

## Number of accesses to cluster

This describe, the number of accesses to a same cluster per snapshot during a workload. I focus here on chain of 50 and 500 snapshots. We noticed that in vanilla version, the number of accesses is larger because we are always looking for all the past snapshots in order to find a specific cluster.

With our direct-access method, we only have the exact number of accesses which give us the real index and then we can jump to that snapshot index, omitting access to intermediate snapshots.

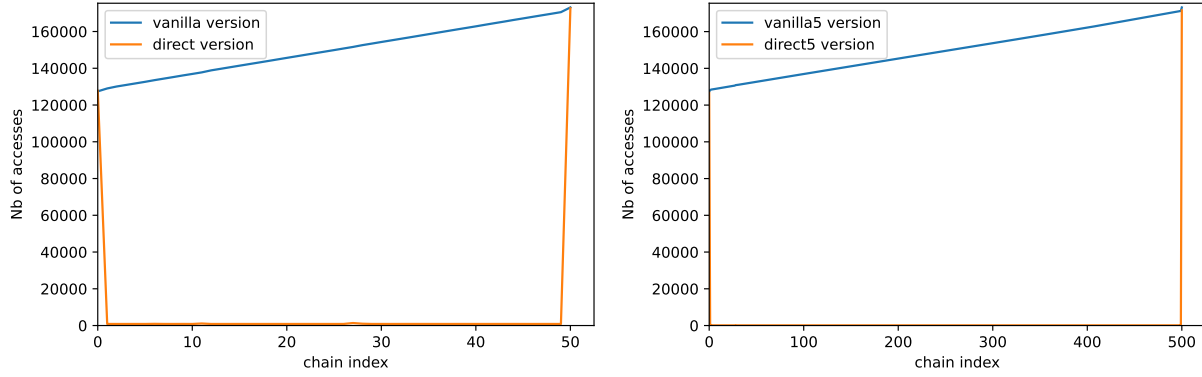


FIGURE 1. Accesses in a chain of 50 and 500 snap

## Details on type of accesses to cluster (events)

Here, we focus on explaining why there are more accesses in vanilla version, and we are doing it by distinguishing all the type of access results (here we call it *event*) in each chain.

- unallocated => the cluster is not allocated in the current snap
- normal => the cluster is allocated in the current (occured when we resolve an unallocated event)
- hit/missed => relative to the fact that cluster metadata are presented/not presented in the cache

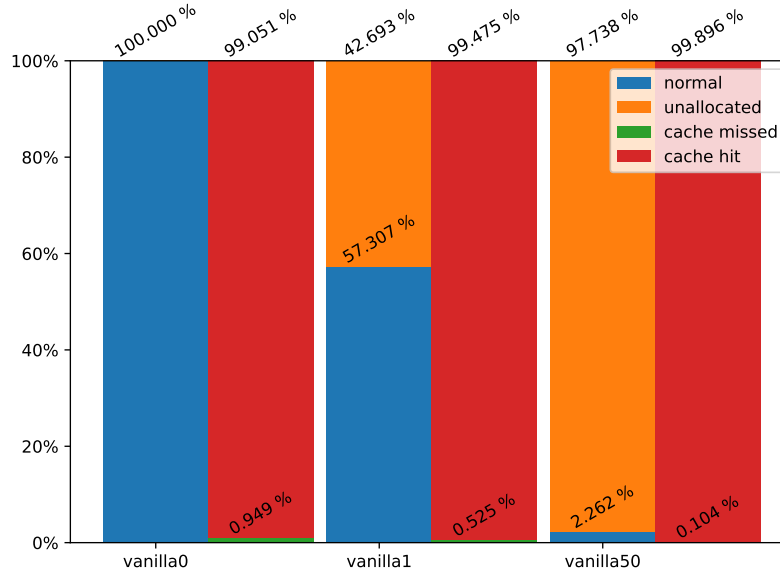


FIGURE 2. Number of events (unallocated, missed, hit, normal) for the vanilla version, the number on the x-axis indicate the chain length corresponding

We noticed that in the vanilla version, the number of unallocated events that we get in order to get the current location of the cluster (normal event occurrence) increased exponentially with the chain of the length. While in our direct-access version, we always approximatively have the same number of unallocated and normal event regardless of the chain length, this due to the fact that a same "unallocated event" is not propagate all over the chain like in the vanilla version.

