

# Tárhely igény és tároló eszközök I/O teljesítményének méretezése a használó alkalmazások jellemzői, igényei alapján

**Téma leírása:** Az alkalmazás fejlesztők által igényelt és a ténylegesen szükséges tárterület sok esetben eltér mind méretében, mind performanciájában. Érdeemes lenne kidolgozni egy olyan méretezési módszertant, ami az ismert adatok (adatbázis típus, tranzakciók mennyisége és mérete, mentési igény stb.) alapján megbecsüli a szükséges tárterületek méretét és a szükséges performancia igényt.

Ebben a fázisban az irodalom kutatás volt az elvégzett feladat. E alapján ki kell dolgozni egy méretezési modellt a következő fázisban.

## Fogalmak

**Input/Output Operations Per Second (IOPS):** ez egy mérőszám, amely megmondja, hogy egy adott számítástechnikai tárolóeszköz mennyi műveletet tud végrehajtani másodpercenként. Például: egy HDD, SSD, vagy egy SAN.

**Logical Unit Number (LUN):** Ez egy logikai egységet jelent, amit a SCSI protokollal vagy a SAN protokollal meg lehet címezni.

**Small Computer System Interface (SCSI):** olyan szabványegyüttes, melyet számítógépek és perifériák közötti adatátvitelre terveztek.

**Storage Area Network (SAN):** blokk szintű hozzáférést biztosít merevlemezekhez, hálózaton keresztül.

**Redundant Array of Inexpensive (or Independent) Disks (RAID):** Egy tárolási technológia, mely segítségével az adatok elosztása vagy replikálása több fizikailag független merevlemezen, egy logikai lemez létrehozásával lehetséges.

**Hard Disk Drive (HDD):** mágneses alapú adattároló.

**Solid State Drive (SSD):** félvezető alapú meghajtó.

## Adattárolás

A legideálisabb az lenne, ha nem is lenne szükség IO műveletekre egy feladat elvégzéséhez. Persze ez igen valószínűtlen, de a második legjobb az, ha minél közelebb vagyunk a processzorhoz, mert annál gyorsabban el tudjuk érni a szükséges adatokat. A harmadik legjobb az, ha egyszerűen csak olcsóbban, vagy rövidebb idő alatt el tudjuk végezni a feladatot.

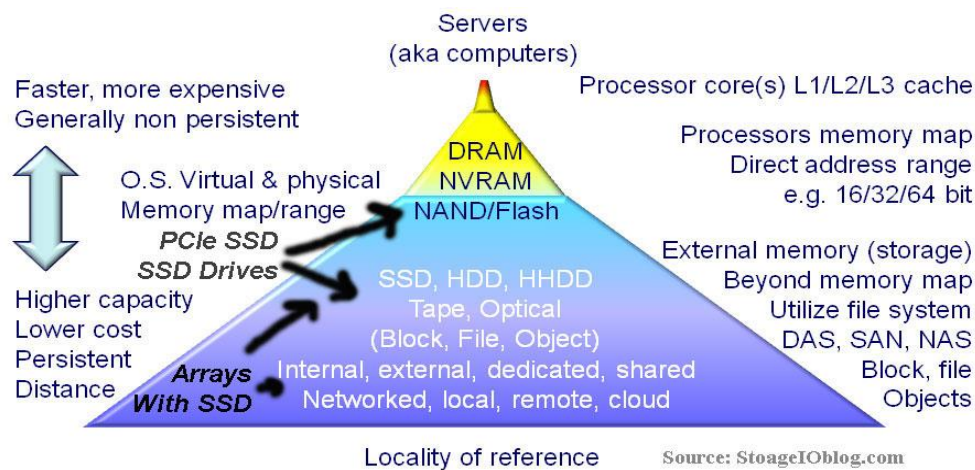


Figure 1: Memoria and tárhely stack

Az ábrán az látható, hogy minél közelebb helyezkedünk el a processzorhoz, annál gyorsabban elérhető a memóriában lévő adat, de igencsak drága és kis méretű (ez processzor gyorsítótára). Utána következik a RAM, ami már megfizethetőbb áron, nagyobb méretet szolgáltat, de lassabban érhető el. Még „távolabb” a HDD-k vannak, amelyek nagy tároló kapacitással bírnak, viszont lassan érhetőek el (a processzor órajeléhez viszonyítva). Említésre méltóak még a felhőszolgáltatások, ahol még nagyobb tárterületet vehetünk igénybe.

Az IO műveletek kihagyhatatlan részei a számítástechnikának. Egy IO művelethez szükség van processzor időre, memóriára az eredmények feldolgozásához és hálózatra az adatok továbbításához vagy fogadásához. Bár el nem hanyagolhatóak, de az erőforrásigényeiket lehet optimalizálni, javítani, például kevesebbet hajtunk végre, mert gyorsítótárból be tudunk tölteni adatokat, vagy például csoportos írás, olvasás is segítheti a helyzetet.

