



Oszlopcsaládok a NoSQL világban

Gábor Bernát
bernat@primeranks.net



Adatbázisok haladóknak 2012.

2012. október 16.



Miről lesz szó?

Miért szükséges az oszlopcsalád és NoSQL?

A CAP háromszögről ismét

Cassandra és HBase-nak:

- Adat partíció
- Adat struktúra

Kérdések, válaszok
Amikor és ahol felmerülnek



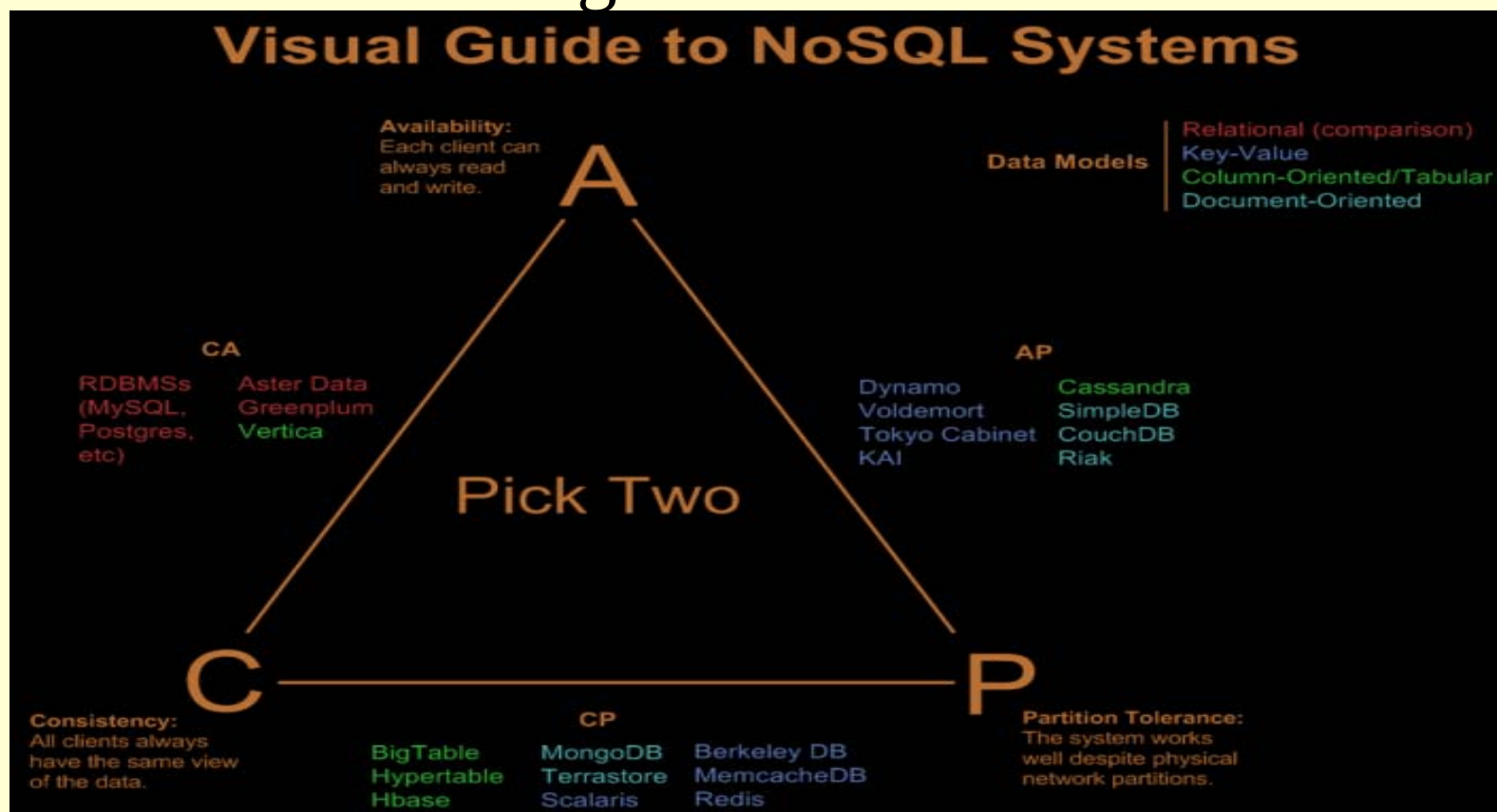
Miért szükséges az oszlopcsalád?

