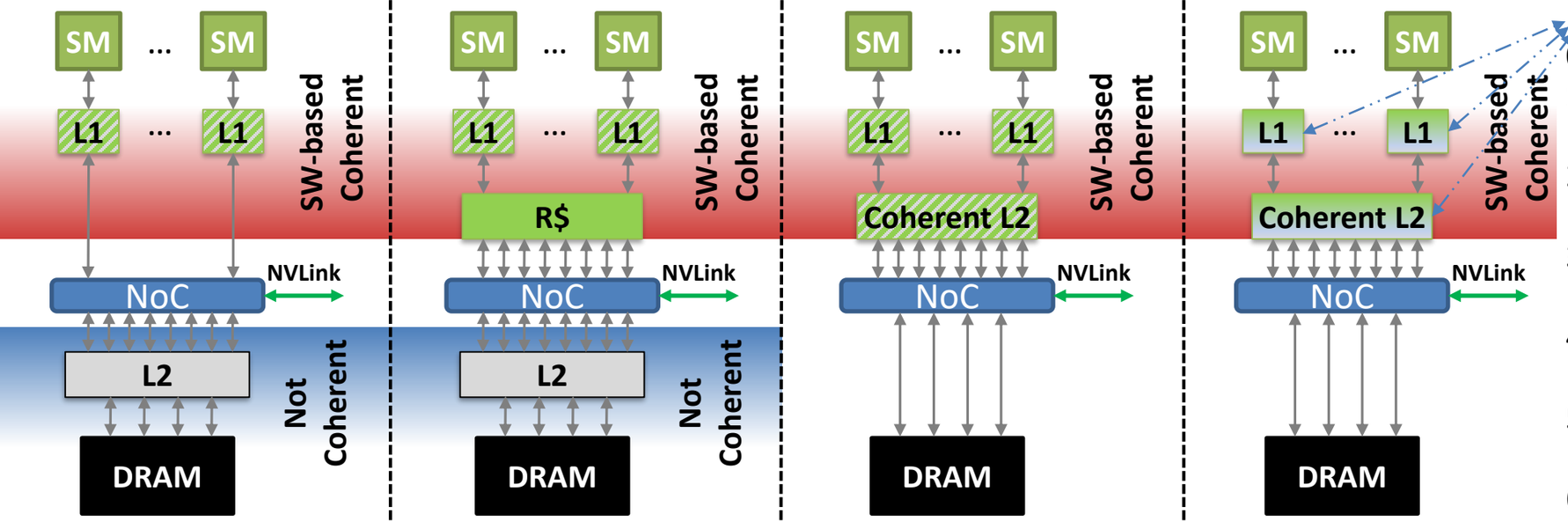

(a) Mem-Side Local Only L2

(b) Static R$

(c) Shared Coherent L1+L2

(d) NUMA-Aware L1+L2

Cache Partitioning Algorithm

0) Allocate ½ ways for local and ½ for remote data
1) Estimate NVLink incoming and monitor local DRAM outgoing BW
2) If NVLink is saturated and local DRAM BW not
   - *RemoteWays++* and *LocalWays--*
3) If local DRAM BW is saturated and NVLink not
   - *RemoteWays--* and *LocalWays++*
4) If both are saturated
   - *Equalize allocated ways (++ and --)*
5) None of them is saturated
   - *Do nothing*
6) Go back to 1) after *SampleTime* cycles