

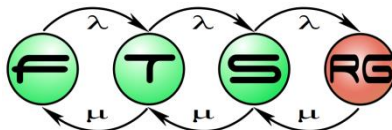
NoSQL RENDSZEREK

Szárnnyas Gábor

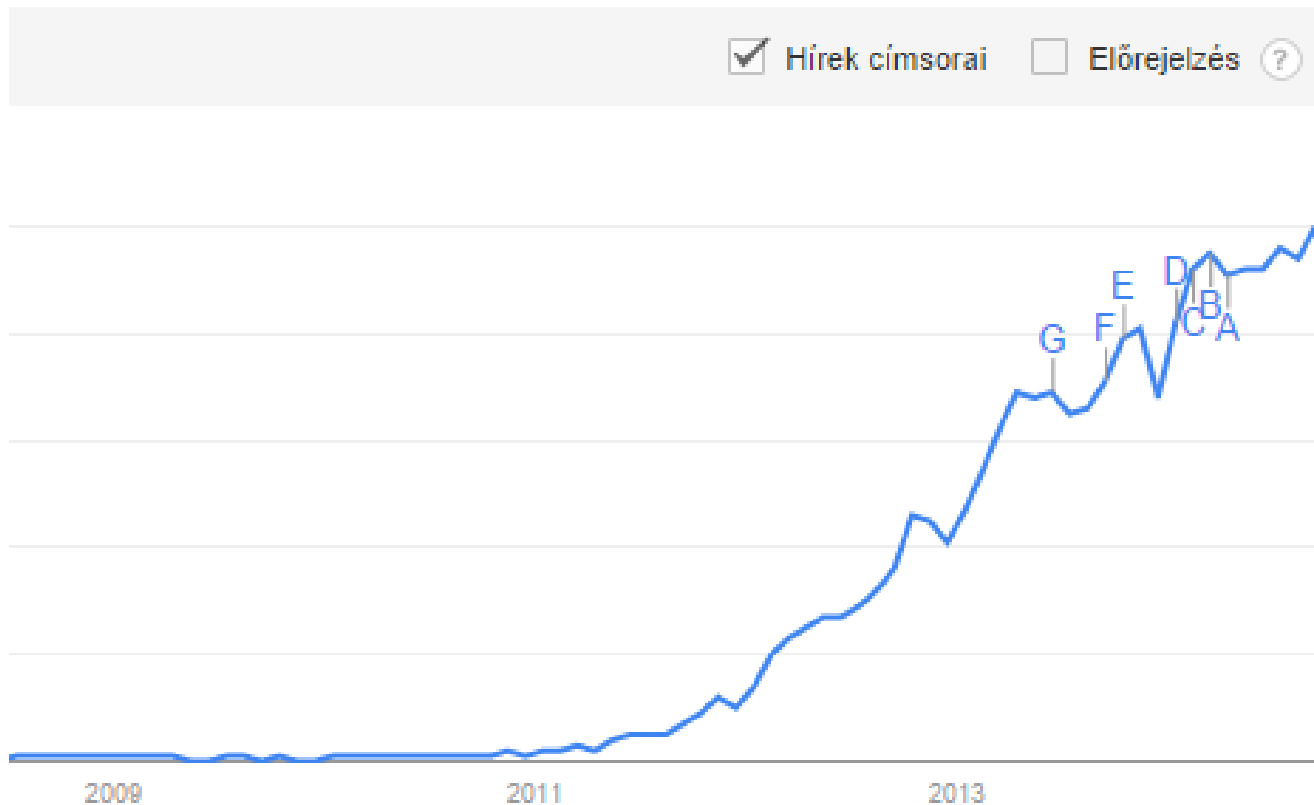
szarnnyas@mit.bme.hu

2014. december 3.

Budapesti Műszaki és Gazdaságtudományi Egyetem
Hibatűrő Rendszerek Kutatócsoport

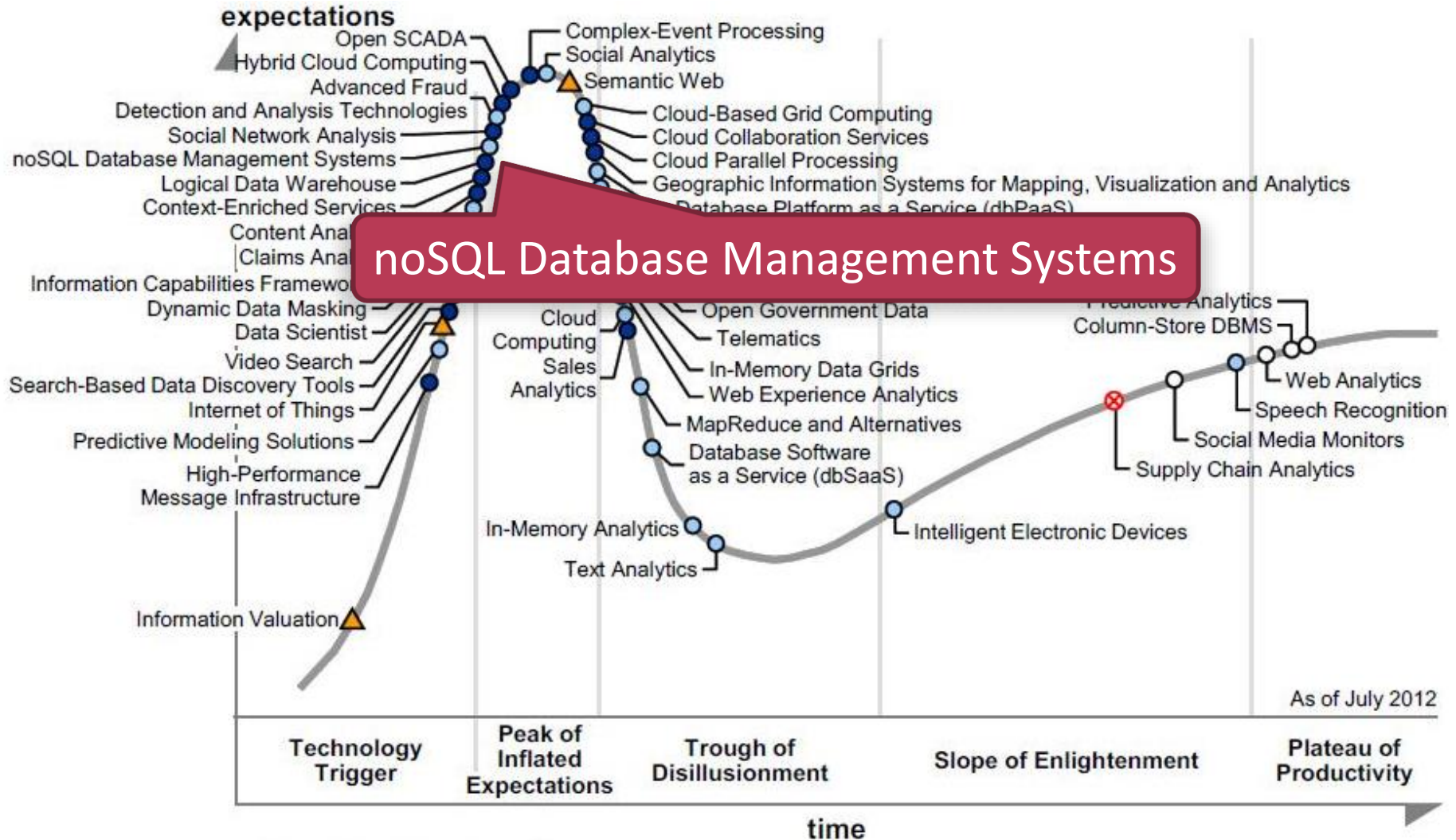


Google Trends: „Big Data”