- 2000-ben a Google – skálázhatósági gondok
- Relációs adatbázisok ☺ és ☹
- Code 13 szabálya
- Meddig skálázható egy relációs adatbázis?



A CAP háromszög

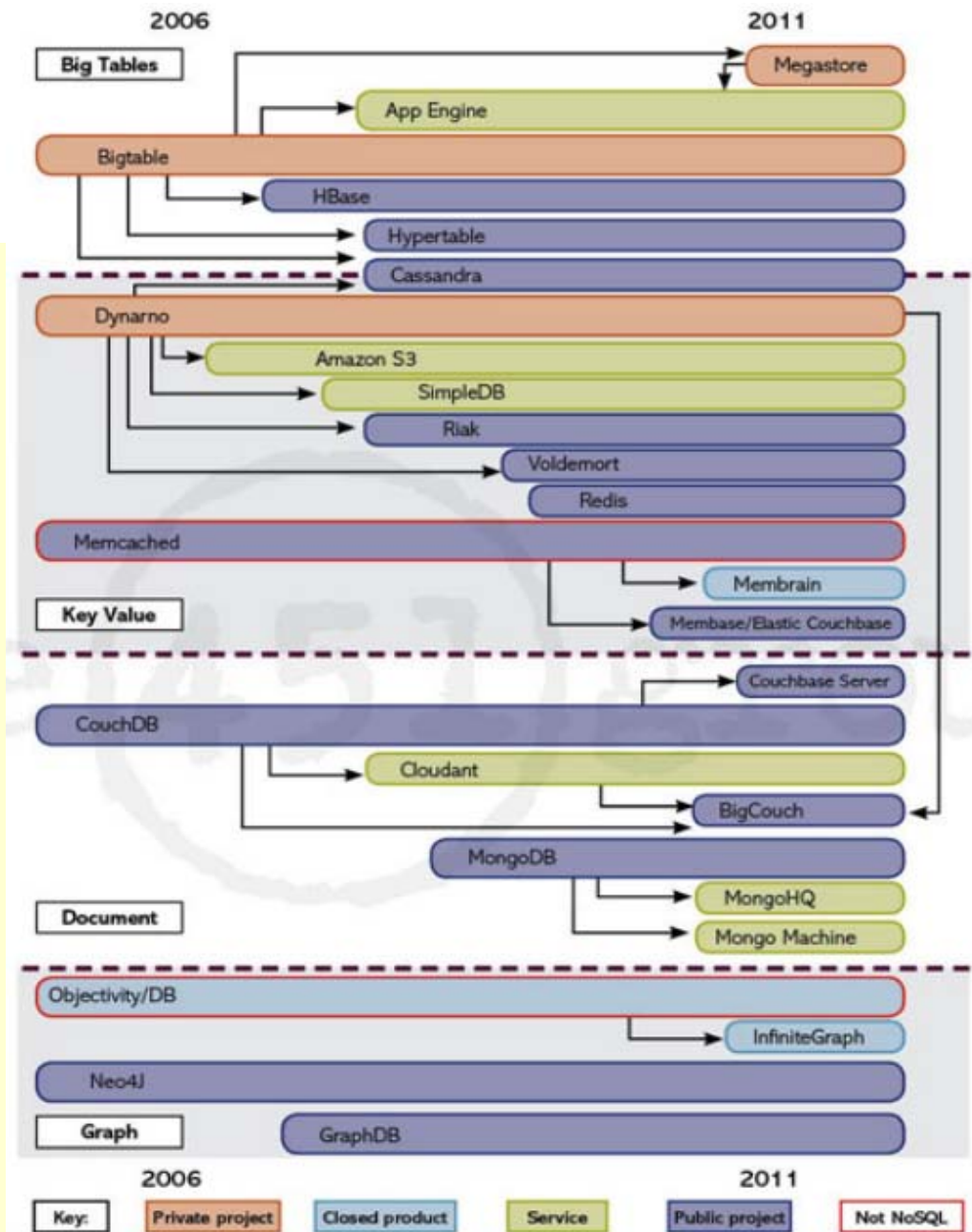
Honnan is származik egy relációs adatbázis skálázási nehézsége?





