History of cache evolution and future trends

Group members

- Ishaan yadav(18114032)
- K vivekveman(18114037)
- G mahidhar(18114028)
- Aniket kumar(181160)
- Sourabh singh(181160)
- Rohit kumar(181160)
- Vishal babal(181160)

History of Cache Evolution and Future Trends

Description of the topic:
During the the early days of microcomputer technology, memory access was only slightly slower than register access. But since the 1980s the performance gap between processor and memory has been growing. Over time Microprocessors have advanced much faster than memory, especially in terms of their operating frequency, so memory became a performance bottleneck. While it was technically possible to have all the main memory as SRAM which could be as fast as the CPU,but a more economically viable path use taken: use plenty of low-speed memory, but also introduce a small high-speed cache memory to close the performance gap.
As CPUs become faster compared to main memory, stalls due to cache misses displace more potential computation; modern CPUs can execute hundreds of instructions in the time taken to fetch a single cache line from main memory.
Cache performance has become important in recent times where the speed gap between the memory performance and the processor performance is increasing exponentially. The cache was introduced to reduce this speed gap. Thus knowing how well the cache is able to bridge the gap in the speed of processor and memory becomes important, especially in high-performance systems. The cache hit rate and the cache miss rate play an important role in determining this performance. It is also important to see how these developments were made and what factors influence the design of caches in current computers.
Early cache designs focused entirely on the direct cost of using various cache designs on execution speed and efficiency in reducing the gap between processor and RAM . More recent cache designs also consider energy efficiency, fault tolerance, and goals.
what we might expect in the future based on current research and areas where room for improvement is present.

1. importance of the topic

Using direct DRAM access from the CPU slows down the overall process of program execution, because of the disadvantages of the DRAM. Using cache memory speeds up the process so that it matches or synchronizes with the CPU to achieve best results. The importance of Cache Memory in today's Microprocessor industry cannot be stressed enough. It has become an integral part of all CPUs designed and manufactured worldwide and has, in fact, been so for many decades. The Cache Memory concept has throughout been evolving and will definitely continue to do so. It has turned out to be a fundamental element of Microprocessors and one which we cannot do without. Hence its important to study the development of cache and also see if any further improvement can be made to cache implementations of CPUs. Cache hierarchy is a form and part of memory hierarchy, and can be considered a form of tiered storage. This design was intended to allow CPU cores to process faster despite the memory latency of main memory access

2. References

Intel (July 1997) Embedded Intel486 Processor Hardware Reference Manual (273025-001)

Pryce, Dave 486 32-bit CPU breaks new ground in chip density and operating performance. (Intel Corp.) (product announcement) EDN — May 11, 1989 —

https://en.wikichip.org/wiki/intel/microarchitectures/p5

https://en.wikipedia.org/wiki/Intel80386

https://en.wikipedia.org/wiki/Intel80486

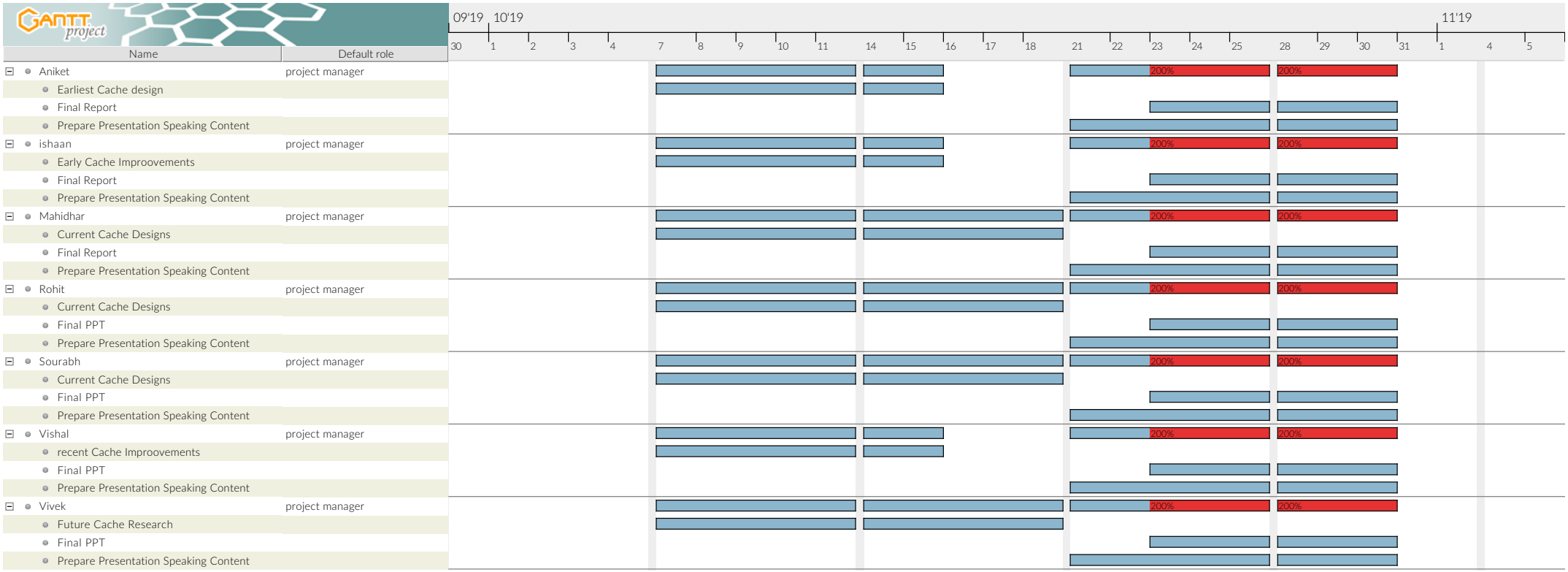
"Product Change Notication 777" (PDF). Intel. February 9, 1999. Archived from the original (PDF) on January 27, 2000. Retrieved October 14, 2019.

View Processors Chronologically by Date of Introduction, Intel, retrieved August 14, 2007

Intel Pentium Processor Family, Intel, retrieved August 14, 2007

D. Alpert and D. Avnon p. 21, "Architecture of the Pentium microprocessor", , IEEE Micro, 13, 3 (June 1993), pp. 11–21, doi:10.1109/40.216745

Resources Chart



RESULT AND DISCUSSION :

Almost all current CPUs with caches have a split L1 cache. They also have L2 caches and, for larger processors, L3 caches as well. The L2 cache is usually not split and acts as a common repository for the already split L1 cache. Every core of a multi-core processor has a dedicated L1 cache and is usually not shared between the cores. The L2 cache, and higher-level caches, may be shared between the cores. L4 cache is currently uncommon, and is generally on (a form of) dynamic random-access memory (DRAM), rather than on static random-access memory (SRAM), on a separate die or chip (exceptionally, the form, eDRAM is used for all levels of cache, down to L1).