- N.B. ez is egy Big Data feladat
- Gartner hype cycle: HF

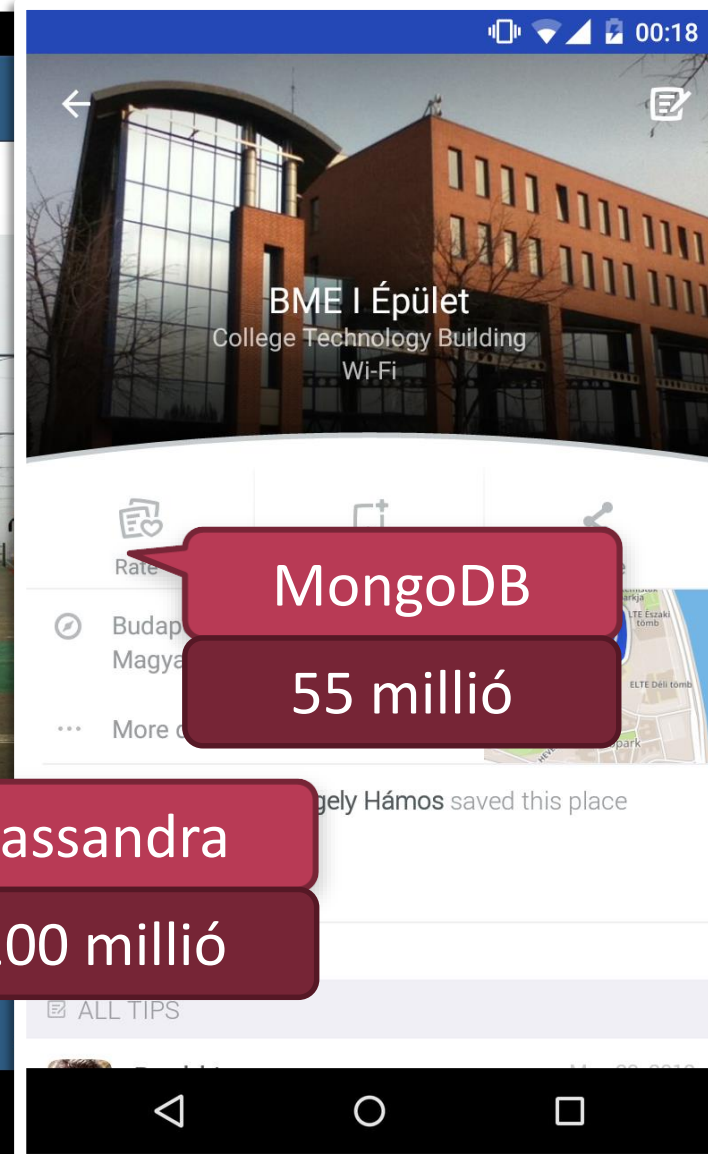
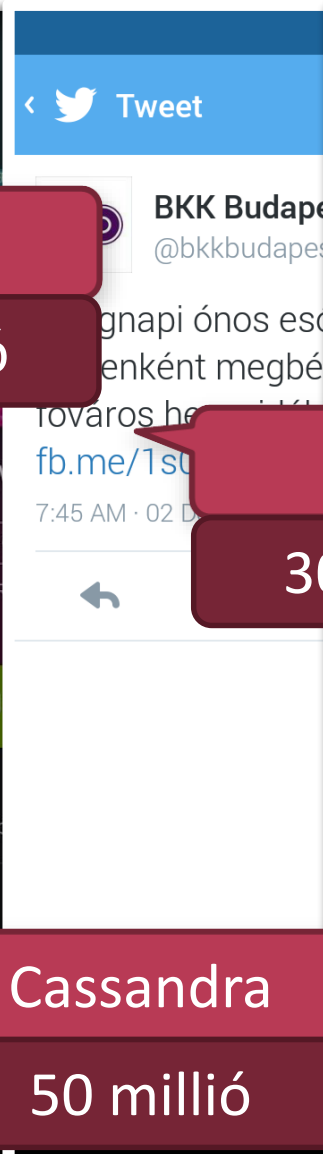
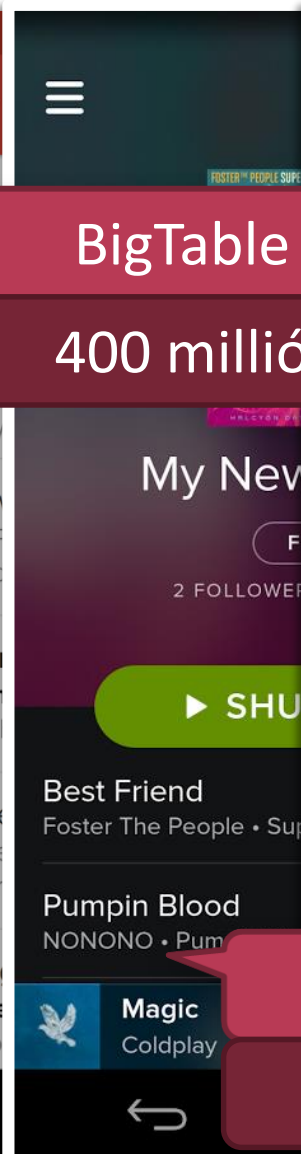
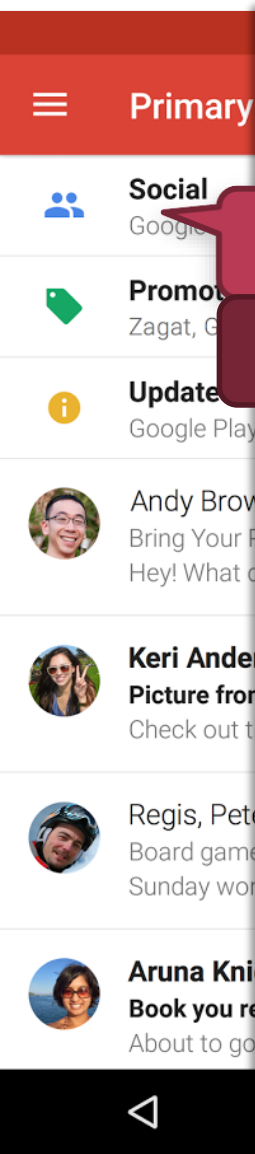
Gartner: Hype Cycle for Big Data – 2012



DB-Engines Ranking – 2014

Rank	Last Month	DBMS	Database Model	Score	Changes
1.	1.	Oracle	Relational DBMS	1459.79	+7.67
2.	2.	MySQL	Relational DBMS	1268.58	-10.50
3.	3.	Microsoft SQL Server	Relational DBMS	1200.05	-20.15
4.	4.	PostgreSQL	Relational DBMS	254.01	-3.35
5.	5.	MongoDB	Document store	246.52	+1.78
6.	6.	DB2	Relational DBMS	210.25	+4.02
7.	7.	Microsoft Access	Relational DBMS	139.89	+1.06
8.	8.	SQLite	Relational DBMS	94.70	-0.58
9.	9.	Cassandra	Wide column store	94.06	+2.07
10.	↑	Redis	Key-value store	87.88	+5.53

http://db-engines.com/en/ranking_trend



BigTable

400 millió

Redis

300 millió

MongoDB

55 millió

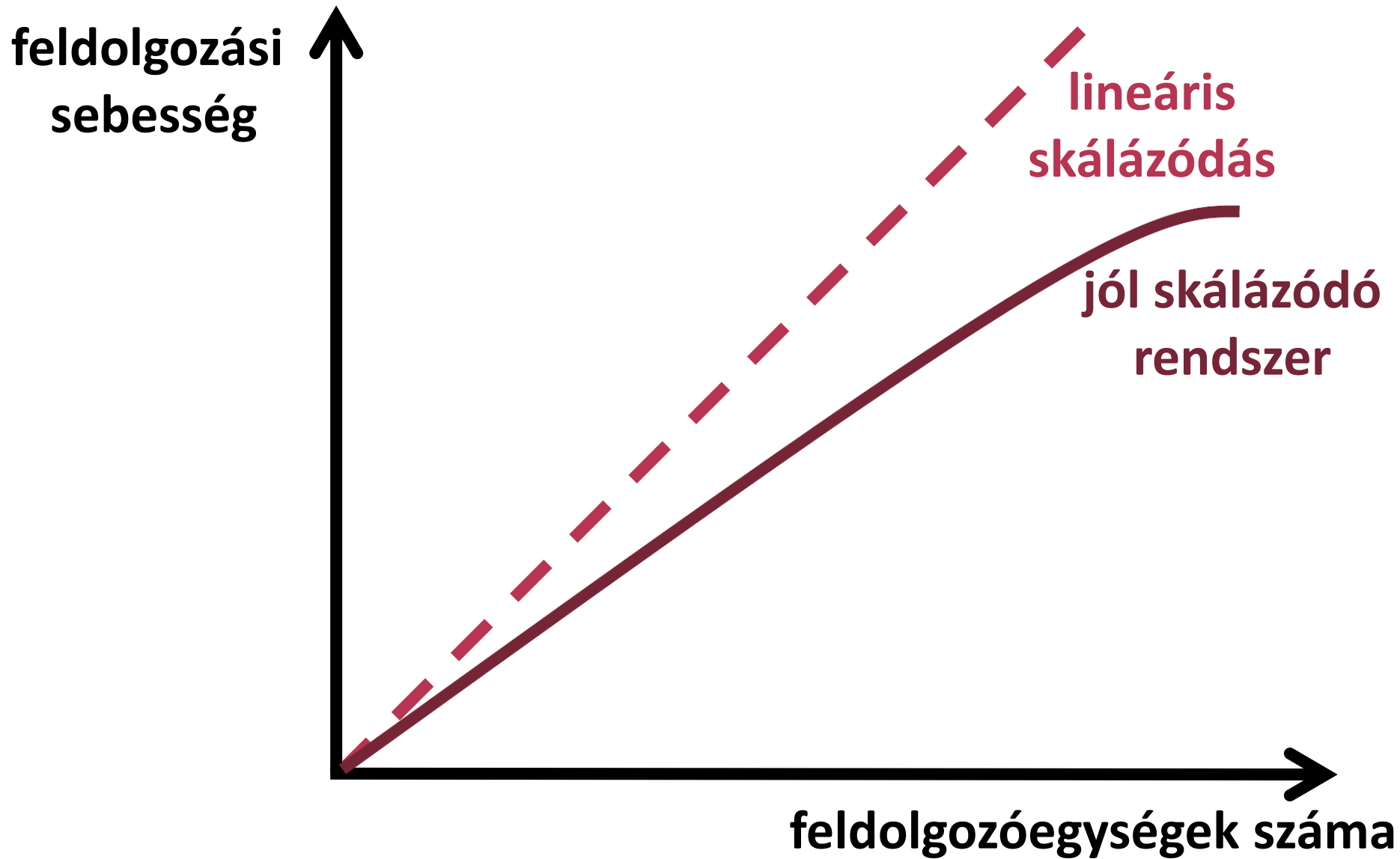
Cassandra

200 millió

Cassandra

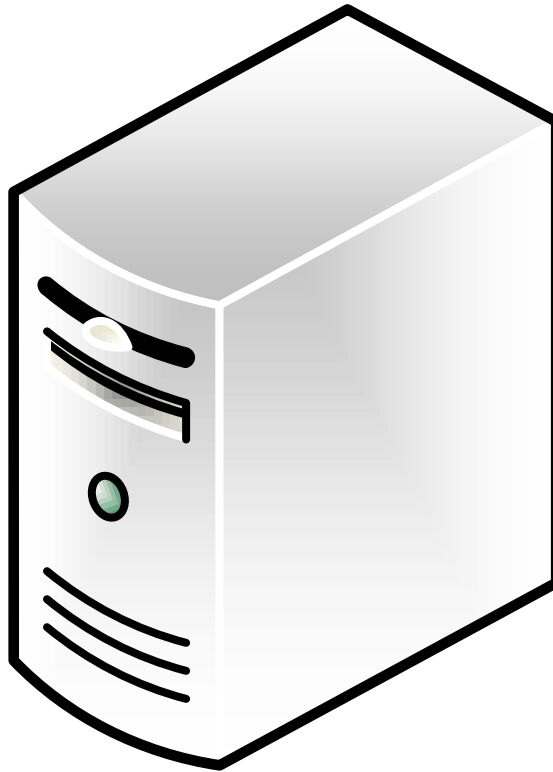
50 millió

Skálázhatóság



Skálázódás

- Vertikális skálázódás



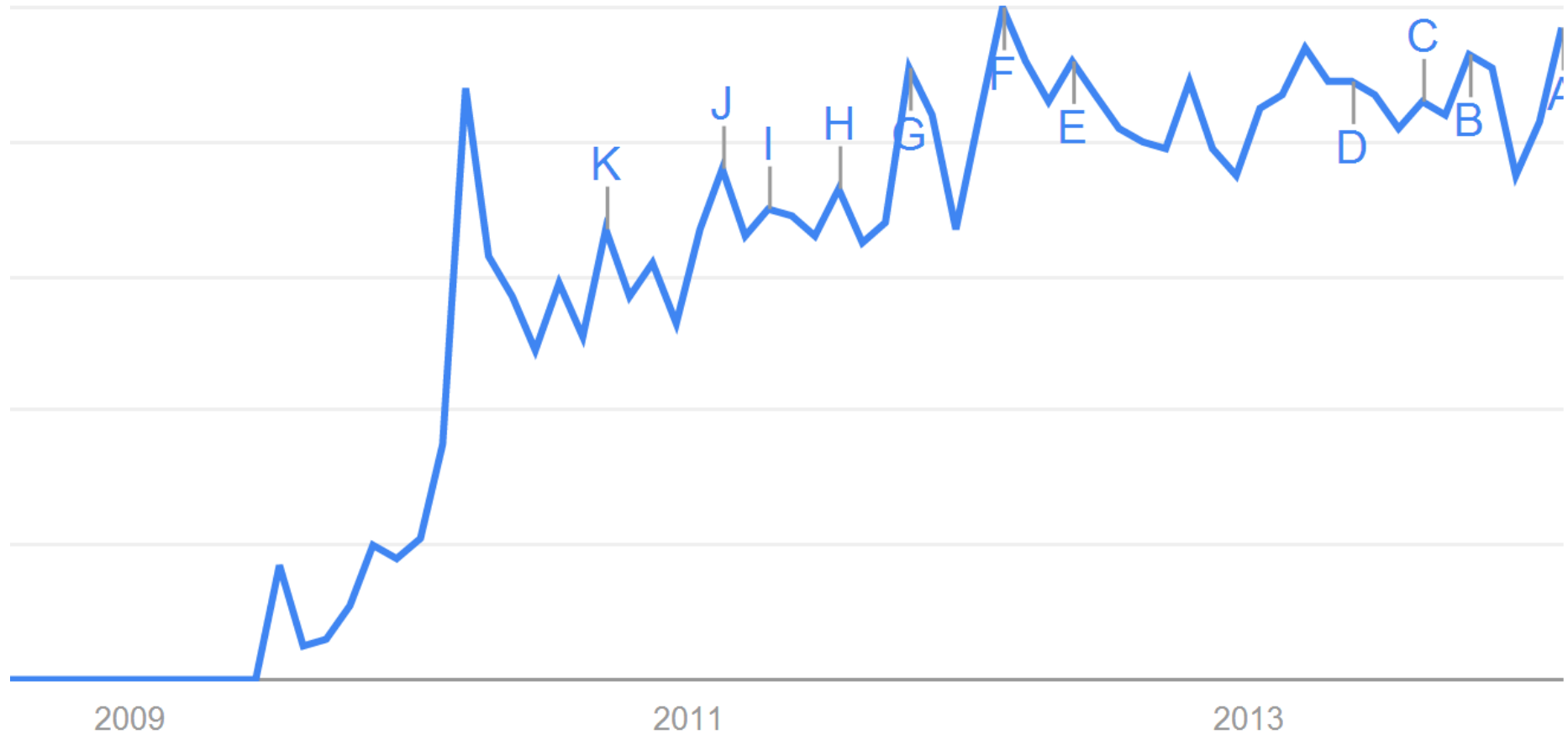
Skálázódás

- Vertikális skálázódás
- Horizontális skálázódás
 - Commercial off-the-shelf (COTS) hardver

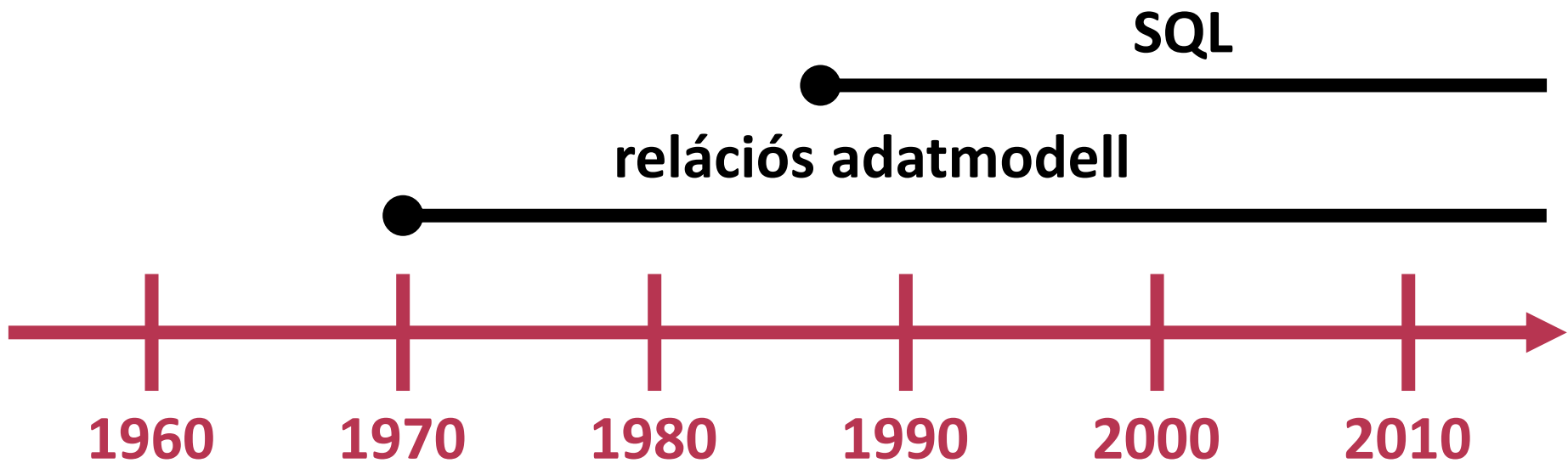


NOSQL RENDSZEREK

Google Trends – NoSQL



Relációs rendszerek



Relációs rendszerek

Rank	Last Month	DBMS	Database Model	Score	Changes
1.	1.	Oracle	Relational DBMS	1459.79	+7.67
2.	2.	MySQL	Relational DBMS	1268.58	-10.50
3.	3.	Microsoft SQL Server	Relational DBMS	1200.05	-20.15
4.	4.	PostgreSQL	Relational DBMS	254.01	-3.35
5.	5.	MongoDB	Document store	246.52	+1.78
6.	6.	DB2	Relational DBMS	210.25	+4.02
7.	7.	Microsoft Access	Relational DBMS	139.89	+1.06
8.	8.	SQLite	Relational DBMS	94.70	-0.58
9.	9.	Cassandra	Wide column store	94.06	+2.07
10.	↑	11. Redis	Key-value store	87.88	+5.53

http://db-engines.com/en/ranking_trend

Dominancia

- Relációs modellben gondolkozunk
- Legyőzte a hálós modellt a '70-es években
- Legyőzte az objektum modellt a '90-es években

„Az SQL adatbázis tovább él, mint az
elérésre használt programozási nyelv”

„Az SQL adatbázis tovább él, mint az azt
futtató operációs rendszer”

Reneszánsz

- 2005 előtt: RDBMS/SQL legtöbb weboldal mögött
- 2005 környéke: Google cikkek
 - Ghemawat, Gobioff, Leung: *Google File System*
 - Chang, Dean, Ghemawat, et al.: *Bigtable*
 - Mike Burrows: *Chubby*
 - Chandra, Griesemer, Redstone: *Paxos made live*
 - Dean, Ghemawat: *MapReduce*
- DeCandia et al., Amazon: *Dynamo*
- Facebook: *Cassandra*

HDFS

Hadoop

Google cikkek

- Nincs forradalmi algoritmus vagy adatszerkezet
- Létező ötletek újfajta alkalmazása
- Rendszertervezésről szóló cikkek
- **Új, elosztott architektúra**
- **Nem relációs modell, nincs SQL**

NoSQL „forradalom”

- Google, Amazon technológiák
 - nem nyílt forráskódúak
- 2005–2010: új nyílt forráskódú NoSQL technológiák a Google és az Amazon ötletei alapján
- Maximum: 2009, hetente új projektek

NoSQL definíció

- Nyílt forráskódú
- Nem-relációs
- Elosztott
- Horizontálisan skálázható
- Sémamentes adatmodell
- Replikáció támogatása
- Fokozatos konzisztencia

NoSQL definíció

- Nyílt forráskódú
- Nem-relációs
- **Elosztott**
- **Horizontálisan skálázható**
- Sémamentes adatmodell
- Replikáció támogatása
- Fokozatos konzisztencia

SKÁLÁZHATÓSÁGI KIHIVÁSOK

Szerverek száma – 2010



Google: 1 000 000+

Intel: 100 000

Facebook: 30 000

COMPANIES THAT DON'T DISCLOSE NUMBERS BUT ARE ESTIMATED TO HAVE OVER 50,000:

● MICROSOFT

● EBAY

● HP/EDS

● GODADDY

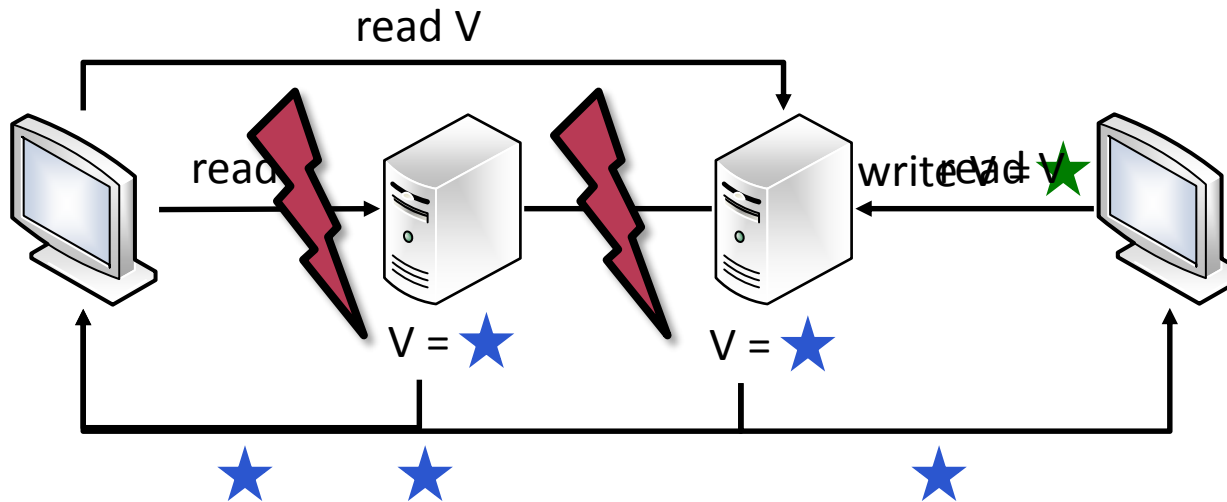
● AMAZON

● YAHOO

● IBM

Replikáció

■ Többpéldányos tárolás



Elméleti korlát: a CAP tétel

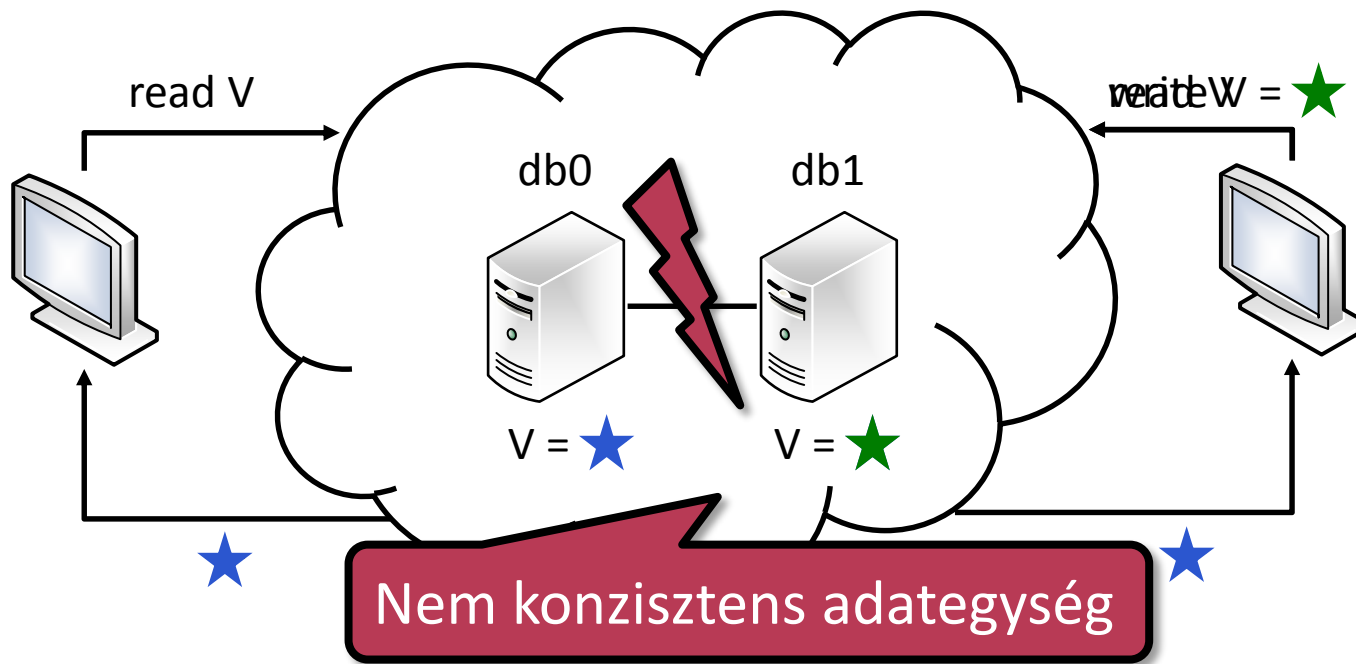
- Sejtés: Eric Brewer, 2000
- Tétel: Nancy Lynch, Seth Gilbert, 2002
- Tulajdonságok:
 - **C**onsistency
 - **A**vailability
 - **P**artition tolerance
- „Elosztott rendszerben egy időben nem garantálható mindhárom tulajdonság.”

CAP tétel precízebben

- Egy elosztott rendszerben nem biztosítható, hogy a rendszer *mindig* (üzenetek elvesztése esetén is) garantálja az alábbi tulajdonságokat:
 - atomi konzisztencia (**c**onsistency),
 - rendelkezésre állás (**a**vailability).

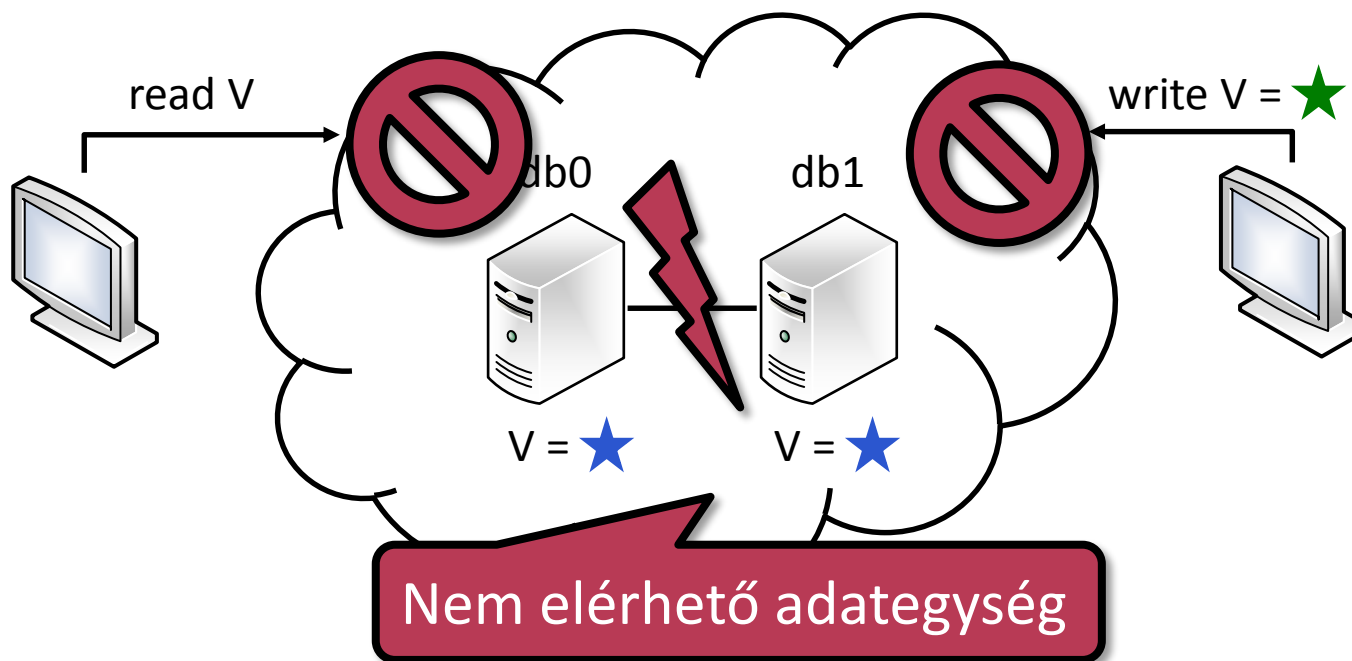
Consistency – konzisztencia

- Egy adataegység értékét bármely csomóponttól lekérdezve ugyanazt az értéket kapjuk.



Availability – rendelkezésre állás

- A rendszer minden működő csomóponthoz érkező kérésre válaszol



CAP tétel precízebben (ism.)

- Egy elosztott rendszerben nem biztosítható, hogy a rendszer *mindig* (üzenetek elvesztése esetén is) garantálja az alábbi tulajdonságokat:
 - atomi konzisztencia (**c**onsistency),
 - rendelkezésre állás (**a**vailability).