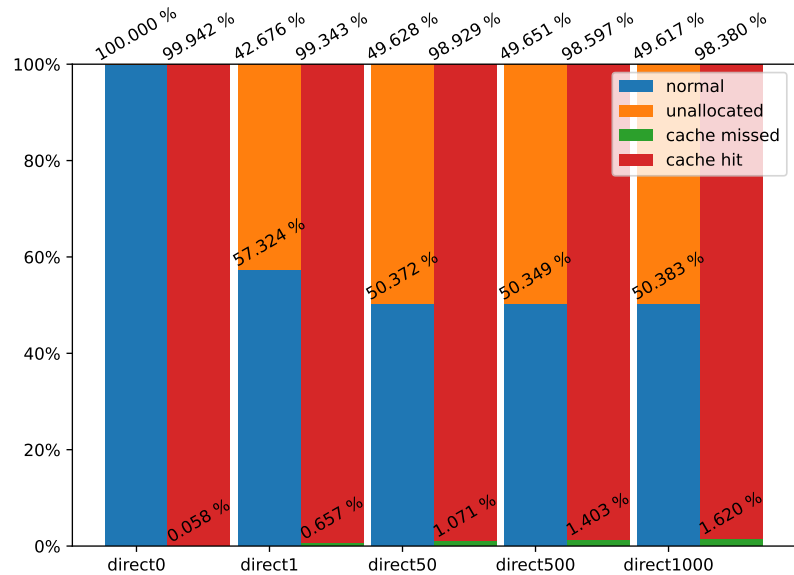


FIGURE 3. Number of events (unallocated, missed, hit, normal) in the direct-access version

Memory footprint

Comparison of memory footprint and throughput during workload dd, between our version and the vanilla one.

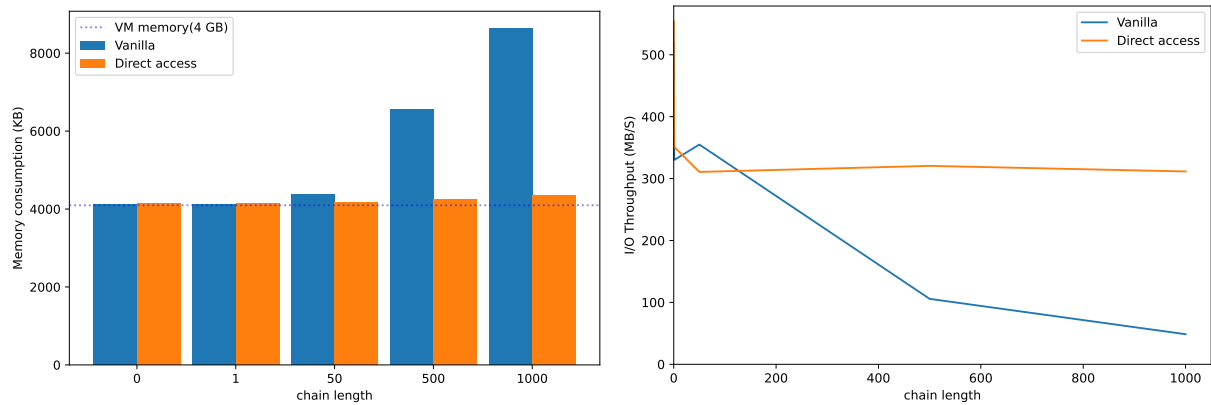
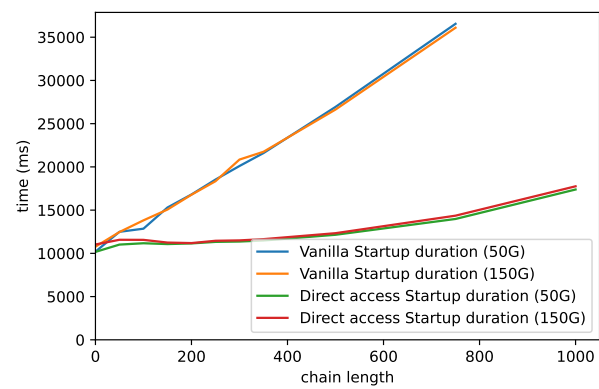


FIGURE 4. Memory consumption and Throughput (I/O) during all the workload

Startup duration

The first evaluation i did :



**FIGURE 5.** Startup, vanilla and direct-access version on 50G and 150G disk