

# Tiered Storage

## Introducció: què és el *tiered storage*?

[1] El *tiered storage* és una forma d'assignar diferents categories de dades a diferents tipus de mitjans d'emmagatzematge amb l'objectiu d'augmentar el nivell de servei per aplicacions i conjunts de dades crítiques, i alhora reduir el cost general de l'emmagatzematge de dades. Una arquitectura d'emmagatzematge per nivells (*tier storage architecture*) organitza les dades en una jerarquia, d'acord amb el seu *valor comercial*. Els diferents nivells de la jerarquia estan determinats pel rendiment i el cost dels mitjans, mentre que les dades es classifiquen, principalment, segons la freqüència amb la qual els usuaris accedeixen a elles. En general, les dades més importants es guarden als mitjans d'emmagatzematge més ràpids, que acostumen a ser els més costosos. Habitualment, com més petit és el nombre del *tier* en la jerarquia, més car és el mitjà d'emmagatzematge i menor és el temps necessari per recuperar les dades d'aquell nivell. Així, per exemple, una empresa que requereix un accés molt ràpid a certes dades pot utilitzar un costós sistema de discs *Enterprise* d'estat sòlid, que formaria part del que es coneix com a *Tier 0*.

El *tiered storage* forma part de la base de la gestió del cicle de vida de la informació (ILM) i ajuda a les empreses a reduir els costos totals d'emmagatzematge, al temps que contribueix a augmentar el seu rendiment i fiabilitat.

## Una mica d'història:

[2] La IBM va ser l'empresa pionera en utilitzar arquitectures d'emmagatzematge per nivells en les seves computadores *mainframe*. El primer (proto-)exemple conegut d'aquesta pràctica es relaciona amb el *IBM 3850 Mass Storage Facility (MSS)*, una biblioteca de cintes *on-line* que s'utilitzava per emmagatzemar grans quantitats de dades a les quals s'accedia de forma poc freqüent [3], l'any 1974.



Fig. 1: Data cartridge del MSS. Cada cartutx contenia un rodet de cinta de 20 m, amb una capacitat d'emmagatzematge de 50 MB [4].

Durant molts anys, quan el concepte encara havia d'evolucionar, el *tiered storage* implicava col·locar “manualment” dades en diferents configuracions de discs durs (com ara SAS i SATA). Les dades s'escriuen en blocs en disc utilitzant tècniques com el *short stroking* o el *disk stripping*. Això va acabar resultant en nivells d'emmagatzematge amb diferents característiques: capacitat, cost i rendiment. L'auge de la gestió d'emmagatzematge jeràrquic (HSM) va impulsar la reducció del procés *manual* del *storage tiering*. L'automatització del *software* trans porta les dades dinàmicament entre diferents sistemes d'emmagatzematge, tipus d'unitats o grups RAID en temps real, de forma transparent a l'usuari.

## Tier 0:

[5] El nivell 0 (Tier 0) és el nivell d'emmagatzematge de dades més ràpid i més costós d'entre tots els que hi ha. La *creació* d'aquest nivell és més recent que la dels altres, i ha estat motivada en gran part per l'augment de l'emmagatzematge en estat sòlid i a la tecnologia *flash*. Cal a dir, però, que quan el HDD era dominant, els administradors col·locaven les dades de Tier 0 en unitats HDD més ràpides i costoses, mentre que les dades *menys importants* s'assignaven, lògicament, a unitats de disc més lentes i més econòmiques. La idea, com hem comentat, s'ha mantingut, però els discs SSD han reemplaçat als HDD com a medi principal per a les dades de Tier 0.

El Tier 0 és més ràpid que l'emmagatzematge *tradicional* de Tier 1, i la realitat és que moltes dades que abans es consideraven de Tier 1 ara s'emmagatzemen en un Tier 0. L'emmagatzematge en Tier 0 és adequat per a aplicacions amb molt poca tolerància al temps d'inactivitat o a la latència. Les dades que s'assignen a aquest nivell inclouen, entre d'altres: dades financeres, *scale-up transactional databases*, dades relacionades amb aplicacions mèdiques o de seguretat. La *Figura 2* mostra una petita taula on es pot veure quin tipus de dades s'assigna a cada nivell de la jerarquia.

Sample data storage hierarchy		
TIER	DATA	STORAGE MEDIA
0	Transactional data requiring extremely high performance	Single-level cell solid-state storage
1	Mission-critical application data	Fibre Channel storage area network or SAS hard disk drives
2	Less critical data that could be recovered after mission-critical application data	SATA disk arrays
3	Data that is seldom used	CD-R or tape
4	Archived email retained for compliance	Public cloud

Fig. 2: Tipus de dades associades a cada Tier d'una tiered storage architecture. Font: TechTarget [2]

## Tier 1:

[2] Les dades de Tier 1 inclouen *mission critical data*, *hot data* o informació delicada, que ha de mantenir-se fortament protegida. Aquestes dades s'emmagatzemen en sistemes costosos i de gran qualitat, com el *double-parity RAID* (RAID 6).

El Tier 1 es reserva per a aplicacions que depenen de lectures i escriptures ràpides, com per exemple, aquelles relacionades amb els ingressos i les operacions comercials (pensem en una base de dades transaccional *on-line* que respon a *queries* fetes per aplicacions d'alta velocitat, en temps real).

El mitjà d'emmagatzemament més habitual d'aquest *tier* són els discs SAS *Enterprise*.

## Tier 2 i Tier3:

Fins i tot quan es completen les operacions primàries, les dades de Tier 1 acostumen a escriure's simultàniament en un nivell secundari de dispositius de suport basats en disc o en cinta magnètica. Com sabem, els centres de dades implementen *un nivell de suport* per assegurar la continuïtat del seu negoci i la recuperació de les seves dades davant de

possibles desastres mitjançant restauracions ràpides de dades essencials per al negoci i *hardware* d'emmagatzematge específic.

Moltes organitzacions mantenen parts de les seves còpies de seguretat en disc durant un cert període de temps determinat, després del qual es passen completament a cintes per a guardar-les durant el temps que faci falta. Generalment, les dades que s'emmagatzemen en Tier 2 contenen informació històrica (com, per exemple, informes financers de l'històric d'una organització o històries clíniques dels pacients d'un hospital, que s'han de guardar, per qüestions legals, durant un gran nombre d'anys, però que, com a norma general, no s'han de consultar amb freqüència), *cold data* o arxius classificats. Les dades de Tier 2 es guarden en mitjans de menor cost en una xarxa d'àrea d'emmagatzematge (SAN) convencional.

Les còpies de seguretat de nivell 2 també poden incloure sistemes de planificació de recursos empresarials (ERP), correu electrònic corporatiu o aplicacions administratives. En resum, l'emmagatzematge de Tier 2 protegeix les dades de les aplicacions que requereixen alta fiabilitat i seguretat, però que no necessiten una latència baixa.

D'altra banda, si es disposa d'un Tier 3 al sistema d'emmagatzematge, s'utilitza com a una espècie de *support*. Les dades que van a Tier 3 poden ser arxius que s'utilitzen amb molt poca freqüència o, fins i tot, arxius sense classificar, que poden guardar-se perfectament en discs durs molt lents o, més habitualment, en cintes. En principi, es poden guardar en Tier 3 còpies de qualsevol contingut que es consideri que té algun valor *estratègic* per al negoci, per petit que aquest sigui. El contingut pot emmagatzemar-se indefinidament o pot configurar-se per a què s'esborri en un cert temps.

La *Figura 3*, a mode de visió general, mostra una taula on s'estimen (entre altres valors) els percentatges de dades digitals que es guarden en cadascun dels *Tiers* que hem vist, així com la tecnologia emprada a cadascun dels nivells i el temps d'accés nominal a les dades que s'emmagatzemen, a nivell global.

**Inside the Storage Tiers - The Physical View 2020**

Storage Tier	Tier 0 Ultra-high Performance	Tier 1 Performance	Tier 2 Active Archive	Tier 3 Archive, Long-term
Amount of Data in Each Tier	10%	10%	20%	60% (or more archival data)
Primary Technology and Interface	NVM (DRAM, 3D-Flash SSD, PCM, 3D-Xpoint)	Enterprise disk arrays	Midrange disk arrays – scale out	Tape libraries, offsite data vaults, cloud services
Nominal Access Time	1-10 $\mu$	5-10 ms	5-20 ms	25-121 sec
Data Transfer Rates	550/520 MB/s R/W speed 3,500/2,300 MB/s R/W speed	160-220 MB/s	80-220 MB/s	360 MB/s LTO 400 MB/s Enterprise
Typical File Access	Random/Seq.	Random/Seq.	Random/Seq.	Sequential only
Data Classification Category	I/O intensive, response-time critical, OLTP, ultra high-performance	Mission-critical, OLTP, revenue generating applications	Vital, sensitive, business important applications	Archives, fixed content, big data, reference data, govt. regs, high data rates
2020 est. price *	~\$140/TB	~\$40/TB	~\$30/TB	~\$4/TB
Reliability (BER)	$1 \times 10^{17}$	$1 \times 10^{16}$	$1 \times 10^{15}$	$1 \times 10^{19}$
Media Life	3-5 years	4-5 years	4-5 years	>30 years
Power rating	3-5 W	6-15 W	6-15 W	Lowest

Fig. 3: Taula resum amb l'ús i les característiques dels diferents Tiers que hem vist. Font: HORISON [6]

## Protecció vs Cost:

Com hem vist anteriorment, si el rendiment fos l'únic factor a tenir en compte en el disseny d'un sistema de dades, aleshores tots els negocis farien servir *all-flash array* de Tier 0 per guardar les seves dades. Però, la realitat és que el cost també és un element que s'ha de tenir en consideració; i per això és útil disposar d'una jerarquia com la que hem presentat a les pàgines anteriors. Establint una jerarquia per *tiers* podem implementar el sistema d'emmagatzematge menys costós per a satisfer les nostres necessitats, a partir d'un bon anàlisi de requisits del nostre sistema. Establint una divisió de les nostres dades segons la seva importància i la freqüència amb la qual és accedida, podem classificar en diferents nivells el nostre sistema d'emmagatzematge i obtenir així un balanç adequat entre el rendiment i el cost total del sistema.

D'igual manera, podem utilitzar *software* en una organització que ja està en funcionament per determinar la freqüència amb la qual les seves dades són accedides i diferenciar, en conseqüència, entre *hot data* i *cold data*. D'aquesta forma, podem intentar fer canvis en l'organització del nostre sistema de dades i ajustar-lo a les nostres necessitats reals, reduint així el cost total del sistema.

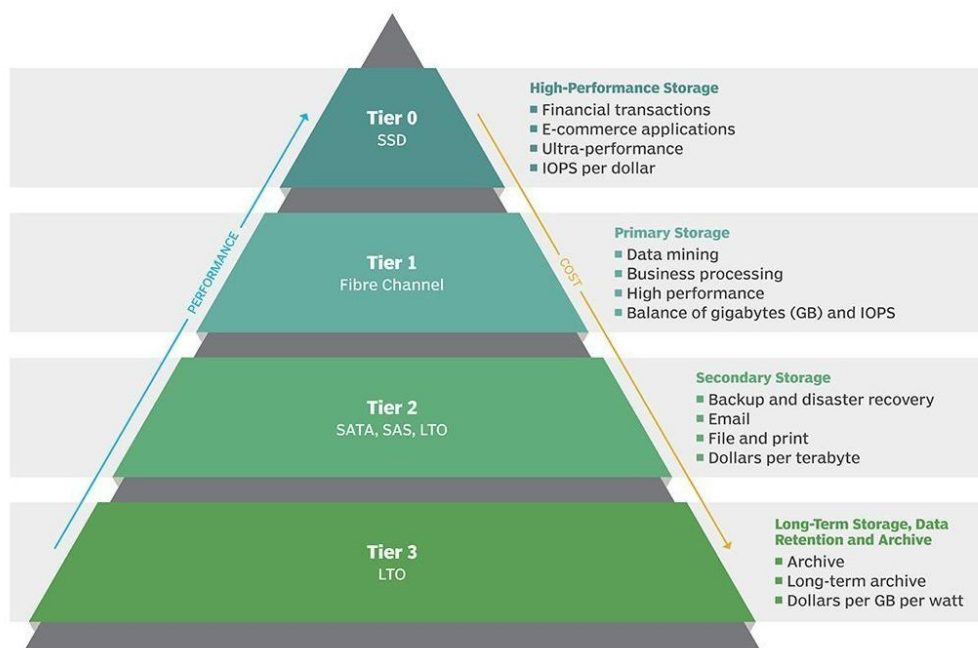


Fig. 4: Un bon disseny d'un sistema d'emmagatzematge requereix d'un balanç entre rendiment i cost.

Font: TechTarget [5]

## Referències

- [1] Durva: "[Tiered storage](#)", Druva Inc., 2021
- [2] Garry Kranz, Dave Raffo: "[Tiered storage](#)", TechTarget, 2020
- [3] Pugh, Emerson W, Johnson, Lyle R, Palmer, John H: "*IBM's 360 and early 370 systems*", MIT Press. pp. 536–540, 1991. [ISBN 9780262161237](#)
- [4] [Introduction to the IBM 3850 Mass Storage System](#). Systems (Third ed.). IBM. July 1975. GA32-0028-2.
- [5] Rodney Brown: "[Tier 0](#)", TechTarget, 2017
- [6] Fred Moore: "[Tiered Storage: Building the Optimal Storage Infrastructure](#)", Horison Information Strategies, Boulder, CO.