A CAP tétel kritikája

- Egyszeres hibapontok (single point of failure, SPOF)
- Adatbázis-kezelő összeomlása
- Alkalmazáshibák
- Teljesítmény figyelmen kívül hagyása
 - Áteresztőképesség [adategység/s]
 - Késleltetés [s]

A késleltetés ára



100 ms késleltetés

1% csökkenés az eladásokban

A késleltetés ára



500 ms késleltetés
20% bevételcsökkenés

Konzisztenciamodellek

- A CAP tétel következménye

gyenge konzisztencia

Gyakori kompromisszum

fokozatos konzisztencia

erős konzisztencia

erős konzisztencia több adategységen

Demó

- Fokozatos konzisztencia a gyakorlatban: **Tumblr**
<http://eventualconsistency.tumblr.com/>

Tranzakciók – ACID garanciák

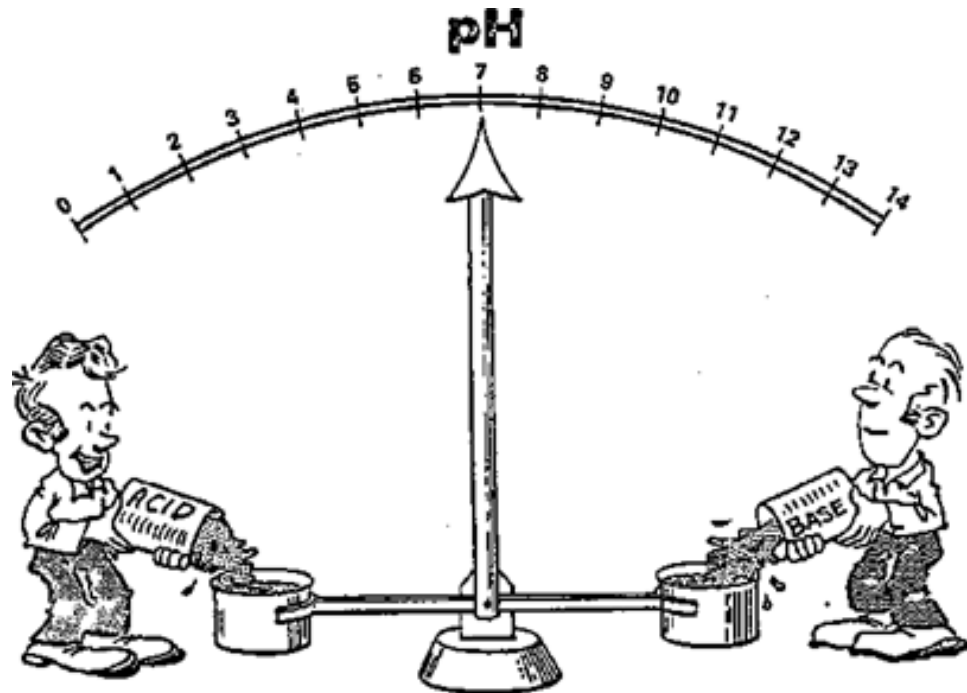
- Atomicity
- Consistency
- Isolation
- Durability



Ld. [http://en.wikipedia.org/wiki/Acid test \(gold\)](http://en.wikipedia.org/wiki/Acid_test_(gold))

ACID–BASE

- Atomicity
- Consistency
- Isolation
- Durability
- Basically Available
- Soft state
- Eventually consistent



NOSQL ADATBÁZISOK TÍPUSAI

NoSQL definíció

- Nyílt forráskódú
- **Nem-relációs**
- Elosztott
- Horizontálisan skálázható
- Sémamentes adatmodell
- Replikáció támogatása
- Fokozatos konzisztencia

NoSQL adatbázisok típusai

- Kulcs-érték tárolók
- Dokumentumtárolók
- Oszlopcsaládok
- Gráfadatbázisok

NoSQL adatbázisok típusai

- **Kulcs-érték tárolók**
- Dokumentumtárolók
- Oszlopcsaládok
- Gráfadatbázisok

Kulcs-érték tárolók

- Nagyon egyszerű programozási felület:
 - `get(key)`
 - `put(key, value)`
 - `delete(key)`

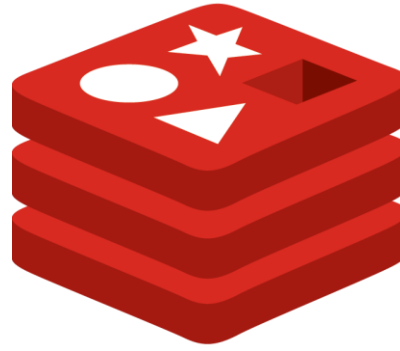
Kulcs-érték tárolók felhasználása

- Munkamenetek tárolása
- Egyszerű felhasználói profilok
- Képek
- Gyorsítótárazás (cache)

Kulcs-érték tároló implementációk

ORACLE®

BERKELEY DB



redis

32 millió dollár

ORACLE®

**NOSQL
DATABASE**



riak



NoSQL adatbázisok típusai

- Kulcs-érték tárolók
- **Dokumentumtárolók**
- Oszlopcsaládok
- Gráfadatbázisok

Dokumentumtárólok

- Szemistrukturált adatok
- Nincs előre definiált séma
- Nincs illesztés művelet

Dokumentumtárolók – JSON

```
{ "document": [  
  {  
    "firstname": "Klemens",  
    "city": "Stuttgart",  
    "age": "42"  
  },  
  {  
    "firstname": "Rajesh",  
    "city": "Delhi",  
    "age": "29"  
  },  
  {  
    "firstname": "Colin",  
    "company": "Oracle"  
  },  
  {  
    "cars": ["BMW 320d", "Jaguar XF"]  
  }  
]}
```

szemistrukturált formátum

Dokumentumtárolók

- Naplózás
- Tartalomkezelő rendszerek (CMS)
- Valós idejű adatelemzés

Dokumentumtároló implementációk



mongoDB



CouchDB
relax



Couchbase

NoSQL adatbázisok típusai

- Kulcs-érték tárolók
- Dokumentumtárolók
- **Oszlopcsaládok**
- Gráfadatbázisok

Oszlopcsaládok

- Táblák

- Sorok = kulcs-érték párok

oszlopkulcs 1	oszlopkulcs 2	...
érték 1	érték 2	...

- Nincs illesztés művelet

- Előre definiált, de rugalmas séma

- Például:

- Memória
- Kijelző átmérője
- Szín
- Mobiltelefon szolgáltató
- stb.



Oszlopcsalád implementációk



cassandra

A P A C H E
HBASE

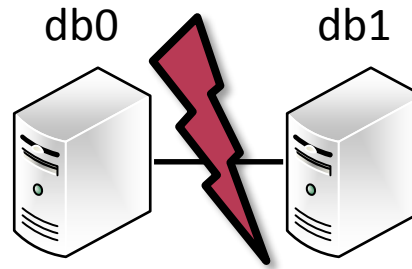
Oszloptárolók használata

- Dokumentumtárolókhoz hasonló
 - Naplózás
 - Tartalomkezelő rendszerek (CMS)
 - Kötegelt adatfeldolgozás (MapReduce)



Vásárlói kosár

name	product
John Doe	Mac Pro



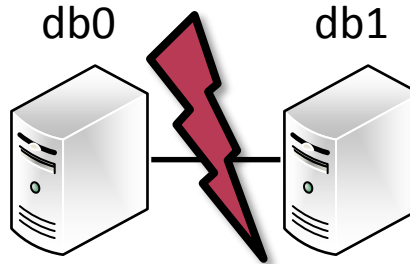
name	product
John Doe	iPad Air



Vásárlói kosár

name	product	color
Jane Doe	iPhone 5s	gold

name	product	color
Jane Doe	iPhone 5s	space gray



NoSQL adatbázisok típusai

- Kulcs-érték tárolók
- Dokumentumtárolók
- Oszlopcsaládok
- **Gráfadatbázisok**

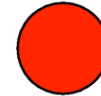
Gráf adatmodell

$$G = (V, E)$$

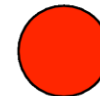
"textbook graph"

Forrás: Titan: The Rise of Big Graph Data, <http://www.slideshare.net/slidarko/titan-the-rise-of-big-graph-data>

