1. Caches can drastically alter the speed of a program. With caches enabled, the CPU can take advantage of the data close to it, reduce power consumption, and the program can speed up performance.
2. A write-through cache is a cache where all copies of the data are update straight way. A write-back cache is a cache where any writes are sent to memory only when the cache entry is evicted.
3. Full-Associative cache management is when the cache is organised into a single cache set with multiple cache lines. N-Way Set-Associative cache management is when the cache is divided into ‘N’ sets and ‘M’ cache lines, where a memory block is first mapped onto a set and then placed into any cache line of that set.
4. Caches can achieve maximum efficiency by taking into account their size, structural complexity, and performance benefit, and making sure that it is the best option compared to other caches.
5. Cold cache: 10ns for first mov, cmp, beq, ldr, str, first add, b, 1ns for second mov and second add = 71 ns  
   Warm cache: 1ns for all instructions = 9ns  
   No cache: 10ns for all instructions = 90ns

mov r1, 0

mov r2, 10

for:

cmp r1, r2

beq endFor

ldr r4, =[r8]

add r4, r4, r1

str r4, [r8]

add r1, r1, #1

b for

endFor: