

# Teljesítmény mérése és hangolása adatbáziskezelő rendszerekben

Marton József

marton@db.bme.hu  
BME-TMIT

Adatbázisok elmélete VITMMA13  
2019. április 4.

# Miről lesz szó?

- ▶ Motivációs példa
- ▶ Benchmarkok
  - ▶ TPC
  - ▶ Linear road: a stream data management benchmark
  - ▶ LDBC SNB: gráfadatbázis benchmark
- ▶ Elméletben
- ▶ A gyakorlatban
  - ▶ Oracle Real Application Testing
  - ▶ Esettanulmány: DW gyorsítás

## Motivációs példa: TPC-B-like vs. KAISER patch

Kérdés: a Meltdown ill. Spectre sérülékenységeket javító Linux kernel patch hogyan befolyásolja a DBMS teljesítményt.

Mérési környezet

- ▶ PostgreSQL 9.6
- ▶ Debian Linux 9.x amd64, 4.9.0 kernel
- ▶ Intel Core i5-4310U, 2 mag, HyperThreading bekapcsolva  
Dell Latitude E5440, 8GB RAM, Kingston KC300 240GB SSD

Olvasnivaló: Meltdown, Spectre Attacks: Exploiting Speculative Execution, PostgreSQL 9.6 pgbench

# Benchmark: mi az?

Mérési elrendezés: jól definiált

- ▶ környezet (hálózati jellemzők, userek)
- ▶ elvárások (funkcionalitás)
- ▶ tesztelési protokoll

Alkalmas a különböző rendszerek összehasonlítására → metrikák

- ▶ mérhető: idő, pénz, lekérdezések száma
- ▶ metrika:
  - ▶ tranzakció/másodperc (tps)
  - ▶ USD / tps

## Benchmark: milyen?

Technikailag szakterület-specifikus

- ▶ többféle metrika
- ▶ ha túl általános, akkor semmire sem jó
- ▶ egy terhelési profil lényeges elemeit tartalmazó teszt

Elvárások

- ▶ releváns: legyen jelentése a területen
- ▶ érthető
- ▶ jó metrikák: lineáris, ortogonális
- ▶ skálázható: többféle architektúrán és HW-n fut, lépést tart a fejlődéssel
- ▶ lefedés: nem egyszerűsít túl
- ▶ elfogadott: szállítók és felhasználók

Forrás: SIGMOD'97: Standard Benchmarks for Database Systems



# Benchmark: előnyök, felelősség, élettartam

A jó benchmark

- ▶ megadja, hogy *mire szánják*, és *mire nem*
- ▶ ösztönzi a fejlődést: mérhető, reprodukálható
- ▶ fejlődést mutat a mérnöknek és a menedzsernek is

Abuse

- ▶ benchmarking
- ▶ marketing vs. fejlesztés

Élettartam

- ▶ egy ideig hajtja a fejlődést, később korlát lehet (“rátanulás”)

Forrás: SIGMOD'97: Standard Benchmarks for Database Systems

# Transaction Processing Performance Council (TPC)

- ▶ 1988-ban alapított
- ▶ non-profit szervezet
- ▶ tagjai a nagy ipari szállítók (HW, SW)
- ▶ adatbázis teljesítménymérésre de-facto szabvány

## Néhány TPC benchmark

- ▶ TPC-A: OLTP (end of life)
- ▶ TPC-B: terhelési teszt (end of life)
- ▶ TPC-C, TPC-E: OLTP (TPC-E eredmények)
- ▶ TPC-DI: data integration
- ▶ TPC-H, TPC-DS: DSS rendszerek (TPC-DS eredmények)
- ▶ TPC-VMS: "C+E+H+DS virtualizált környezetben"

## TPC-A: OLTP (1989-1995)

Felhasználói viselkedéseket szimuláló OLTP benchmark.

- ▶ párhuzamos kliens sessionök válaszidő korláttal
- ▶ hálózat a kliens és szerver között (LAN, WAN)
- ▶ IO intenzív (update), egyféle tranzakció
- ▶ alacsony system és app terhelés → DBMS-t mér

Mértékek

- ▶ tps (tpsA)
- ▶ price/tps, ahol price: 5 éves TCO

Forrás: TPC-A



## TPC-B: terhelési teszt (1990-1995)

Batch jellegű feldolgozáson alapuló terhelési és stabilitási teszt.

- ▶ IO intenzív (update), egyféle tranzakció
- ▶ alacsony system és app terhelés → DBMS-t mér
- ▶ elvárás: tranzakciós integritás

TPC-A tranzakciós profil és adatbázisséma.

- ▶ kis tranzakciók, de folyamatosan

## TPC-B: terhelési teszt (1990-1995)

Batch jellegű feldolgozáson alapuló terhelési és stabilitási teszt.

