

# Benchmarking

Kritikus Rendszerek Kutatócsoport

2021

## Tartalomjegyzék

1. Alapfogalmak	2	2.3. Gráftranszformáció benchmarkok . . . . .	3
2. Példa benchmarkok*	3	<b>Irodalomjegyzék</b>	4
2.1. Hardver benchmarkok . .	3	<b>Tárgymutató</b>	4
2.2. Szoftver benchmarkok . .	3		

## Bevezetés

Melyik a legjobb ár-érték arányú processzor egy PC-be? Melyik a leggyorsabb relációs adatbázis-kezelő? Melyik kollektiókat kezelő függvénykönyvtár használja a legkevesebb memóriát? Gyakran szükséges, hogy ezekhez hasonló kérdésekre megbízható választ adjunk. A kérdések megválaszolásában általában komoly segítséget nyújt valamilyen metrika vizsgálata. Például számítsuk ki a másodpercenként elvégzett lebegőpontos utasítások számának és az árnak a hányadosát, mérjük le adott mennyiségű lekérdezéshez szükséges időt egy előre definiált adathalmazon, vagy hozzunk létre egy megadott kollektiót különböző függvénykönyvtárakkal és vizsgáljuk meg a memóriaforrásukat. Megfelelő mérések elvégzésével jó képet kaphatunk egy rendszer adott jellemzőiről.

**Megjegyzés.** Benchmarkokat abban az esetben érdemes alkalmazni, ha egy rendszer objektív jellemzőire vagyunk kíváncsiak. Szubjektív jellemzők esetén (pl. mennyire használható egy programozási nyelv, mennyire alkalmas egy adatmodell bizonyos feladatokra) a kérdés megválaszolása jóval összetettebb lehet. Ezekben az esetekben általában *felhasználói tanulmányok* () végzése célravezető. Utóbbiak elvégzése azonban rendkívül idő- és költségigényes, ezért kevés tudományos eredmény áll rendelkezésre ezekben a kérdésekben.

Fontos megemlíteni, hogy az informatika a klasszikus mérnöki tudományokkal – pl. építészet, gépészet, vegyészet – szemben még fiatal területnek számít. Ennek egyik fő jele, hogy gyakran szubjektív kérdésekről (melyik a legjobb programozási nyelv, melyik a legjobb operációs rendszer

vagy melyik a legjobb adatmodell) is véget nem érő „vallási viták” zajlanak.<sup>1</sup> Ezeket érdemes messziről elkerülni.

Az elmúlt évtizedek intenzív kutatómunkája ellenére az informatikai projektekben továbbra is kiemelkedően magas a sikertelen, elvetett (cancelled) vagy költségtervet túllépő (over budget) projektek száma [3]. Az informatikai projektek menedzsment háttere iránt érdeklődőknek javasoljuk a mára „klasszikusnak” számító műveket, mint a *Peopeware* [2] és a *The Mythical Man-Month* [1].

## 1. Alapfogalmak

A benchmarkolás elsődleges célja egy rendszer teljesítményének mérése. A kapott eredmények felhasználása többféle lehet: hasonló célú rendszerek teljesítményének összehasonlítása, egy rendszer teljesítményének felmérése, annak optimalizálása stb.

**Definíció.** A *benchmarkolás*

- egy *program* (programok, vagy más műveletek) *futtatása*,
- *szabványos tesztekkel* vagy bemenetekkel,
- egy objektum *relatív teljesítményének* felmérése érdekében.

**Megjegyzés.** A definíció eredetije az angol Wikipédia definíciója szerint [7] (kiemelések a jegyzet szerzőitől):

In *computing*, a benchmark is the *act of running* a computer program, a set of programs, or other operations, in order to *assess the relative performance* of an object, normally by running a number of *standard tests* and trials against it.

A benchmarkokkal szemben többféle elvárást támasztunk. Nyilvánvalóan nem sokat ér egy olyan benchmark, amit csak egyszer tudunk lefuttatni vagy nem tudunk később (valamilyen pontossággal) reprodukálni.

**Definíció.** *Ismételhetőség (repeatability)*: a benchmarkot lehessen egymás után többször futtatni, hogy a mérési eredmények szórása csökkenthető legyen.

**Definíció.** *Reprodukálhatóság (reproducibility)*: a benchmark legyen hasonló környezetben, hasonló eszközökkel megismételhető.

**Definíció.** *Érthetőség (comprehensibility)*: átlag felhasználó számára értelmezhető legyen az eredmény.

---

<sup>1</sup>Érdekes olvasmány a témában Luis Solano „Why Does Programming Suck?” c. cikke: <https://medium.com/@luisobo/why-does-programming-suck-6b253ebfc607>

**Definíció.** *Relevancia (relevance):* a benchmarkban megvalósított terhelési profil hasonlít arra a valós terhelésre, amely alatt a rendszer teljesítményéről információt szeretnénk kapni.

