

Exercici 2

Motius pels que sorgeix aquesta nova tecnologia de base de dades

Podem separar els motius pels quals aquestes tecnologies han hagut de ser creades dels factors tecnològics que les han fet possibles. Els principals motius són:

1. **Temps de resposta més curts.** Les noves aplicacions requereixen de uns temps de resposta més curts. Ja no és suficient amb proporcionar la resposta correcta. Aquesta s'ha de retornar en el menor temps possible. Qualsevol temps afegit perjudica de forma directa al negoci. Les tecnologies de bases de dades tradicionals són molt fiables però relativament lentes. Un bon exemple és el de Netflix: té milers de versions de cada vídeo distribuïdes a diferents datacenters per tal de proporcionar el vídeo més adequat de la forma més ràpida possible. Una base de dades en memòria va un pas més enllà i elimina el delay introduït per l'accés a disc.
2. **Tractament de gran quantitat de dades.** Les aplicacions tradicionals tractaven amb un nombre limitat de dades. Els nous dispositius (per exemple el mòbil) generen dades en ordres de magnitud molt superiors¹. Com a exemple: Youtube reproduïx mil millions d'hores de vídeo cada dia²; Netflix, cada setmana. Netflix té més de 110 millions de clients. Ja no és necessari pensar en gegants tecnològics: avui dia un partit de fútbol normal genera més de 8GB de dades³. Tractar aquest volum de dades amb una base de dades en memòria és molt més eficient, inclús si es comprimeix la informació.
3. **Intel·ligència de negoci.** La gran quantitat de dades fa que tinguem la possibilitat de conèixer millor als nostres clients o usuaris. Per fer això, no només és necessari poder tractar una gran quantitat de dades, sinó també fer-ho de la forma més ràpida possible. El producte personalitzat és el present i el futur. Un exemple és com Netflix adapta el seu producte a cada client⁴. Les bases de dades en memòria poden fer aquesta feina d'una forma molt més eficient.

Factors tecnològics que l'han fet possible:

1. **Abaratiment de les memòries RAM.** La Llei de Moore⁵ no aplica només al nombre de transistors per CPU. La podem aplicar tant a la capacitat RAM com al seu abaratiment. Un exemple és com les memòries SSD estan baixant de preu⁶. Això fa que el que era impensable fa 10 anys (tenir una base de dades enterament en memòria) avui sigui una realitat.
2. **Creació de datacenters (cloud) que permeten tenir una estructura adequada des del primer moment.** Alguns gegants tecnològics (Amazon, Google, Microsoft, etc.) han creat una sèrie de datacenters per donar suport al seu negoci. No només això: han desenvolupat les tecnologies necessàries per tal de tractar la seva immensa quantitat de dades. Com a part del seu negoci han obert aquestes plataformes a tercers. Un exemple pot ser l'UOC, que utilitza gmail com a servidor de correu electrònic. Així doncs, qualsevol empresa que ho pugui pagar pot tenir aquesta infraestructura a la seva disposició. És un gran pas: ja no és necessari adquirir servidors, plataformar-los, comprar llicències per separat, crear estructures de rollback o de tractament de desastres, polítiques de xifrat, etc.
3. **Memòries RAM no volàtils (NVRAM).** Ja es poden utilitzar com a "discs durs", incrementant dramàticament la velocitat d'accés a les dades.
4. **Bases de dades en columnes.** Les bases de dades columnars permeten reduir el volum de dades de forma dramàtica. Els algorismes de compressió són molt més efectius a l'haver de tractar amb tipus de

¹Molt interessant el recopilatori de dades que mostra na Paloma Llaneza a Datanomics <https://www.planetadelibros.com/libro-datanomics/289284>

²<https://www.xataka.com/streaming/la-compleja-infraestructura-detras-de-netflix-que-pasa-cuando-le-das-al-play>

³<https://www.xataka.com/empresas-y-economia/asi-tecnologia-detras-liga-futbol-ia-para-predecir-mejores-horarios-var-datos-iguales-para-todos-equipos>

⁴<https://www.20minutos.es/noticia/3533237/0/algoritmo-netflix-manipula-veas-series-peliculas/>

⁵https://es.wikipedia.org/wiki/Ley_de_Moore

⁶<https://www.minitool.com/news/ssd-prices-fall.html>

dades homogenis. A més, redueixen tant l'espai com el temps necessaris per a gestionar les dades.

Avantatges i punts febles que veieu de l'ús d'aquesta nova tecnologia

Punts febles

Si ens centrem en els punts febles d'una base de dades en memòria RAM:

1. **Les dades són volàtils.** S'ha d'implementar qualche estratègia de persistència per tal de poder recuperar-les en cas de desastre.
2. **Stack overflow.** Si s'intenta fer servir més memòria de la disponible, falla. S'ha d'implementar qualche sistema de control i escalat.
3. **Gestionar les estructures de memòria.** Pot ser necessari optimitzar l'utilització de la pila de memòria (registres, caché, RAM, etc.) per tal d'evitar tenir un rendiment pobre.
4. **El preu és molt més elevat.** Evidentment, la memòria RAM és encara molt més cara.
5. **S'ha d'adaptar al nostre model de negoci.** Moltes aplicacions no han de menester aquestes velocitats tan elevades. Fer servir aquesta tecnologia en lloc d'altres amb més rodatge i menor preu només ens pot ficar en problemes.

Punts forts

Si ens centrem en els punts forts d'una base de dades en memòria RAM:

1. **És molt més ràpida.** Recordem que un accés a memòria RAM pot ser 1.000 vegades més ràpid que un accés a disc⁷.
2. **Ocupa molt menys espai.** La marca produïda per les dades és molt menor. Tampoc és necessari, per exemple, mantenir un segment de REDO. No s'ha de fer res per tal de garantir la consistència disc-memòria.
3. **Permet OLAP i OLTP a una única plataforma,** evitant haver de crear la infraestructura per a tenir la informació de l'OLAP actualitzada.
4. **ACID.** HANA, per exemple, garanteix l'atomicitat de les operacions, la seva consistència, el seu aïllament i la seva persistència.
5. **Són modernes i ja estan dissenyades per al tractament de dades.** Me crida molt l'atenció que HANA (per exemple) ja tingui un motor d'R incorporat. Ja no és necessari crear un dataset a partir de la base de dades. La pròpia base de dades és el dataset.

⁷<https://www.youtube.com/watch?v=JqzttF3rMjk>