

Elosztott relációs adatbázis kezelő rendszerek vizsgálata cloud környezetben

Budai Péter

Dr. Goldschmidt Balázs

*Budapesti Műszaki és Gazdaságtudományi Egyetem
Irányítástechnika és Informatika Tanszék*

Tartalom

- Adattárolás a felhőben
- Vizsgálati terület
- MySQL Cluster ismertető
- MySQL Cluster teljesítményvizsgálata
- Összefoglalás, további célok
- Kérdések

Adattárolás a felhőben

- Cloud computing
 - Számítási erőforrás, *tárolási kapacitás* és különböző szoftverek egyikének vagy mindegyikének szolgáltatásként való megjelenítése
- Adattárolás
 - Dedikált szolgáltatások
 - Saját megoldások

Adattárolás a felhőben

- Dedikált szolgáltatások
 - Blob
 - Nagyméretű bináris objektumok
 - Fastruktúra
 - Table, NoSQL
 - Sok adat, minimálisan strukturált
 - Limitált lekérdezési lehetőségek
 - Relációs adatbázis
 - Erős struktúra
 - Változatos lekérdezések

Adattárolás a felhőben

- Saját megoldások
 - Hagyományos adatbázis szoftverek futtatása virtuális kiszolgálókon
 - Pl. (elosztott) relációs adatbázisok
 - Előnyök:
 - Több funkció, jobban integrálható
 - Meglévő alkalmazás gyorsabban migrálható
 - Hátrányok:
 - Karban kell tartani
 - Monitorozás
 - Backup

Vizsgálati terület

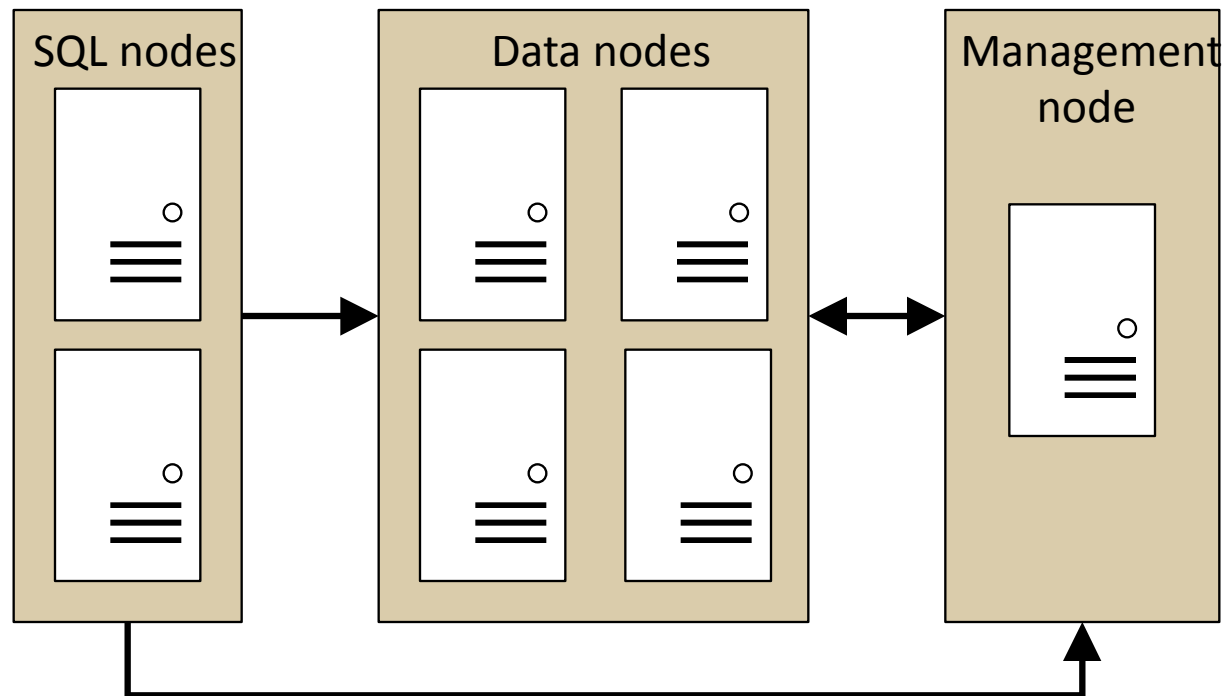
- Elosztott relációs adatbázis kezelő rendszerek (DRDBMS)
 - Nagyobb hibátűrés
 - Felhőben könnyebben menedzselhető
 - Kiszámíthatatlan környezet
 - Skálázható
- Tulajdonságok:
 - Homogén – heterogén
 - Replikáció – partícionálás
- MySQL Cluster