És elindultunk...

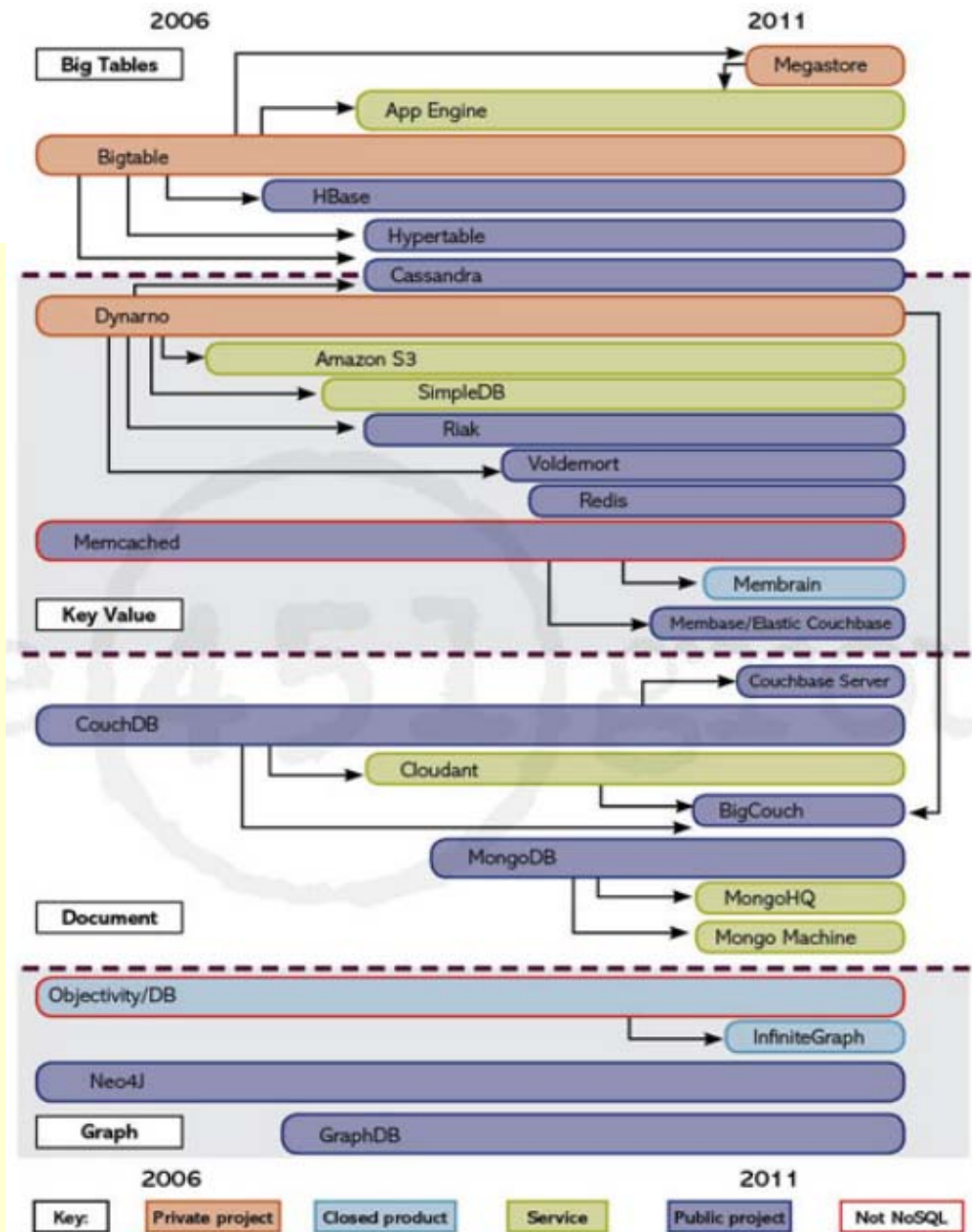
- Közelítsünk a skálázhatóság felől





És elindultunk...