A relevancia biztosításához fontos, hogy:

- tényleg azt az alkalmazást mérjük, amit kell,
- terhelésgenerálás jellege közelítse a *valódi* terhelést, valamint
- minimalizáljuk a zavaró tényezőket, pl. megfelelően ürítjük a *gyorsítótárak* (pl. diszk cache, CPU cache) tartalmát és az operációs rendszeren futó többi felhasználói folyamatot leállítjuk.

**Megjegyzés.** Érdeklődőknek javasolt olvasmány a *Benchmarking Handbook* [4] 1. fejezete.

## 2. Példa benchmarkok\*

Többféle területen is specifikál benchmarkokat a SPEC (*Standard Performance Evaluation Corporation*)<sup>2</sup>.

### 2.1. Hardver benchmarkok

A PassMark<sup>3</sup> különböző benchmarkokat definiált, pl. processzorok teljesítményméréshez a PassMark CPU benchmarkot<sup>4</sup>.

Böngészőmotorok teljesítményméréshez: Octane<sup>5</sup>, Kraken<sup>6</sup>.

Okostelefonok teljesítményméréséhez: Antutu<sup>7</sup>.

### 2.2. Szoftver benchmarkok

Relációs adatbázis-kezelő rendszerek teljesítményének méréséhez a TPC (*Transaction Processing Performance Council*) szervezet definiált különféle benchmarkokat. Például a TPC-C<sup>8</sup> célja tranzakciós rendszerek (OLTP (*Online Transaction Processing*)) teljesítménymérése.

**Megjegyzés.** Az OLTP rendszerekkel szemben az OLAP (*Online Analytical Processing*) rendszerekben (általában) nagyobb mennyiségű adaton összetettebb, analitikus jellegű lekérdezéseket futtatnak. OLAP rendszerek teljesítménymérésére tervezték a TPC-DS benchmarkot.

<sup>2</sup><https://www.spec.org/benchmarks.html>

<sup>3</sup><https://www.passmark.com/>

<sup>4</sup><https://www.cpubenchmark.net/>

<sup>5</sup><https://developers.google.com/octane/>

<sup>6</sup><http://krakenbenchmark.mozilla.org/>

<sup>7</sup><http://www.antutu.com/en/index.shtml>

<sup>8</sup><http://www.tpc.org/tpcc/>

### **2.3. Gráftranszformáció benchmarkok**

A kutatócsoportunkban készült el az egyik első gráftranszformációk teljesítményét vizsgáló benchmark [6]. A területen azóta is aktív kutatási munkát végzünk [5].

# Hivatkozások

- [1] F.P. Brooks: *The Mythical Man-Month, Anniversary Edition: Essays On Software Engineering*. 1995, Pearson Education. ISBN 9780132119160.  
URL <https://books.google.hu/books?id=Yq35BY5Fk3gC>.
- [2] T. DeMarco–T. Lister: *Peopleware: Productive Projects and Teams*. 2013, Pearson Education. ISBN 9780133440737.  
URL <https://books.google.hu/books?id=TVQUAAAAQBAJ>.
- [3] Bent Flyvbjerg–Alexander Budzier: Why your IT project may be riskier than you think. <https://hbr.org/2011/09/why-your-it-project-may-be-riskier-than-you-think/ar>, 2011. October.
- [4] Jim Gray (szerk.): *The Benchmark Handbook for Database and Transaction Systems (2nd Edition)*. 1993, Morgan Kaufmann. ISBN 1-55860-292-5.
- [5] Gábor Szárnyas–Oszkár Semeráth–István Ráth–Dániel Varró: The TTC 2015 Train Benchmark case for incremental model validation. In Louis M. Rose–Tassilo Horn–Filip Krikava (szerk.): *Proceedings of the 8th Transformation Tool Contest, a part of the Software Technologies: Applications and Foundations (STAF 2015) federation of conferences, L’Aquila, Italy, July 24, 2015.*, CEUR Workshop Proceedings konferenciasorozat, 1524. köt. 2015, CEUR-WS.org, 129–141. p. URL <http://ceur-ws.org/Vol-1524/paper2.pdf>.
- [6] Gergely Varró–Andy Schürr–Dániel Varró: Benchmarking for graph transformation. In *Proc. IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 05)* (konferenciaanyag). Dallas, Texas, USA, 2005. September, IEEE Press, 79–88. p.
- [7] Wikipedia: Benchmark (computing). [https://en.wikipedia.org/wiki/Benchmark\\_\(computing\)](https://en.wikipedia.org/wiki/Benchmark_(computing)), 2016. április.

## Tárgymutató

érthetőség comprehensibility 2

ismételhetőség repeatability 2

TPC-C –, TPC-C 3

**OLAP** Online Analytical Processing;  
Online Analytical Processing 3

benchmarkolás benchmarking 2

**OLTP** Online Transaction Processing;  
Online Transaction Processing 3

gyorsítótár cache 3

**relevancia** relevance 2

**reprodukálhatóság** reproducibility 2

**SPEC** Standard Performance Evaluation Corporation; Standard Performance Evaluation Corporation 3

**TPC** Transaction Processing Performance Council; Transaction Processing Performance Council 3

**valódi** real-world 3