Gráf adatmodell

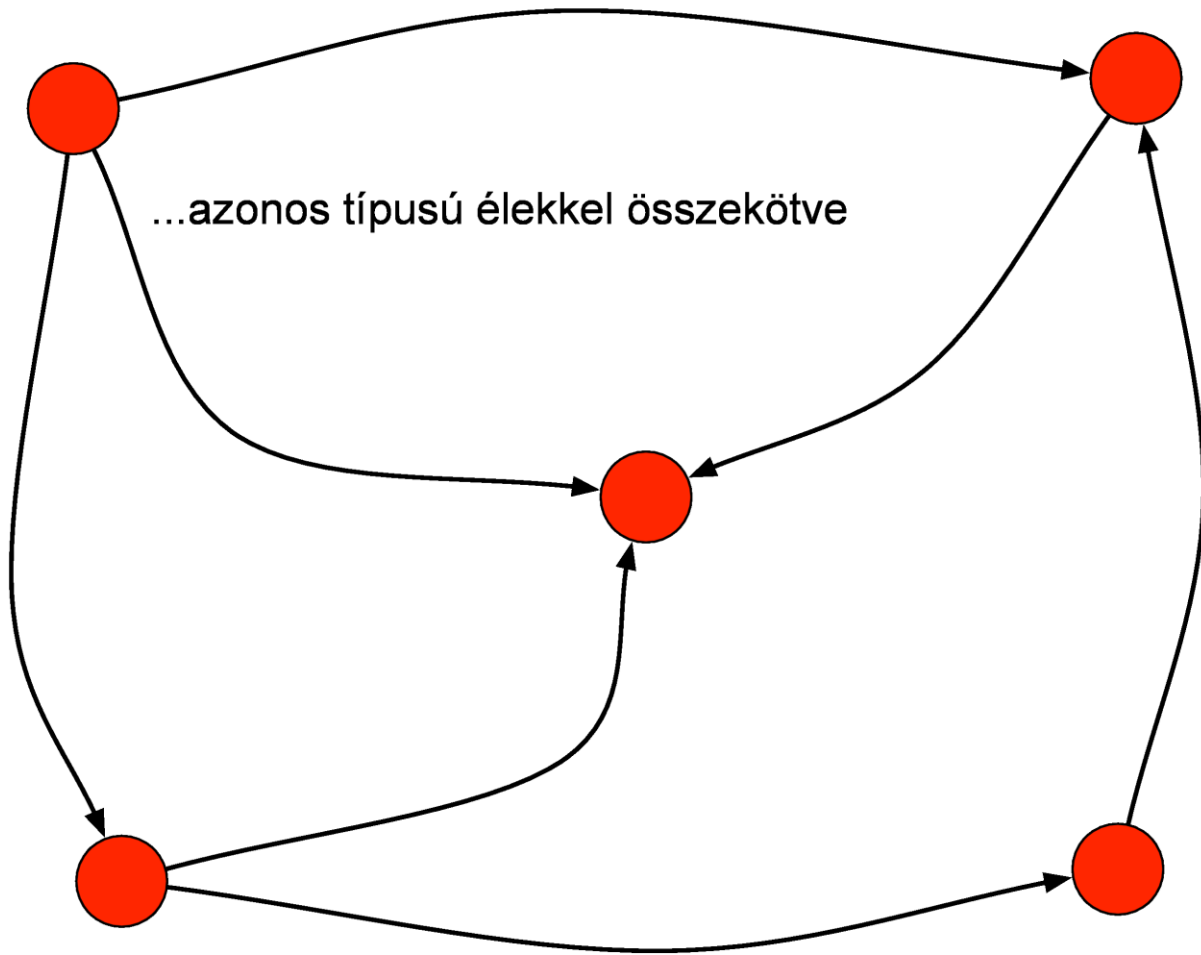


Azonos típusú csomópontok



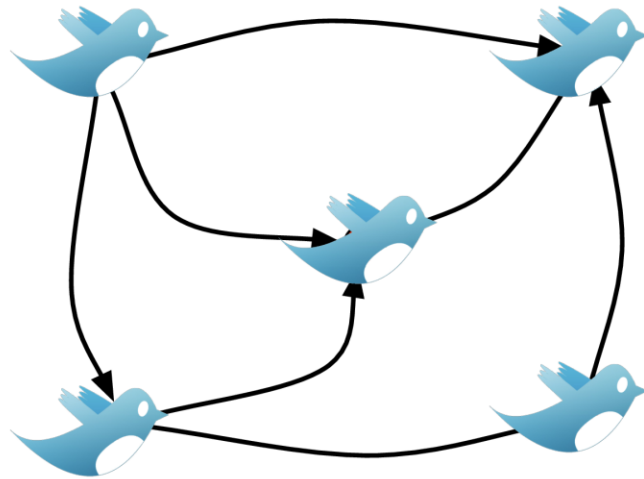
V

Gráf adatmodell



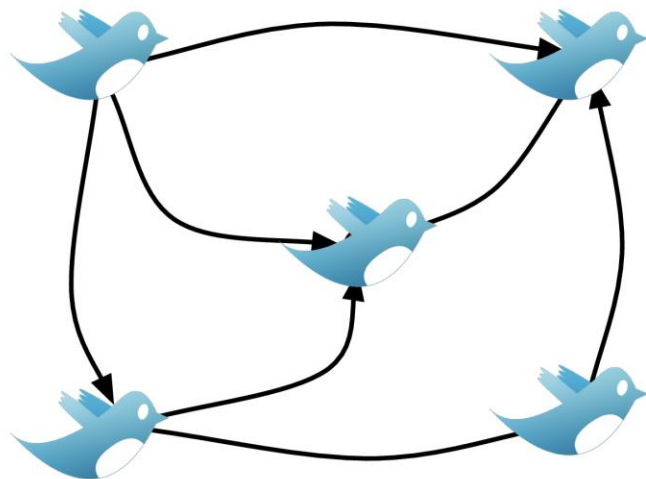
E

Korlátozott kifejezőerő



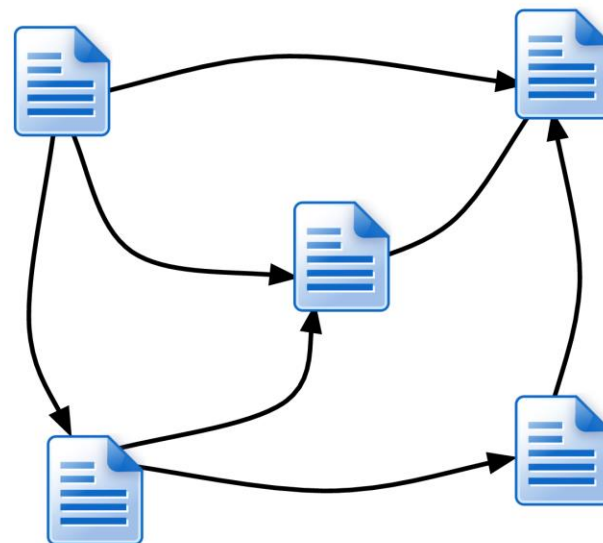
Twitter: ki kit követ

Korlátozott kifejezőerő

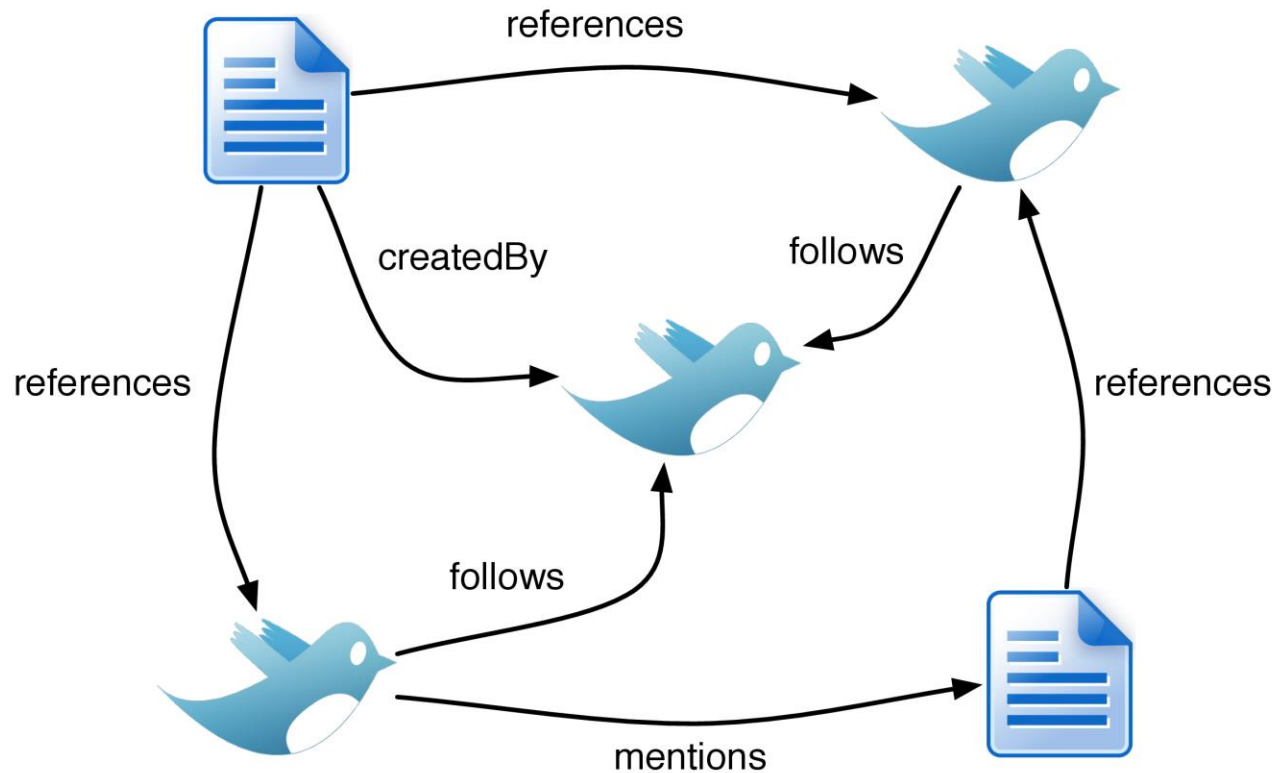


Twitter: ki kit követ

...weboldalak és linkek



Gyakorlatiasabb modell



Gyakorlatiasabb modell

Tulajdonsággráf

$$G = (V, E, \lambda)$$

- * Irányított gráf, élek: címkék, csomópontok: tulajdonságok
- * Többszörös élek lehetségesek