- Közelítsünk a skálázhatóság felől





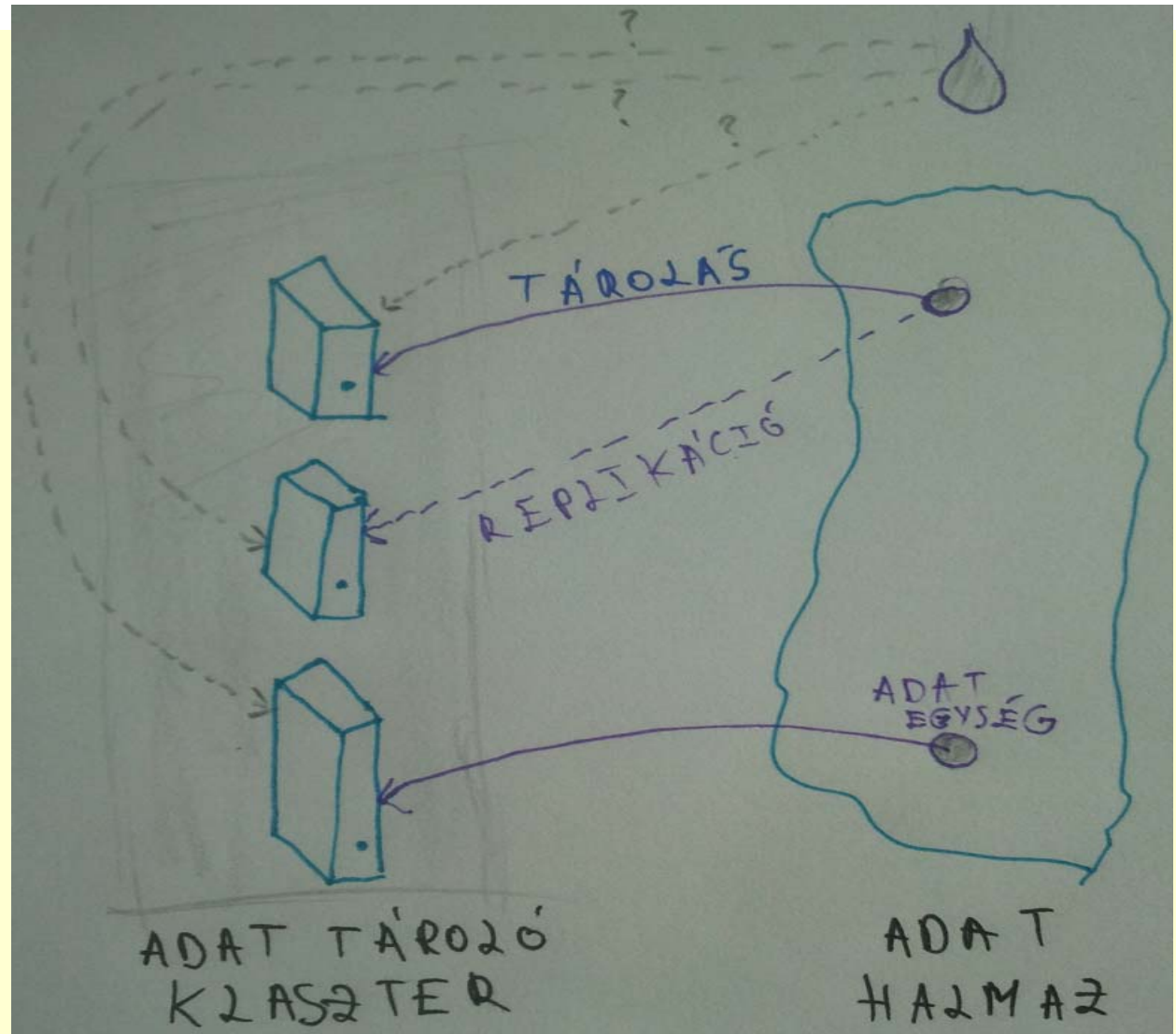
Relációs vs. NoSQL oszlopcsaládok

- Közös: rendezzük oszlopba az adatot
- “Széles” oszlop családok (strukturált szótár)
- Skálázhatóság: vertikális vs. Horizontális