Talán a leglényegesebb a lokalitás, hogy a szükséges erőforrás minél közelebb legyen a felhasználóhoz. Például a processzorhoz a legközelebbi memóriaterület a regiszterek, utána jönnek az L1, L2, L3 gyorsítótárak, RAM, SSD és a többi.

## Háttértárak szűk keresztmetszete<sup>1</sup>

### A lehetséges okok

A háttértárak szűk keresztmetszete a következő okok miatt jelenkezik:

#### Cserehely (SWAP)

Amennyiben egy szerveren megfelelően sok virtuális gépet futtatunk, akkor előfordulhat az, hogy a szerver memóriája betelik és ilyenkor elkezd a háttértárakra adatokat kimenteni. A swappelés azt jelenti, hogy azokat a program adatokat, amik nem férnek el az operatív memóriában, azokat megpróbálja ideiglenesen a háttértáron elhelyezni. Ez értelemszerűen IOPS műveletekkel jár. Nem várt következményei is lehetnek a megnövekedett műveletszámoknak, például lejár a válaszra szánt várakozási idő (time out), ilyenkor akár le is fagyhat a virtuális gép.

#### Alkalmazás szintű szűk keresztmetszet

Előfordulhat, hogy kevesebb adattároló van, mint amennyi kérést fel tudnak dolgozni és mint amennyi beérkezik a rendszerbe. Az is lehet, hogy a sávszélesség kicsi és nem tudja a nagy mennyiségű adatot továbbítani, tehát a két lehetőség, ami közül választhatunk: vagy IOPS-ra optimalizálunk, vagy az átvitt adatmennyiségre. Meg kell találni az egyensúlyt a kettő között. Ha sokan használnak egy-egy alkalmazást, akkor arra is figyelni kell, hogy ne egy adattárolót használjon mindkét alkalmazás, hanem legyenek különválasztva.

#### IO válasz idő

A megnövekedett válaszügy szintén gond lehet. Feltorlódhatnak a beérkező kérések.

#### Rosszul tervezett rendszer

Legfőbb okai a következők lehetnek:

- Alacsony fordulatszámú háttértárolók használata,
- Terhelés elosztás (CPU, HDD) nem megfelelő működése, hiánya,
- Nem megfelelő RAID technológia használata magas erőforrásigénnyel,
- Kisméretű cache

### A probléma megelőzése

Érdemes a bajt megelőzni és előre gondolkodni, mintsem a gond felmerülésekor foglalkozni vele:

- A kritikus alkalmazásokhoz rendeljünk nagy erőforrásokat, hogy minimalizáljuk a problémák előfordulásának esélyét. Keressük meg a céljainknak megfelelő hardvereket, nézzünk körül a piacon. Ne legyen két kritikus alkalmazás egy lemezhez rendelve.
  - Válasszuk ki a nekünk megfelelő RAID technológiát.
  - Biztosítsunk elegendően nagy gyorsítótárat, amellyel javítható az írás és olvasás sebessége.

---

<sup>1</sup> [http://web.swcdn.net/creative/pdf/Whitepapers/Top\\_4\\_causes\\_of\\_storage\\_io\\_bottlenecks.pdf](http://web.swcdn.net/creative/pdf/Whitepapers/Top_4_causes_of_storage_io_bottlenecks.pdf)

## Monitorozás

Egy jól menedzselte rendszerben végponttól-végpontig lehet mérni a rendszer teljesítményét. A következő paramétereket mindenképp érdemes mérni:

- IOPS, írások és olvasások száma.
- Adatátvitel sebessége.
- Késleltetés hardver és logikai szinten.

### Lépések az új eszközök teljesítményének méréséhez<sup>2</sup>

1) *Meg kell tervezni, hogy hogyan és mit fogunk tesztelni.*

Például, milyen fájl rendszer lesz rajta? Engedélyezve lesz-e rajta a gyorsítótárazás? Milyen operációs rendszer fog rajta futni? Mik a fő szempontok: írás, olvasás, méret, véletlenszerű elérés, szekvenciális elérés, válaszidő, proceszor leterheltség, gyorsítótárazott, nem gyorsítótárazott. **VIGYÁZAT!** Az írás tesztelése destruktív lehet, legyen biztonsági mentés, mielőtt tesztelni kezdenék! Legyen elegendő tárterület a mérési naplók, logfájlok elmentéséhez.

