

# Tiered Storage

## Què és el “tiered storage”?

El “tiered storage”, o emmagatzematge en nivells, és un sistema per assignar dades i classificar-les jeràrquicament en funció del seu valor empresarial i segons la freqüència amb què accedeixen els usuaris i les aplicacions. S’assignen a diferents tipus de suports d'emmagatzematge en funció del preu, el rendiment, la disponibilitat i la recuperabilitat de les dades. Aquests factors són els principals per determinar quin tipus de suport (cinta magnètica, emmagatzematge en núvol o discs per exemple) és ideal per emmagatzemar una categoria determinada de dades, que és un dels majors reptes que presenta aquest tipus de sistema jerarquitzat.

Les dades solen estar classificades en 4 tiers, segons la importància i de la freqüència d'accés: “Mission-critical” data , “Hot” data, “Warm” data i “Cold” data; reclassificar aquestes dades de manera rutinària a mesura que passa el temps i es fan operacions és un altre gran repte d’aquest tipus de sistema. Basada en aquestes 4 categories es solen fer 4 nivells de “tiered” storage”, un per cada tipus de dades, encara que es podria dividir en més, al tier 0 sempre tindrà les dades de tipus “mission-critical” i a l’últim tier les dades de tipus “cold”:

Data storage tiering hierarchy			
How many storage tiers an organization has largely depends on how it classifies data.			
Tier	Data category	Data description	Example storage media
0	Mission critical	<ul style="list-style-type: none"><li>■ Data that supports critical, high-performing workloads that cannot afford delays or disruptions in service</li><li>■ Data requirements outweigh storage costs</li></ul>	<ul style="list-style-type: none"><li>■ NVMe SSDs</li><li>■ RAM</li><li>■ Storage-class memory (e.g., Optane)</li></ul>
1	Hot data	<ul style="list-style-type: none"><li>■ Data that is used continuously to maintain day-to-day business operations</li><li>■ Data needs are balanced against storage costs</li></ul>	<ul style="list-style-type: none"><li>■ SSDs</li><li>■ High-performing HDDs</li><li>■ Hybrid storage systems</li></ul>
2	Warm data	<ul style="list-style-type: none"><li>■ Data that's accessed infrequently or not in constant use but might still be required on occasion</li><li>■ Cost considerations are given greater priority</li></ul>	<ul style="list-style-type: none"><li>■ SATA HDDs</li><li>■ Backup appliances</li><li>■ Tape storage</li><li>■ Cloud storage</li></ul>
3	Cold data	<ul style="list-style-type: none"><li>■ Data that is rarely accessed or updated, if at all, or stored only for archival purposes</li><li>■ Uses the least expensive storage</li></ul>	<ul style="list-style-type: none"><li>■ Slow-spinning HDDs</li><li>■ Optical discs</li><li>■ Tape storage</li><li>■ Archival cloud storage</li></ul>

Figura 1: Divisió de dades. Font: Tectarget

En general, les dades més importants es distribueixen des dels mitjans d'emmagatzematge més ràpids, que normalment són els més cars, i com més nivells hi hagi disponibles, més opcions tenen els administradors sobre la col·locació de classes de dades específiques i més eficient es poden utilitzar els recursos d'emmagatzematge. Per exemple, un sistema de dos nivells es una opció bastant limitada, ja que solament podem distribuir les dades en dos categories; una de tres sol ser més eficient, ja que es disposa d'un nivell més i es podria dividir millor les diferents tipus de dades esmentats, i els administradors poden optar per emmagatzemar les dades al nivell més costós només quan el nivell de rendiment requerit ho requereixi. A mesura que augmenta el nombre de nivells d'emmagatzematge, també es desplega un suport d'emmagatzematge més barat.

En el cas més comú, com s'ha dit abans, es tindrà un “four-class tiered storage”:

## **Tier 0**

El nivell 0 s'adapta millor a dades de tipus “mission-critical” amb poca tolerància a la latència i al temps d'inactivitat, i altament volàtils. Són dades necessàries per al funcionament empresarial i requereixen mitjans d'emmagatzematge d'alta velocitat per a un accés i un ús sense interrupcions. La incapacitat d'accedir a aquest tipus de dades crítiques podria produir problemes com un rendiment baix, la pèrdua de negoci, retards o temps d'inactivitat.

Alguns exemples inclouen transaccions amb clients o sistemes de venda d'entrades. Per oferir IOPS d'alta velocitat per a càrregues de treball crítiques, podem utilitzar SSD NVMe per a l'emmagatzematge de nivell 0 per exemple.

## **Tier 1**

El nivell 1 està dissenyat per emmagatzemar “hot” data i càrregues de treball d'alt rendiment on només es poden tolerar retards curts. Aquestes dades s'utilitzen amb freqüència i s'accedeix constantment així que també requereixen mitjans d'emmagatzematge cars i ràpids, però no tant com al tier 0, que el rendiment és important però s'ha d'equilibrar amb el cost.

Alguns exemples inclouen el correu electrònic i aplicacions com ERP, CRM, o qualsevol aplicació en què es basen les operacions comercials diàries de l'organització.

Per a l'emmagatzematge de nivell 1, s'admeten normalment unitats de SAS Enterprise, per exemple.

## Tier 2

El nivell 2 és el més adequat per a dades d'accés poc freqüent, les dades "warm". Aquest tipus de dades solen tenir uns quants dies d'antiguitat, i encara que s'utilitzen amb menys freqüència han de romandre fàcilment accessibles. Els mitjans d'emmagatzematge necessaris se centren més en la durabilitat i la rendibilitat per sobre de l'accessibilitat i les velocitats de lectura/escriptura però el cost és la consideració principal per a les dades "warm", ja que encara que no és la prioritat, s'ha de proporcionar velocitats d'accés raonables.

Alguns exemples inclouen informació sobre bitllets, correus electrònics antics o transaccions de la setmana passada.

Per a l'emmagatzematge de nivell 2, normalment s'ofereixen unitats SATA o emmagatzematge al núvol a Azure, per exemple.

## Tier 3

El nivell 3 és el més adequat per arxivar dades ("cold" data) a les quals s'accedeix rarament o mai, i està creat per a l'assequibilitat i la retenció a llarg termini d'aquestes dades. Són guardades per motius reglamentaris o perquè podria tenir valor en el futur. En altres casos, també es mantenen per facilitar l'anàlisi de big data que no és sensible al rendiment. Les dades fredes són adequades per al nivell d'emmagatzematge de menor cost disponible, amb temps d'accés de minuts o hores.

Alguns exemples inclouen documents de fa molt de temps.

Els mitjans d'emmagatzematge preferits per a "cold" data inclouen matrius de cintes o nivells d'emmagatzematge en núvol fred com ara Azure Cool Blob o AWS S3-IA.

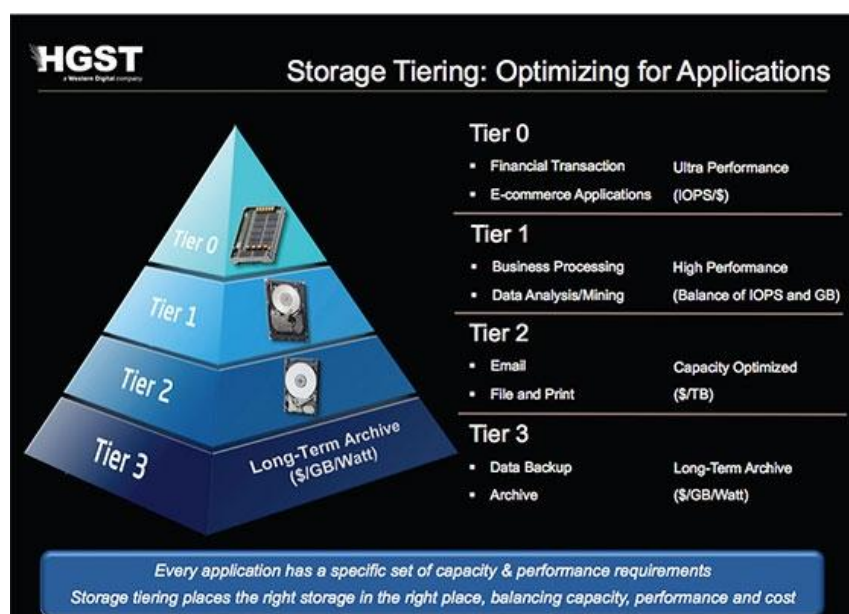


Figura 2: Tiers dividits. Font: TweakTown

## Compromís entre la protecció i el cost

Com s'ha explicat fins ara, podem veure que tant el rendiment com el cost són factors importants a tenir en compte a l'hora de decidir un bon disseny d'un sistema d'emmagatzematge. No podem prescindir de la variable del preu ja que sinó tindrem solament el Tier 0 per tot tipus de dades i això seria molt car. Tenint el sistema jerarquitzat com és el nostre cas, en tiers, ens ajuda a distribuir les diferents tipus de dades (depenent de la importància d'aquestes dades i la freqüència d'accessos) en els diversos nivells, que a mesura que incrementa els tiers el preu va baixant, ja que el tier 0 és molt més car que el tier 3 (en el cas d'un "4-class tiered storage"). Això ens ajuda a reduir el cost total del sistema i també encaixa amb les nostres necessitats.

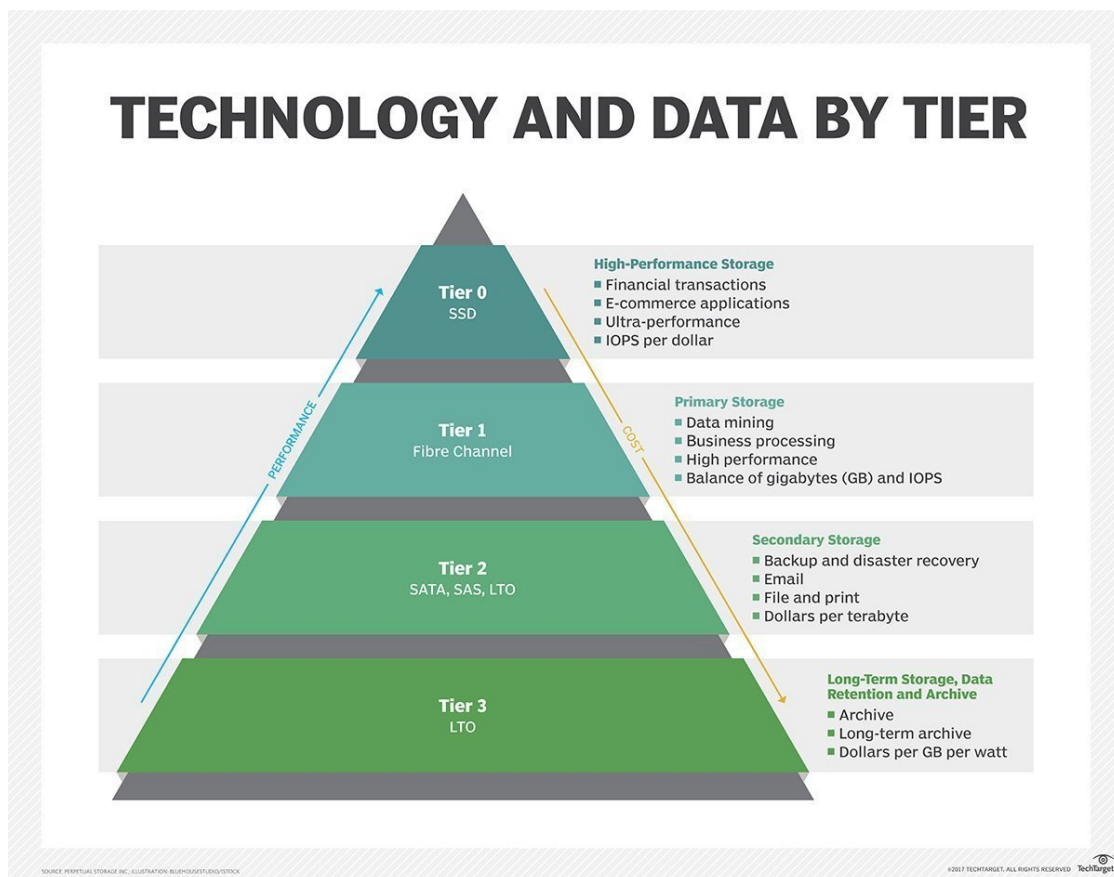


Figura 3: Característiques del "tiered storage" Font: Techtargget

## Bibliografia:

- <https://www.techtargget.com/searchstorage/definition/tiered-storage>
- <https://www.druva.com/glossary/what-is-tiered-storage-definition-and-related-faqs/>
- <https://cloudian.com/guides/data-backup/storage-tiering/>
- <https://stonefly.com/resources/what-is-tiered-storage>