- Egyszerűsített adat modell
- Adattár mennyiség sebességért
- Mit szeretnénk lekérdezni, nem pedig mit szeretnénk tárolni



Cassandra – partíció

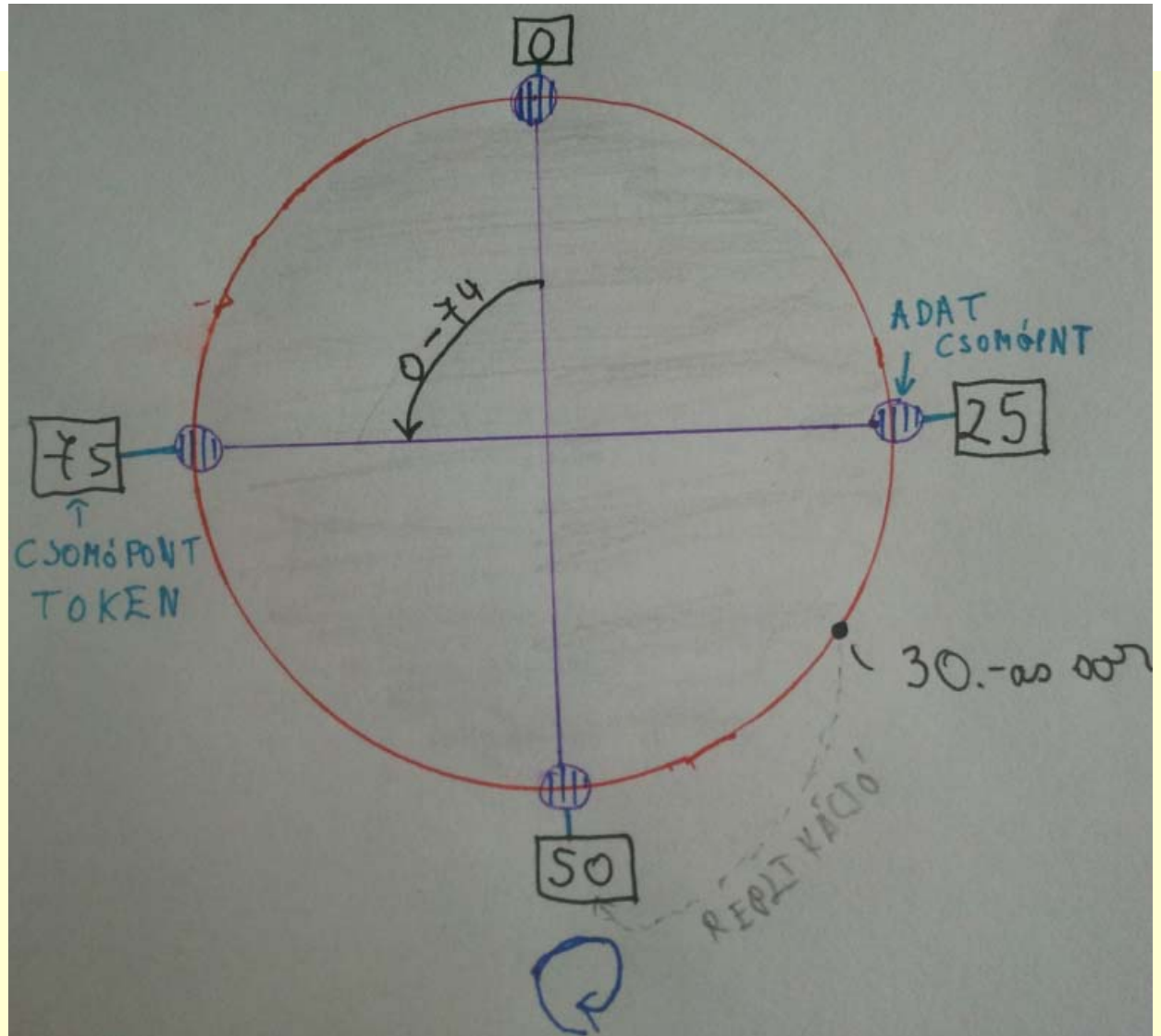
- A probléma:





Cassandra – partíció

- A megoldás:
egy adatgyűrű





Cassandra – partíció

- Írás és olvasási konzisztencia futás időben specifikálható
- Lokális kvórum – miért is kell?



Cassandra – adat modell

- Séma => kulcstér
- Tábla => Oszlopcsalád
- Oszlopok => Oszlop
- $\nexists$ relációk (JOIN), idegen kulcs
- Ami szükséges egy lekérdezéshez egy oszlopcsaládba kerül



Cassandra – oszlopcsalád típusok

- Statikus (*ritka mátrix*)

row key	columns ...			
jbellis	name	email	address	state
	jonathan	jb@ds.com	123 main	TX
dhutch	name	email	address	state
	daria	dh@ds.com	45 2 nd St.	CA
egilmore	name	email		
	eric	eg@ds.com		

- Dinamikus (akár több millió is)

row key	columns ...			
jbellis	dhutch	egilmore	datastax	mzcassie
dhutch	egilmore			
egilmore	datastax	mzcassie		



Cassandra – oszlop típusok

- Statikus

column_name
value
timestamp

The diagram illustrates a static column structure. It consists of a table with three rows. The first row is the header, labeled 'column_name'. The second row contains the value 'value'. The third row contains the timestamp 'timestamp'.

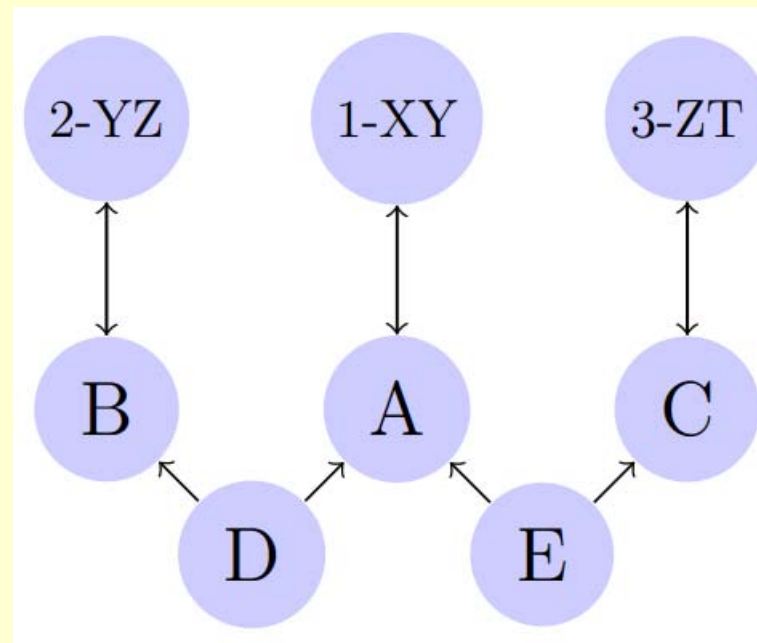
- Összetett – több kulcs egy sornak
- Számláló
- Lejáró