- ▶ IO intenzív (update), egyféle tranzakció
- ▶ alacsony system és app terhelés → DBMS-t mér
- ▶ elvárás: tranzakciós integritás

TPC-A tranzakciós profil és adatbázisséma.

- ▶ kis tranzakciók, de folyamatosan

Megbízhatósági tesztek: teljes terhelés alatt

- ▶ tápfeszültség megszüntetése utáni helyreállítás
- ▶ adat-diszk hibából helyreállítás
- ▶ napló-diszk hibából helyreállítás

Forrás: TPC-B

## TPC-B (folyt.)

Célja: teljesítménymérés és tuning

- ▶ megismételhető
- ▶ konzisztens
- ▶ terheléses teszt

Monitorozási paraméterek

- ▶ CPU load
- ▶ IO load

Tuning:

- ▶ OS paraméterek
- ▶ DBMS paraméterek (osztott memória, napló méret, checkpointing, zártáblák)

## TPC-A + TPC-B: több szem többet lát

Kérdés: hogyan lehet a maximális, nem batch Felhasználói számot elérni?

- ▶ pl. 1000 TPC-A felhasználó
- ▶ hány DBMS session legyen?

## TPC-A + TPC-B: több szem többet lát

Kérdés: hogyan lehet a maximális, nem batch Felhasználói számot elérni?

- ▶ pl. 1000 TPC-A felhasználó
- ▶ hány DBMS session legyen?
- ▶ TPC-B-vel kimérhető a párhuzamosítás optimális foka
- ▶ session multiplexer az OLTP userek és a DBMS közé
- ▶ TPC-A ezután megméri, hogy hány usert bír el

# TPC-C

## Frissebb OLTP benchmark

- ▶ fogalmi rendszer: rendelés, raktárkészlet, teljesítések, fizetés
- ▶ 5-féle tranzakció
- ▶ konkurrens végrehajtás
- ▶ válaszidő kritérium, 90%-ban
  - ▶ raktárkészlet lekérdezés: 20s
  - ▶ többi: 5s

Forrás: TPC-C

## TPC-C: ACID tesztek

A tranzakciókra ACID elvárások, ehhez vannak tesztek:

- ▶ A: commit és abort tranzakciókra is tesztel
- ▶ C
- ▶ I:
  - ▶ raktárkészlet: “read committed”
  - ▶ többire: “ANSI repeatable read” (sorosítható)
- ▶ D: tápfeszültség, memória, médiahiba

Forrás: SIGMOD'97: Standard Benchmarks for Database Systems

# TPC-E

Frissebb OLTP benchmark

- ▶ fogalmi rendszer: brókerek, piacok, ügyfelek
- ▶ 12-féle tranzakció
- ▶ konkurrens végrehajtás
- ▶ OLTP + eseményvezéreltség

Link: Aktuális TPC-E eredmények

Forrás: TPC-E



## TPC-E Advanced Sorting eredmények

System (CPU Socket/Core)	TpsE	USD/TpsE	Submit Date	DB Software Name	Total Price (USD)
Lenovo ThinkSystem SR950 Xeon Platinum 8180 2.5GHz, 4/112	11,357	98.83	2017-11-01	MS SQL 2017 EE	1 122 410
Lenovo System x3850 X6 Xeon E7-8890 v4 2.2 GHz, 4/96	9,068	139.85	2016-05-31	MS SQL 2016 EE	1 268 101
FUJITSU PRIMERGY RX4770 M3 Xeon E7-8890 v4 2.2 GHz, 4/96	8,796	116.62	2016-07-12	MS SQL 2016 EE	1 025 815
Lenovo ThinkSystem SR650 Xeon Platinum 8280 2.7GHz, 2/56	7,012	90.99	2019-03-29	MS SQL 2017 EE	638 052
Lenovo ThinkSystem SR650 Xeon Platinum 8180 2.5GHz, 2/56	6,779	92.49	2018-09-05	MS SQL 2017 EE	626 980
FUJITSU PRIMERGY RX2540 M4 Xeon Platinum 8180 2.5GHz, 2/56	6,606	92.85	2018-03-31	MS SQL 2017 EE	613 391
Lenovo ThinkSystem SR650 Xeon Platinum 8180 2.5GHz, 2/56	6,598	93.48	2017-06-27	MS SQL 2017 EE	616 777