MySQL Cluster

- Nyílt forráskódú
 - MySQL különálló fejlesztési ága
- Homogén, elosztott és redundáns
 - Nincs egyszeres hibahely
- Memória adatbázis
 - Hatékonyabb
 - Kisebb kapacitás
 - Háttértár csak szükség esetén

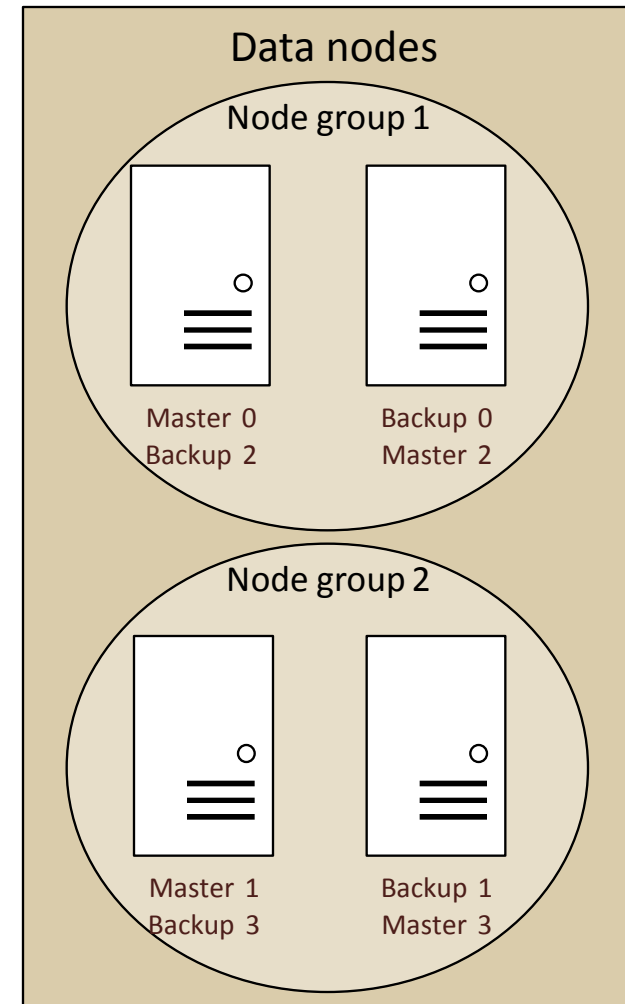
MySQL Cluster

- Felépítés



MySQL Cluster

- Redundancia és partícionálás
 - Automatikusan
- Megválasztható:
 - Redundancia mérték
 - Adatkiszolgálók száma
 - Partícionálás mértéke

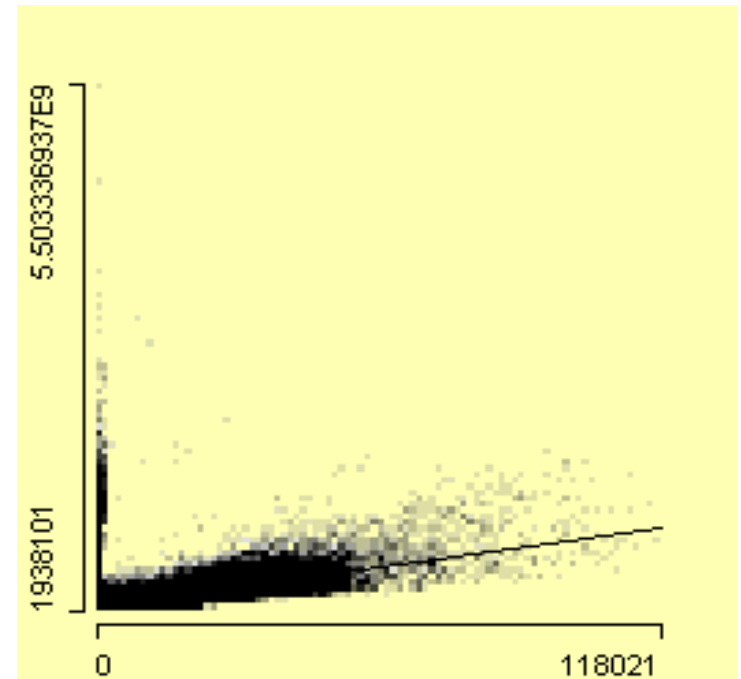


Skálázódási vizsgálat

- Lekérdezések futási ideje
- 17 különböző konfiguráció
 - Redundancia
 - Adatkiszolgálók száma
 - SQL kiszolgálók száma
- 7 különböző típusú lekérdezés
 - Paraméterezhetően
- Adatbázis méret