Cassandra – példa

tweet_id	author	body
1	A	XY
2	B	YZ
3	C	ZT

user_id	tweet_id	author	body
D	2	B	YZ
D	1	A	XY
E	4	C	ZT
E	1	A	XY





Cassandra – példa

- Ami a háttérben

D	[2,author]:B	[2,body]:YZ	[1,author] : A	[1,body] : XY
E	[4,author]:C	[4,body]:ZT	[1,author] : A	[1,body] : XY

- CQL:

```
CREATE TABLE timeline (  
  user_id varchar,  
  tweet_id uuid,  
  author varchar,  
  body varchar,  
  PRIMARY KEY (user_id, tweet_id)  
);
```

```
SELECT * FROM timeline WHERE user_id = D  
ORDER BY tweet_id DESC LIMIT 50;
```



HBase – történelem



- Google File System
- Google MapReduce
- Google BigTable
- Nutch => Hadoop (HDFS & MapReduce)
- Powerset => HBase



HBase – adat struktúra

- Egy HBase adat elérési útvonalát megad-hatjuk mint:
(Tábla, Sor Kulcs, Oszlop család, Oszlop, Időbélyeg) → Érték
- Ami adatstruktúra szinten:

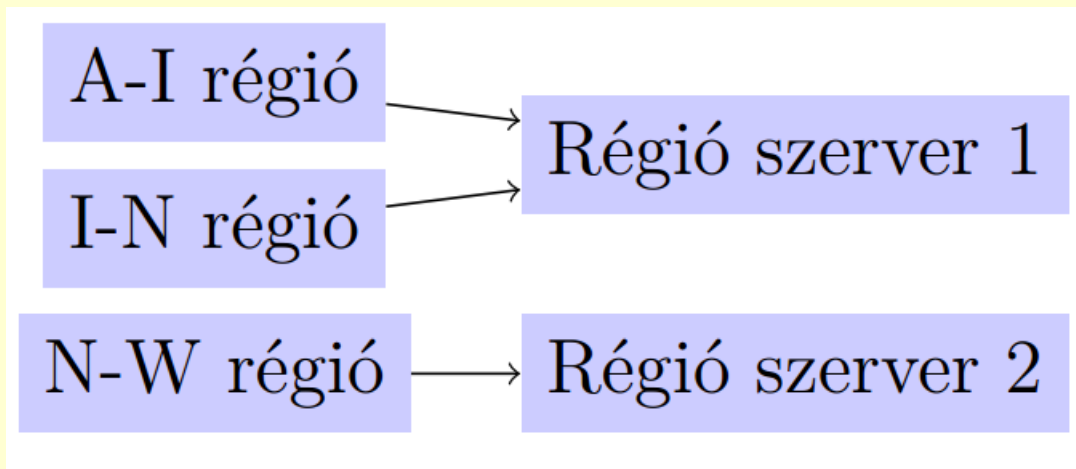
```
Map<Row, List < Map <Column, List<Value, Timestamp> > >
```

- Cella tartalma bájtsorozat, az alkalmazás értelmezi azt



HBase – partíció

- Régió leosztás sor kulcs szerint – szerverek



- A régió manifesztációja egy a Hfile, HDFS tárolás
- Write Ahead Log – robosztusság



HBase – felépítés

- Kliens API - *Java*
- Mester szerver
 - Terhelés elosztás
 - Meta adat menedzselés
- ZooKeeper (Google Chubby ☺)
 - Tulajdon viszony egyeztetés
 - Régió szerverek – heartbeat rendszer
- Régió szerverek – kliens kommunikáció



Kérdések...idő???

- ...válaszok ☺



Köszönöm a figyelmet!

Adatbázisok haladóknak

Gábor Bernát
bernat@primeranks.net



Adatbázisok haladóknak 2012.

2012. október 16.