Tulajdonsággráf



csúcs

Tulajdonsággráf

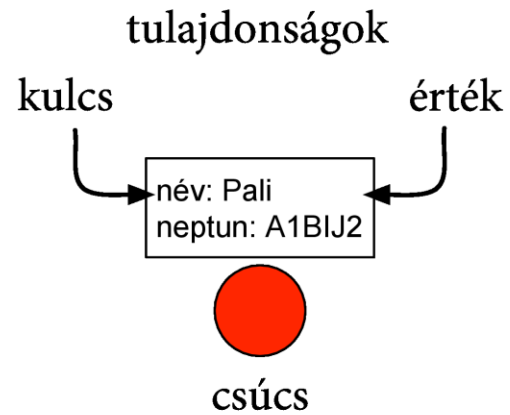
tulajdonságok

név: Pali
neptun: A1BIJ2



csúcs

Tulajdonsággráf

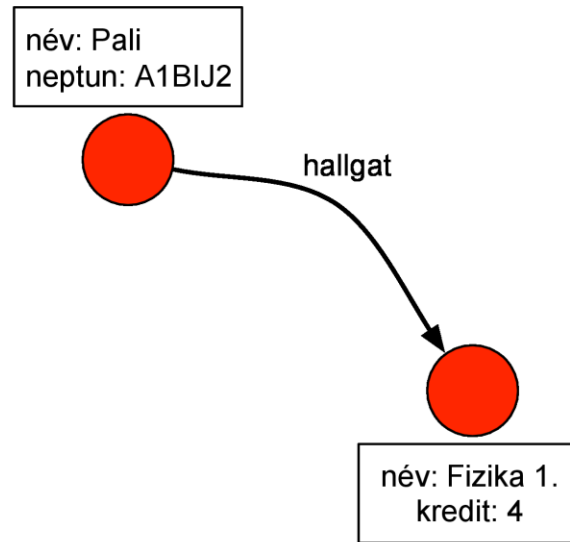


Tulajdonsággráf

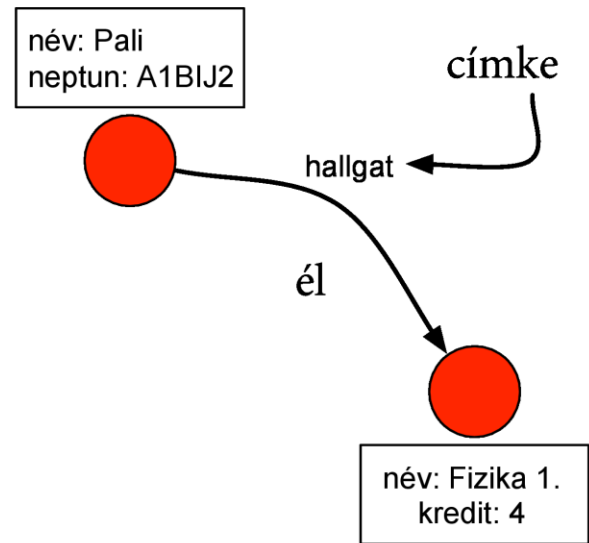
név: Pali
neptun: A1BIJ2



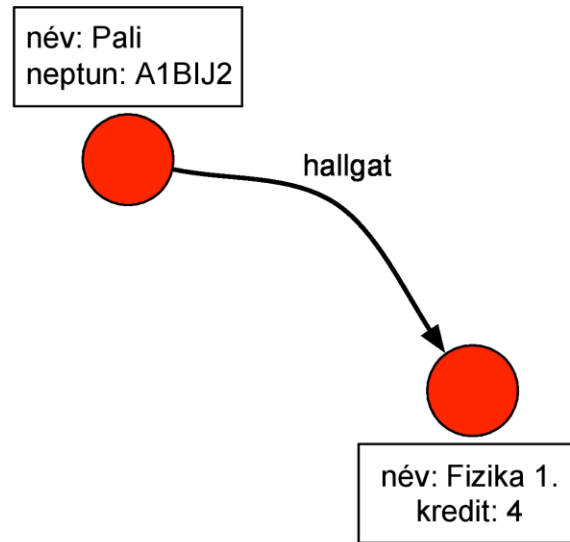
Tulajdonsággráf



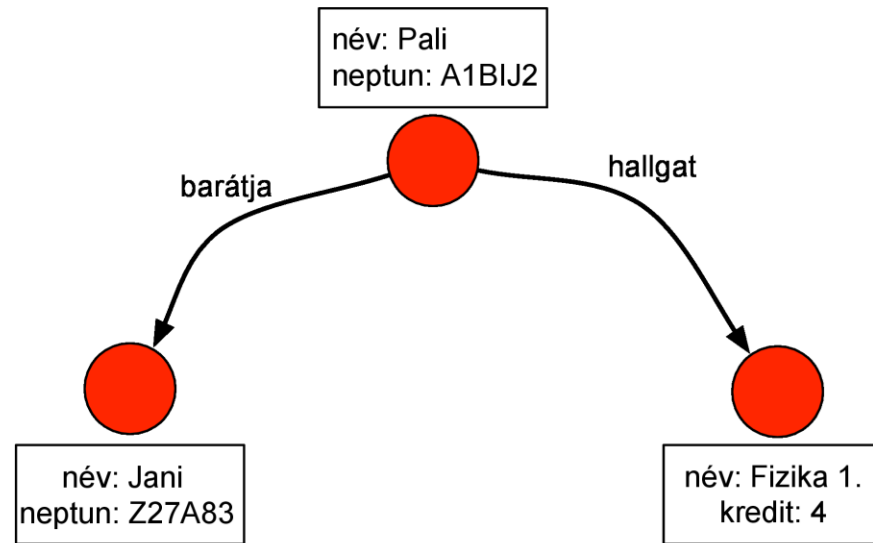
Tulajdonsággráf



Tulajdonsággráf



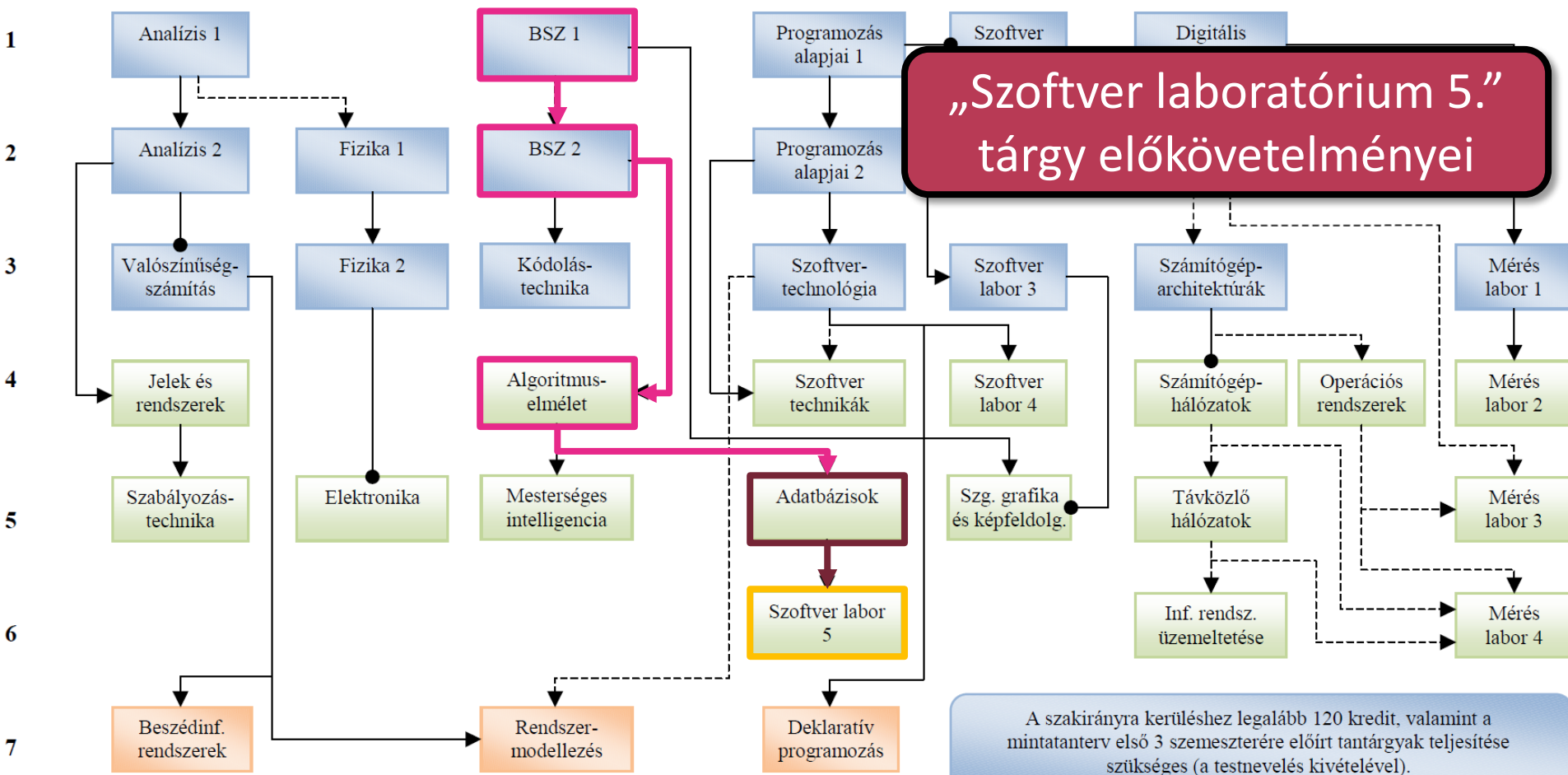
Tulajdonsággráf



Gráfadatbázis implementációk

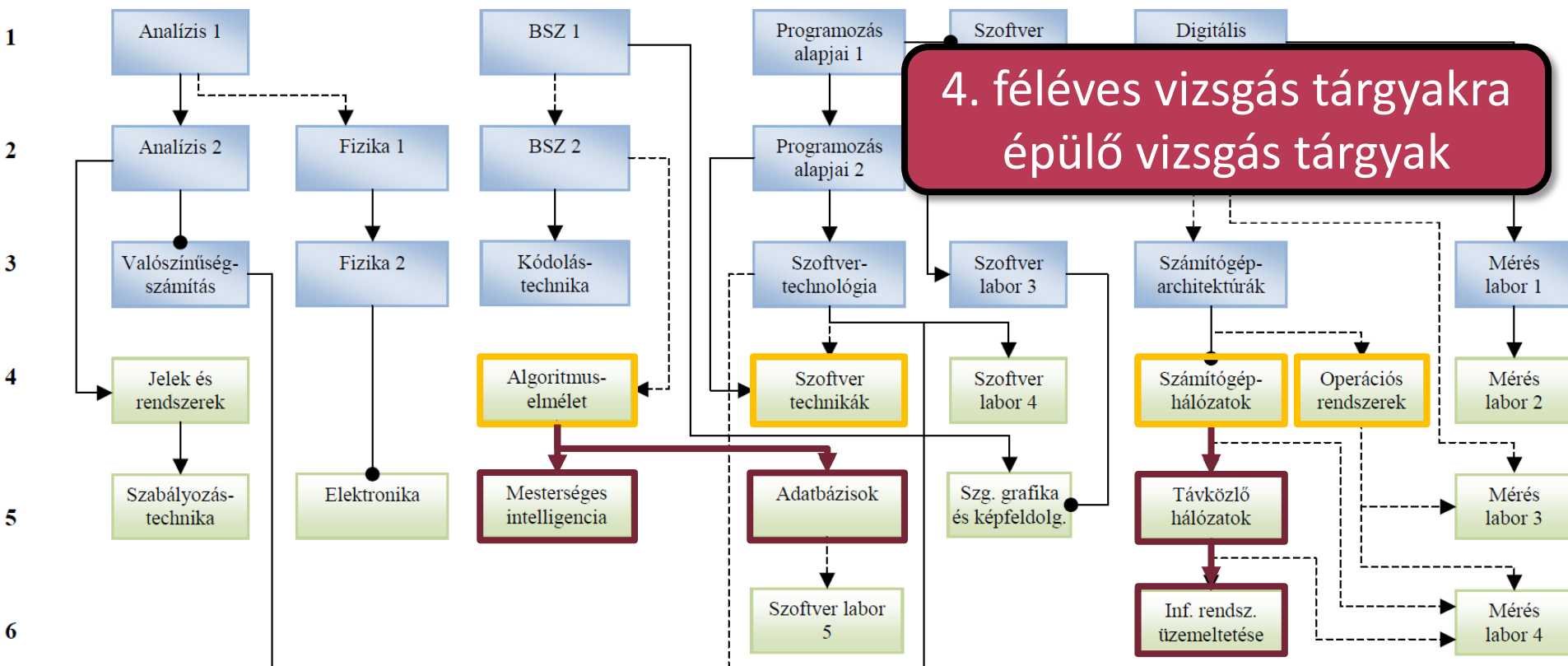


Mérnök informatikus előtanulmányi rend



```
START n=node:node_auto_index(name='Szoftver laboratorium 5.')
MATCH a-[r*]->n
RETURN a, r
```

