VMware Memory Tiering - besparelser, værktøj til memory-krise og forskellen fra traditionel tiering



Memory er blevet en af de største omkostninger i moderne servere.

I memory-tunge konfigurationer kan DRAM i dag udgøre langt størstedelen af den samlede hardwarepris, i nogle tilfælde op mod 90 procent.

Det er i den kontekst, at Memory Tiering i VMware Cloud Foundation (VCF) bliver interessant.

Flere kender tiering-konceptet fra storageplatforme, hvor man arbejder med hot og cold data for at balancere performance og omkostninger for mindre kritiske workloads.

Memory Tiering fra VMware blev lanceret som tech preview i 2023 og blev en fuldt understøttet feature i VMware Cloud Foundation 9 i 2025. Det har længe været en lidt overset feature, men flere og flere har fået øjnene op for den. Ikke mindst på grund af stigende memory-priser (og leveringstider) og et øget fokus på at udnytte eksisterende infrastruktur bedre.

Kort intro til Memory Tiering
Memory Tiering i VMware Cloud Foundation er en teknologi, hvor man kombinerer traditionel DRAM med hurtige NVMe-enheder og præsenterer det som én samlet memory-pool. Systemet analyserer løbende, hvilke data der er aktive, og holder de mest brugte i DRAM, mens mindre aktive data flyttes til NVMe.

Det sker automatisk og transparent, så applikationer ikke mærker forskel. Formålet er at udnytte memory mere effektivt og reducere behovet for dyr DRAM uden at gå på kompromis med performance.

Traditionel tiering og VMware Memory Tiering
Tiering er allerede et kendt koncept fra storageplatforme, hvor man også taler om hot og cold data og det kan lyde som det samme. Den største forskel handler om, hvor i infrastrukturen teknologien arbejder.

Memory Tiering i VMware Cloud Foundation (VCF) ligger i compute-laget og arbejder direkte med RAM inde i hosten.

Storage tiering ligger i storage-laget og arbejder med data på disk, typisk blokke eller filer i et SAN eller software-defined storage som vSAN.

Memory Tiering:
- Flytter memory-sider
- Arbejder med aktive workloads i realtid
- Er tæt koblet til CPU og applikation.

Storage tiering:
- Flytter datablocks eller filer
- Arbejder med persistente data
- Er længere væk fra applikationen.

Memory Tiering er en del af runtime execution. Storage tiering handler om, hvor data ligger, når det ikke er i aktiv brug.

Memory Tiering som økonomisk beslutning

Der er flere tydelige gevinster ved Memory Tiering. Memory Tiering i VMware Cloud Foundation handler ikke kun om performance, men også om økonomi. Ved at supplere dyr DRAM med billigere NVMe kan man reducere hardwareomkostninger markant, samtidig med at man får mere kapacitet ud af den eksisterende infrastruktur.

Det betyder, at man kan køre flere workloads på færre servere, hvilket også reducerer omkostninger til strøm, køling og rackplads. Samtidig udnyttes CPU-ressourcer bedre, fordi de ikke længere står og venter på tilgængelig memory.

Resultatet er en lavere TCO, uden at man nødvendigvis går på kompromis med performance.

Memory Tiering som værktøj under memory-krisen

Memory Tiering i VMware Cloud Foundation kan også ses som et konkret værktøj i den såkaldte SSD- eller memory-krise, hvor priserne på DRAM og højtydende storage er steget markant.

I stedet for at være afhængig af at købe mere og dyrere memory for at skalere, giver Memory Tiering mulighed for at udvide kapaciteten ved hjælp af NVMe, som typisk er billigere og lettere at få fat i.

Det betyder, at man kan udskyde hardwareinvesteringer og reducere presset på supply chain, samtidig med at man fastholder eller endda øger kapaciteten. Fleksibiliteten i konfigurationen gør det muligt løbende at tilpasse sig markedssituationen, så man ikke er låst til én type hardware eller en bestemt prisudvikling.

Memory Tiering er en måde at designe infrastrukturen på. I stedet for at tilføje mere kapacitet får man udnyttet den kapacitet, man allerede har.

Hvis du vil vide mere om Memory Tiering, kan du læse mere her.