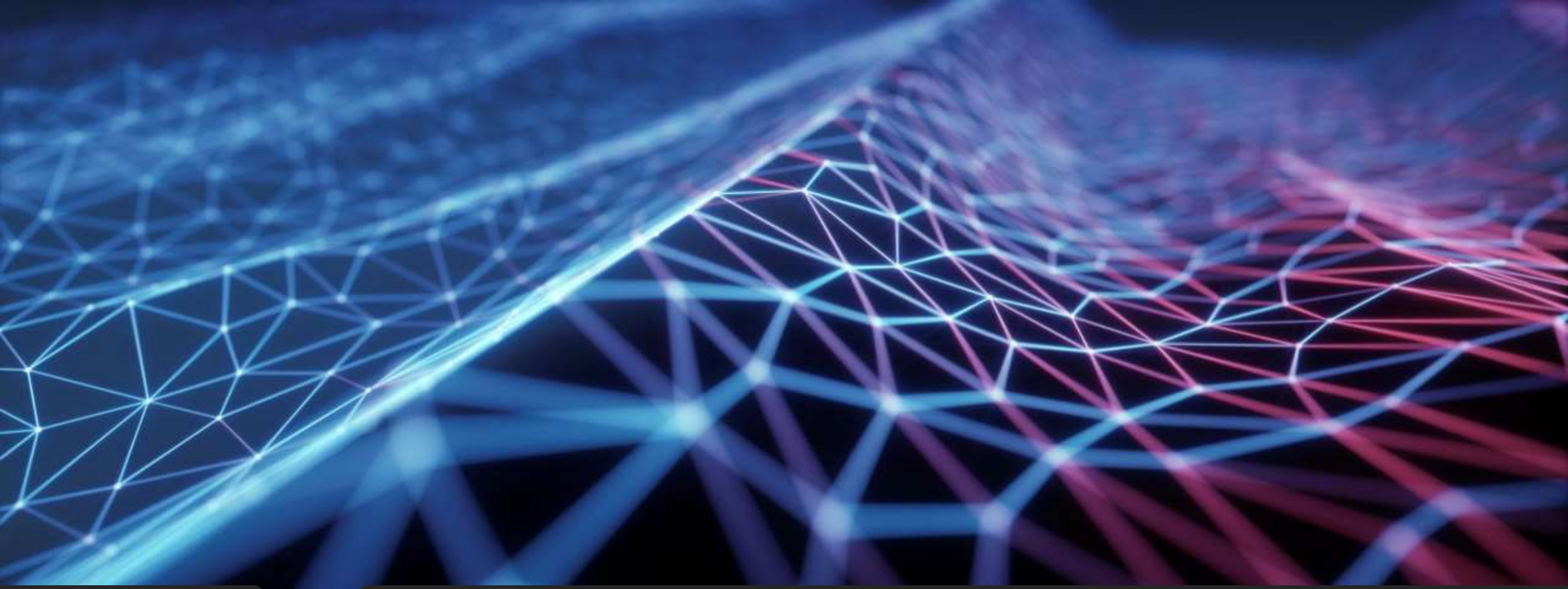
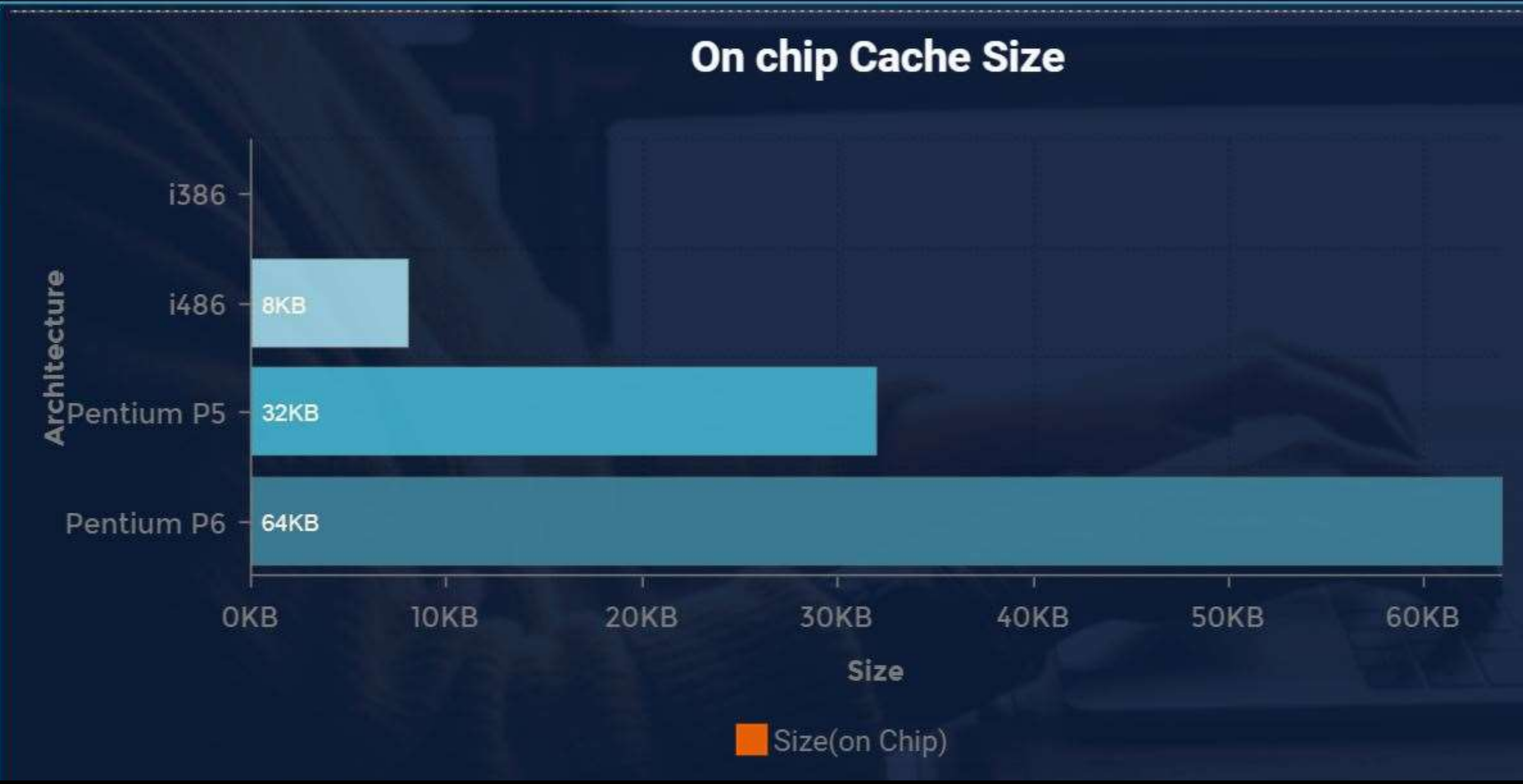


Early Cache designs

And Improvements


On chip Cache Size
Architecture
i386
i486
8KB
Pentium P5
32KB
Pentium P6
64KB
0KB
10KB
20KB
30KB
40KB
50KB
60KB
Size
Size(on Chip)

Introduction of L2 Cache with P6

- P6's most noticeable addition was its on-package **L2 cache**, which ranged from 256 KB at introduction to 1 MB in 1997

- The cache was also "non-blocking", meaning that the processor could issue more than one cache request at a time (up to 4), reducing cache-miss penalties. These properties combined to produce an L2 cache that was immensely faster than the motherboard-based caches of older processors.
- This cache alone gave the CPU an advantage in input/output performance over older x86 CPUs.

Separate Data and Instruction Cache

- Separate data Cache and instruction Cache made it possible to fetch instructions and data in parallel.

- Reduced miss rate of data cache as compared to unified cache
- The split design is useful for pipelined processors where the instruction fetch unit and the memory access unit are physically located in different parts of the chip. With the unified design, it is impossible to place the cache both close to the instruction fetch unit and the memory unit, resulting in high cache access latency=

Thank You

Ishaan Yadav