

A single case

Design a traditional cache for GO benchmark

- design target - achieve a good hit rate and lower bandwidth
- The optimal design: 8 words line size
- Total memory Bandwidth needed: 0.32B/cycle

*A strong rule:
Larger the cache line size is,
higher the memory bandwidth the cache will ask for.
Because you need to feed the cache
with a larger window of data every time
(When fill/replace/WB).*

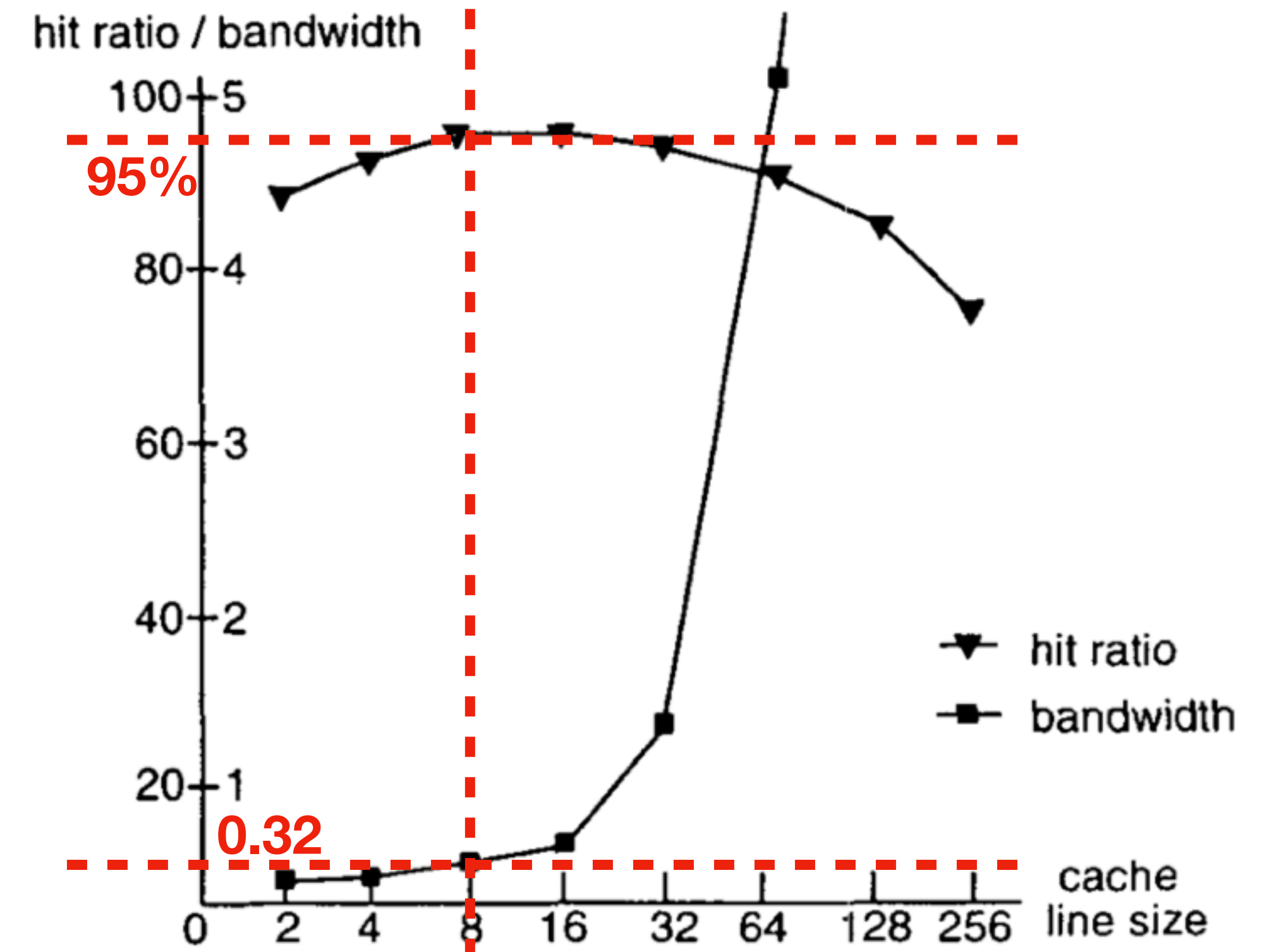
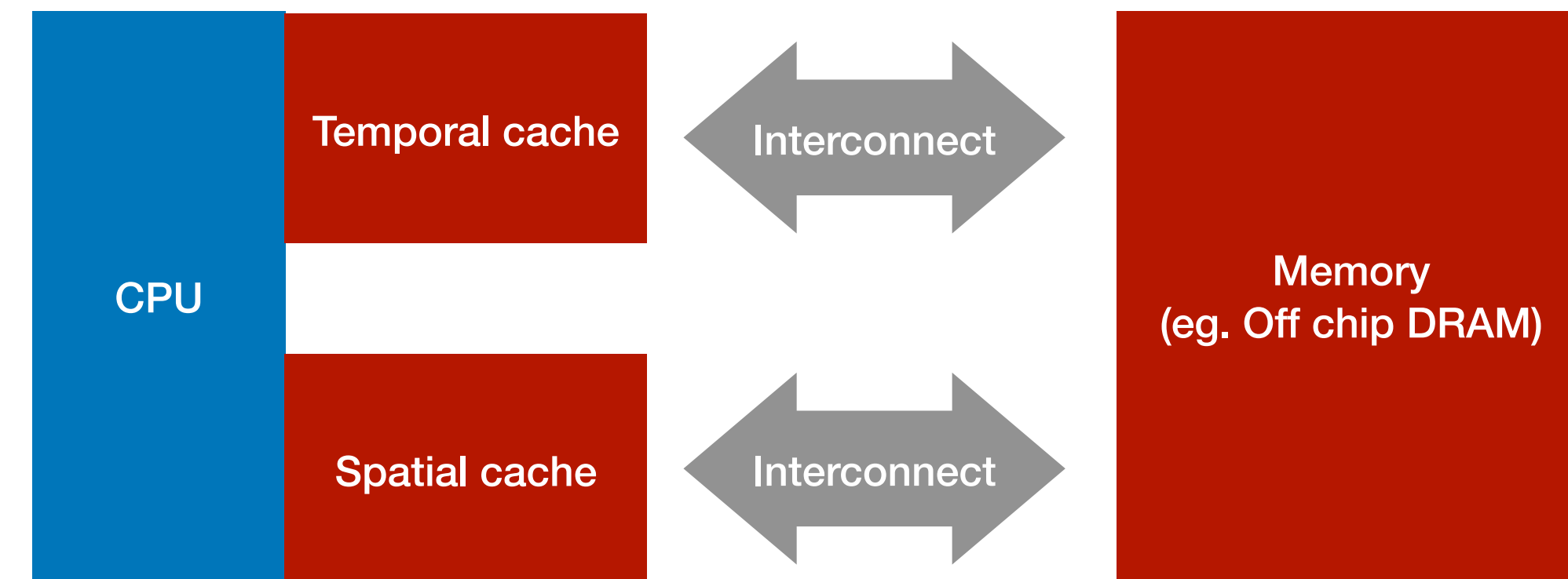


Figure 2. The cache hit ratio and memory bandwidth variation with the cache line size for the go benchmark

A single case

Then use **a temporal/spatial-spilt cache**



- The cache setup (decided by PaLM algorithm) is :
- Temporal cache: 4 words cache line size, 2KB in total
 - Variable with strong temporal locality get cached here
- Spatial cache: 8 words cache line size, 6KB in total
 - Variable with strong spatial locality get cached here
- Final memory bandwidth needed: 0.26B/cycle
- Maintained the same hit rate: 95%

We do achieve a much **SMALLER** bandwidth with NO degeneration on performance, how this contribute to Power Saving then?