Forrás: TPC-E eredmények, 2019-04-04

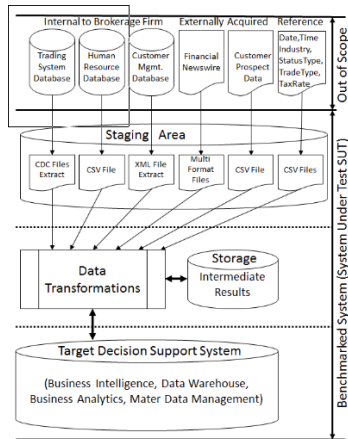
# TPC-DS eredmények

System	SF	Performance (QphDS)	Price/QphDS	Submit Date	DB SW	Total
Alibaba Cloud E-MapReduce (CentOS 7.4)	10	1,824,283	0.31 USD	2019-03-19	Alibaba Cloud E-MapReduce 3.16.1	559 53
Cisco UCS Integrated Infrastructure for Big Data (RHEL 6.7)	10	1,580,649	0.64 USD	2018-03-05	Transwarp Data Hub v5.1	1 008

Forrás: TPC-DS eredmények, 2019-04-04

# TPC-DI: Data integration (ETL)

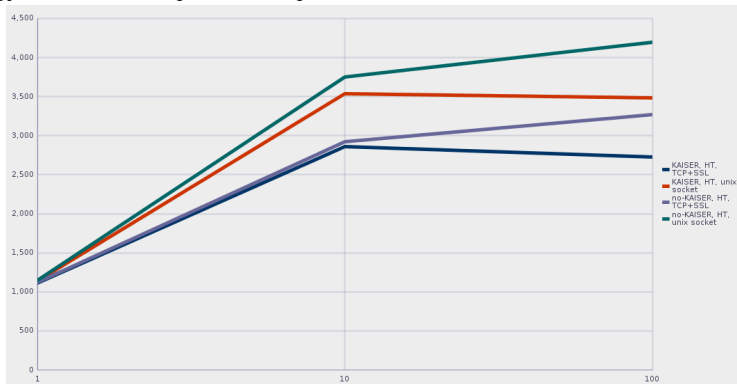
- ▶ nagy adatmennyiségek
- ▶ transzformációk: hibellenőrzés, szurrogátumkulcs-lookup, típuskonverzió, aggregáció, adatmódosítás
- ▶ előzmény és inkrementális töltés
- ▶ konzisztencia elvárások
- ▶ 5, heterogén adatforrás (relációs adatbázis, XML, CSV ...)
- ▶ sok tábla, gazdag kapcsolatrendszer



Forrás: TPC-DI, ábra és egyebek: VLDB 2014: TPC-DI

# KAISER patch mérés újra (TPC-B-like)

Vigyázat! Gondoljunk bele jobban: SF 1 vs. 10



x tengely: "scale factor", y tengely: másodpercenkénti tranzakciószám (tps)

Link: [pgbench](#)

## További benchmarkok

Linear Road Benchmark<sup>1</sup>: streaming benchmark

- ▶ folyamatos lekérdezésekre teljesítménymetriák
- ▶ OLTP és analitikus jellegű lekérdezések egyvelege
- ▶ egy érdekes kísérlet valósidejű információs rendszer benchmarkként

LDBC SNB<sup>2</sup>: gráfadatbázis benchmark

- ▶ LDBC Social Network Benchmark
- ▶ interaktív és BI profil, “chokepoint design”
- ▶ adatgenerátor: szintetikus, de realisztikus közösségi háló
- ▶ skálázási faktor

---

<sup>1</sup> Arasu et. al. Linear road: a stream data management benchmark. In: VLDB 2004

<sup>2</sup> Erling et. al. The LDBC Social Network Benchmark: Interactive Workload. In: SIGMOD 2015

# Oracle Real Application Testing

Probléma: szép az XY benchmark, de nem az én profilomat méri

# Oracle Real Application Testing

Probléma: szép az XY benchmark, de nem az én profilomat méri

Egy megoldás: Oracle Database Real Application Testing opció

- ▶ cél: pathcek, HW/SW upgradek hatásának mérése
- ▶ valódi, saját terhelésen
- ▶ két része
  - ▶ SQL Replay: adatbázis terhelés felvétele, és visszajátszása
  - ▶ SQL Performance Analyzer: riport, ami a végrehajtási tervek változását mutatja

Link: [OTN: Real Application Testing](#)

# Oracle-alapú DW teljesítménytuning esettanulmány

Néhány szempont, trükk

- ▶ HW van, de nagymértékben soros a DW töltés szervezése
- ▶ inter-query párhuzamosítás: dependency gráf az objektumok és folyamatok között
- ▶ intra-query párhuzamosítás
- ▶ 12c: adaptív végrehajtási terv
- ▶ szurrogátumkulcsok: szekvencia cache
- ▶ idegen kulcsok letiltása: ún. *direct path* insert
- ▶ adattáblák tömörítése: mivel CPU intenzív, csak a végeredménynél
- ▶ NOLOGGING vs. redo logok
- ▶ memóriatuning: SGA, PGA vs. TEMP táblatér, SGA poolok