Mérnök informatikus előtanulmányi rend



```
START n=node:node_auto_index(semester='4')
```

```
MATCH n-[r*]->o
```

```
WHERE n.exam = TRUE
```

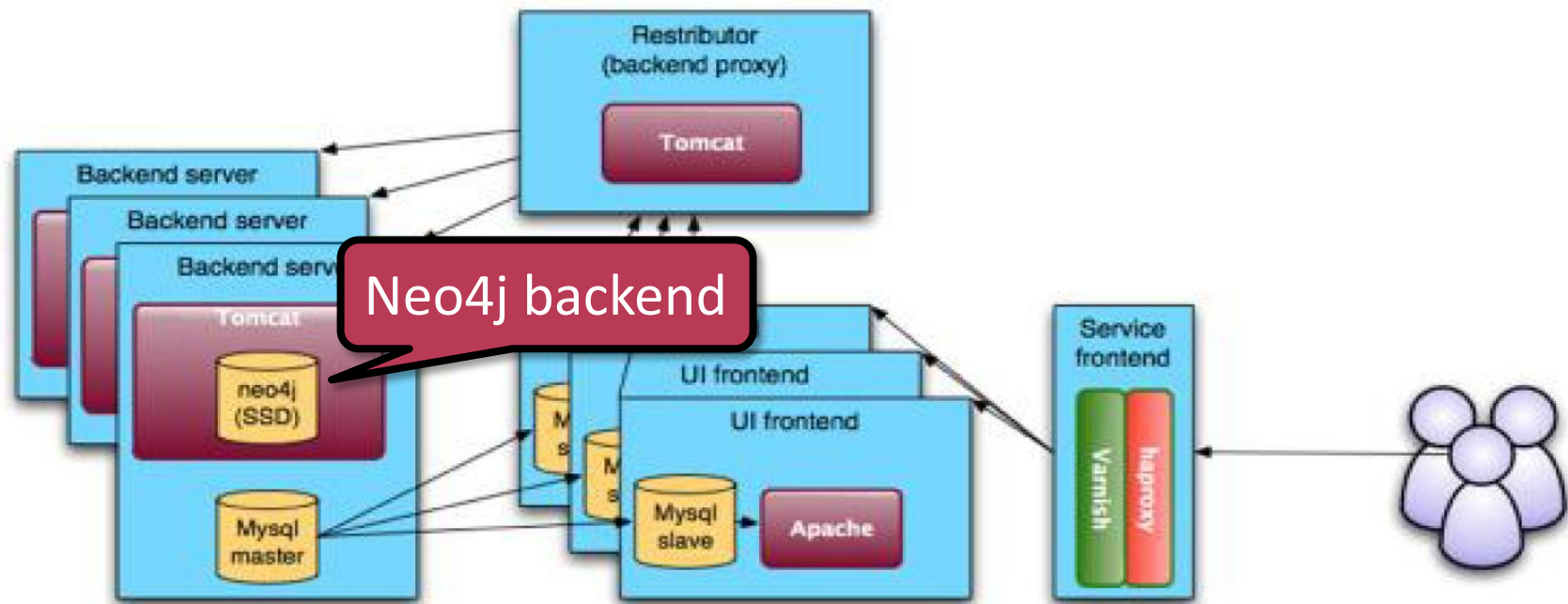
```
AND o.exam = TRUE
```

```
RETURN o
```

Alkalmazás: SZTAKI szótár

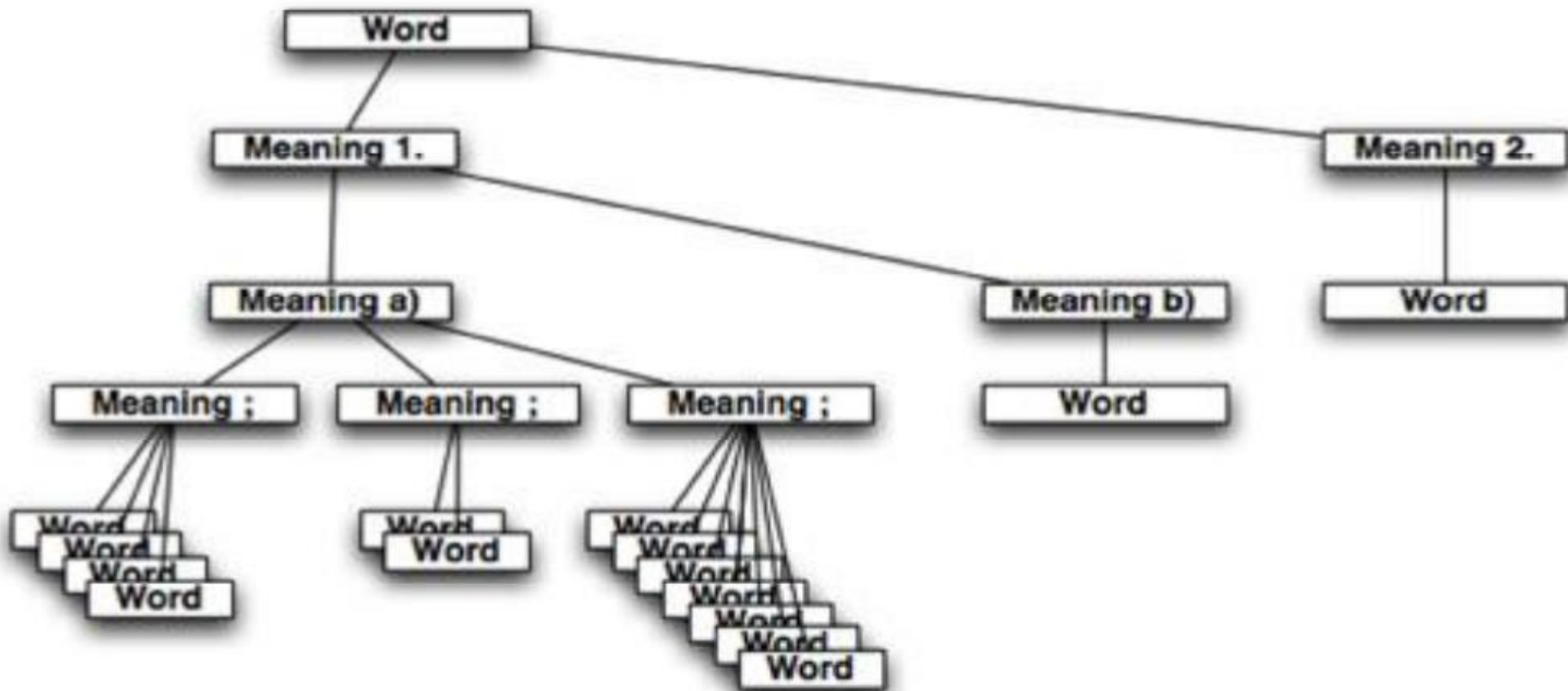
[origo] sztaki szótár

bármiről ▾ —> bármire ▾



SZTAKI szótár – szócikk gráfja

abatement: 1. a) csökkenés, kisebbedés, enyhülés, gyengülés; alábbhagyás, lecsendesedés [viharé] lankadás, megfogyatkozás, apadás, alábbszállás, szűnés, csökkenés, enyhülés [tüneteké]; csillapodás, csökkenés [lázé]
b) megszüntetés [visszaélésé]; 2. címertörés



Források

- NoSQL adatbázis-kezelők:
<http://nosql-database.org/>
- Christof Strauch, NoSQL Databases, <http://www.christof-strauch.de/nosql dbs.pdf>
- SZTAKI Szótár – Olyan jó, hogy nem találom a szavakat,
<http://nws.niif.hu/ncd2012/docs/ehu/022.pdf>
- Titan: The Rise of Big Graph Data,
<http://www.slideshare.net/slidarko/titan-the-rise-of-big-graph-data>
- BME Adatbázisok haladóknak kurzus,
<https://www.db.bme.hu/targyak/adatbazisok-haladoknak>

Technológiák

- Cassandra:
 - <http://cassandra.apache.org/>
 - <http://www.datastax.com/>
- Neo4j: <http://www.neo4j.org/>
- Titan: <http://thinkaurelius.github.io/titan/>
- Massey Architecture Explorer,
<http://xplrarc.massey.ac.nz/>
- Graphviz: <http://www.graphviz.org/>

Érdekessegek

- MongoDB shell: <http://try.mongodb.org/>
- Neo4j konzol: <http://console.neo4j.org/>
- Sakkjátszmák gráfadatbázisban:
<http://gist.neo4j.org/?6506717>