2, *Válasszuk ki a szoftvert, amivel a mérést fogjuk végezni.*

3) *Ellenőrizzük a létrehozott teszt eseteket, konfigurációkat és beállításokat.*

4, *Csak a lényegre fókuszáljunk! Következő teljesítményjellemzőket érdemes mérni:*

- **IOPS** (input/output művelet per másodperc, ami az aktivitást jelenti)
- **I/O size** (átvivendő bájtok száma IO műveletenként)
- **Bandwidth** (a sávszélesség vagy adat, amit időegység alatt átmozgattunk)

Ezek között lehet prioritást felállítani. Például tranzakciós adatbázisnak nagy IOPS igénye lesz, míg egy videó megosztó oldalhoz nagy sávszélesség fog kelleni. Jó hír, hogy igazából bármelyik kettő paraméter ismeretével kiszámítható a harmadik.

$$\text{Bandwidth} = \text{IOPS} * \text{I/O size}$$

Egyéb fontosabb jellemzők:

**Latency** vagy response time (milyen gyorsan végzi el a megadott műveletet)

**Queues** (feladatok száma, amelyek a végrehajtásra várnak)

**Server system CPU** (mennyi kernelidővel jár egy IO művelet kiszolgálása)

Általában egy merevlemezen leggyakrabban az **átviteli sebességet** tüntetik fel megabájt per másodperc mértékegységben, de nem feltétlenül ez a legkritikusabb mérőszám. A második leggyakoribb mérőszám, az **IOPS**, ami azt jelenti, hogy hány írási vagy olvasási műveletet tud elvégezni egy másodperc alatt. A harmadik, a **mérete az egyes IO kéréseknek**, átlagos értéki például 4 KB, 8 KB, 32 KB, stb., de a minimum értéke az 512 bájt. Operációs rendszertől is függ, hogy hány IOPS keletkezik egy másodpercben.

$$\text{Átlagos IO méret} * \text{IOPS} = \text{átviteli sebesség MB/s}$$

---

<sup>2</sup> <https://biztechmagazine.com/article/2015/02/4-ways-performance-test-your-new-hdd-or-ssd>

Minden IO kérés időbe fog teleni, hogy befejeződjön, ezt hívják **átlagos késleltetésnek** általában miliszekundumban (ms) mérik, minél kisebb, annál jobb. Általában fizikai korlátai vannak ennek a mérőszámnak, mivel a hagyományos adattárolók mechanikusan működnek.

## Egy számítási példa

Amikor válaszidőről beszélünk érdemes látni, hogy miről is van szó. Az olvasó fej a lemez külső ívétől a belső ívéig mozoghat. Ezt hívjuk **seek time**-nak (a kisebb jelenti a jobbat). Ez időbe tellik, mint ahogy az is, hogy meg kell várni, míg a szükséges terület a korongon az olvasófej fölé / alá forog. Ez utóbbi a forgási késleltetés (**rotational delay**).

Miután a szükséges terület megérkezett az olvasófejhez, elkezdjük olvasni. Ha kis mennyiségű az IO (minimum 512 bájt) akkor az első szektor kiolvasása után már végezhetünk is, de ha nagyobb, akkor tovább tart az olvasás. Ezt nevezzük **transfer delay**-nek, ennyi ideig tart olvasni / írni.

Az ideális állapot az, hogy amit olvasni szeretnénk, az ugyanabban a sávban van és egymás utáni szektorokban („lineárisan tudjuk olvasni”). A legszerencsétlenebb az, ha állandóan új sávba kell pozícionálni a fejet és meg kell várni míg fölé / alá forog a korong. Ezért hátrányos a **töredezett partíció**.

**IO size = 4 KB, IOPS = 29600**, feltéve, hogy szekvenciálisan helyezkednek el a blokkok.  
(IO size x IOPS = Throughput )  $4096 \times 29600 = 121 \text{ MB/s}$  lesz az átviteli sebesség.

**IO size = 32 KB, IOPS = 3700**, feltéve, hogy szekvenciálisan helyezkednek el a blokkok.  
(IO size x IOPS = Throughput )  $32784 \times 3700 = 121 \text{ MB/s}$  lesz továbbra is az átviteli sebesség.  
Lehet, hogy a nagyobb IO méret miatt kevesebb IOPS-ot tudunk végrehajtani, de az adatátvitelsebesség nem romlott emiatt.

**IO size = 4 KB**, Véletlenszerű elérés esetén, **IOPS = 245**, nagyon alacsony is lehet  $4096 \times 245 = 0,96 \text{ MB/s}$ -re zuhan az átviteli sebesség. Ez javítható azáltal, hogy az IOPS-okat növeljük. Ezt csak a lemez fordulatszámának növelésével tudjuk elérni, azért jó, ha 15 ezer a fordulatszáma egy merevlemeznek. Vagy választhatunk SSD-t, amiben nincs mozgó alkatrész.