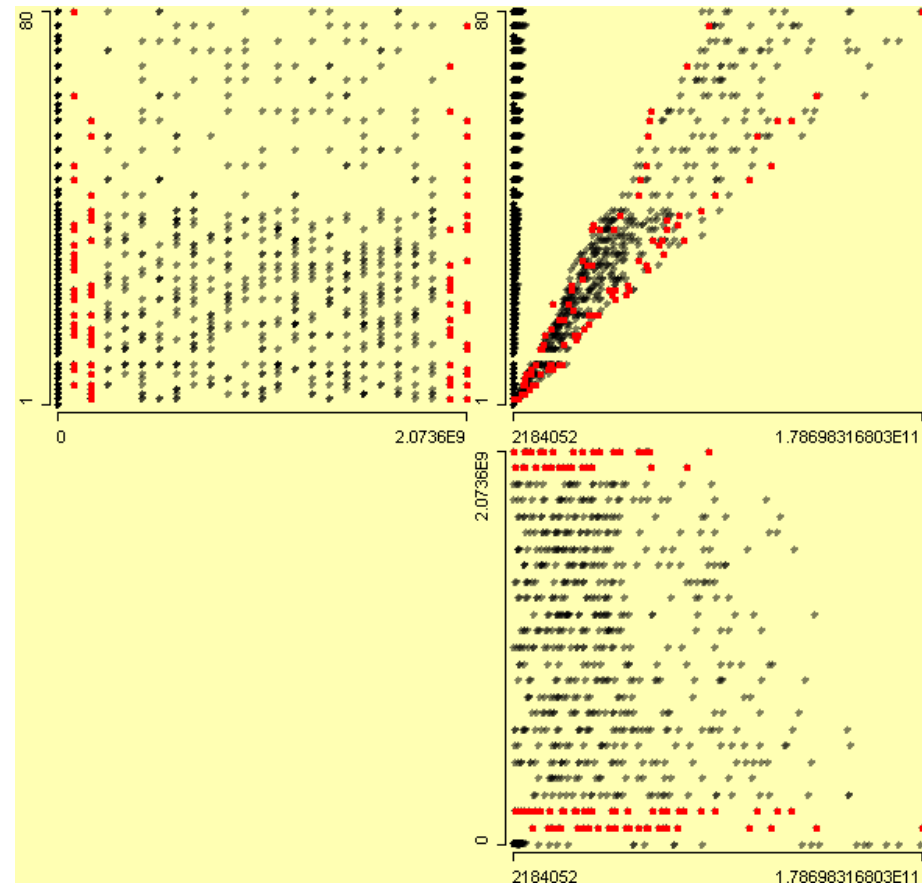
Skálázódási vizsgálat

- Kb. 350 000 adatpont
- Statisztikai vizsgálat
 - 1.93 ms – 2 s: 99 %
 - 2 s – 70,6 s: 1 %
- Többségi adatsor vizsgálata:
 - Skálázódás nincs hatással
 - Csak az eredményhalmaz mérete
 - Hálózati átvitel!



Skálázódási vizsgálat

- Kiugró értékek vizsgálata
 - Egy bizonyos lekérdezés
 - Nem indexelt oszlopra szűrés
 - Korreláció:
 - Adatbázis méret
 - Lekérdezés paraméterek
 - Skálázódástól független



Összefoglalás

- Adattárolási lehetőségek a felhőben
 - Dedikált megoldások
 - Saját megvalósítás
 - Elosztott adatbázis üzemeltetés
- MySQL Cluster
 - Hibatűrő, redundáns
 - Jól skálázódik
 - Elsődleges kulcsra optimalizálva

Köszönöm a figyelmet

- Kérdések
- Kutatásunk az "Új tehetséggyondozó programok és kutatások a Műegyetem tudományos műhelyeiben" című projekt része
 - A projekt megvalósítását a TÁMOP - 4.2.2.B-10/1--2010-0009 program támogatja.