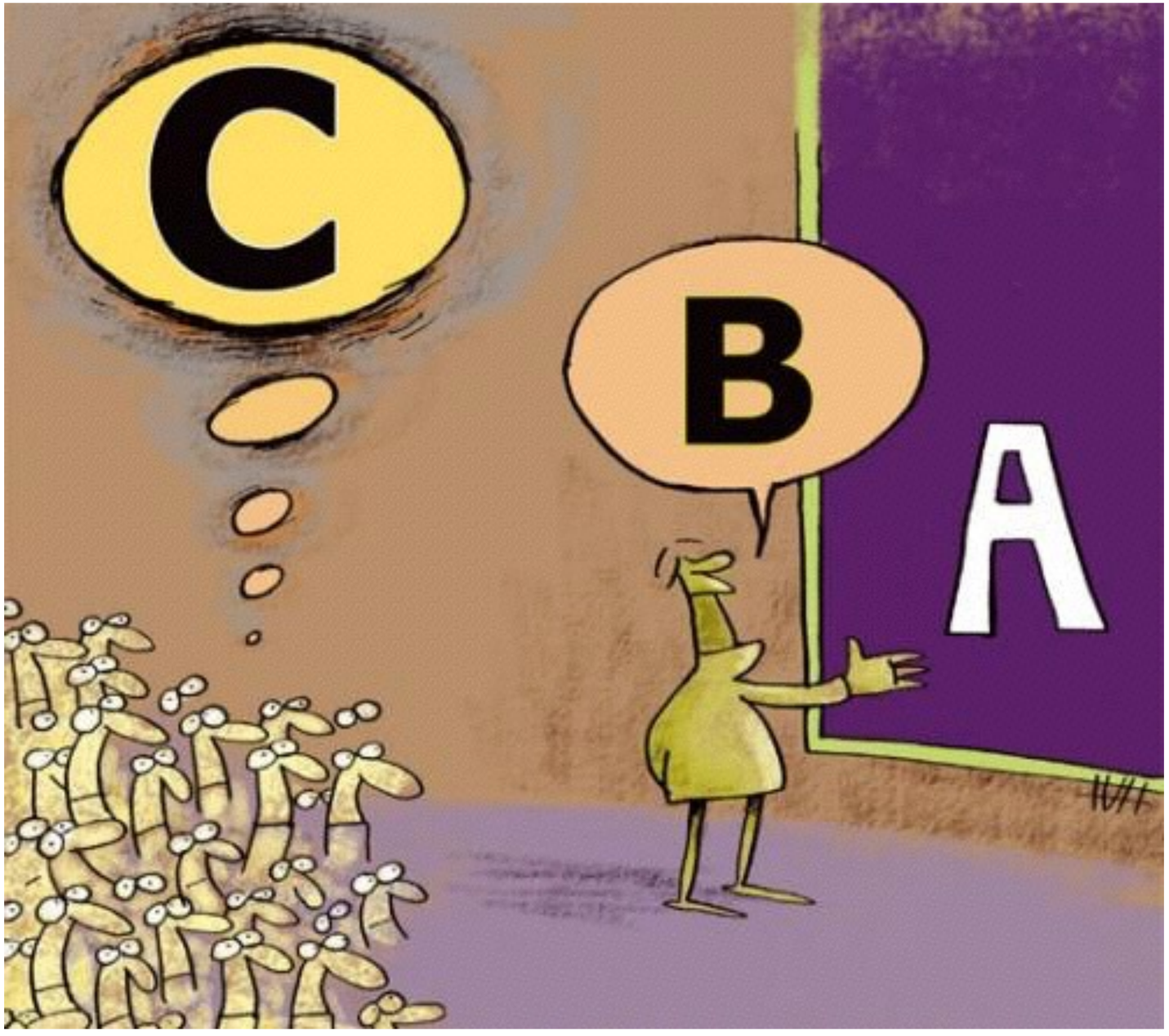
"Business group showing diversity in a meeting"


C
B
A

Z + 6 N z ? (
g
$ r
M + 2
L ?
X G s e 2 < £
Z s 4 & e i
L
@ M X { {
C { $ P 8 8 t
= Q
b
0
M
K
0 j T o x d
d
^ A N s L e 8 | O 7
L v ` P m P u 4 ` s
{ — — O l p S t 0)
g (E D . X C I O o
' 9 { \ 2 s T U s [
H 6 9 J E — Z 8 ~ .
u # h d Z O l w v O
o J y (— e g " O /
J C h 6 n ^ e a x M
j [F ; ! v Y # {
W ; | 1 K 4
W > t W = N 8 $ L :
^ 8 F] ; O 4 ; b h
~ g ( ) ; — I W r u
{ O 5 i 6 Y ` X] A
z f & 0 5 ` ; * * &
" T & 8 c ? 7 g T h
v 2 D 9 4 , [{
E >
z s # # Z J K ` X | T J d &
f # # x c O @ ; q
; d : $ + a
M j f P u
:
0
"


Microsoft®
.net™
Java™


MySQL®
PostgreSQL


Microsoft®
.net™
Java™

Mono culture = clean

Diversity = messy

Embracing Diversity in Databases

Frank @Lyaruu

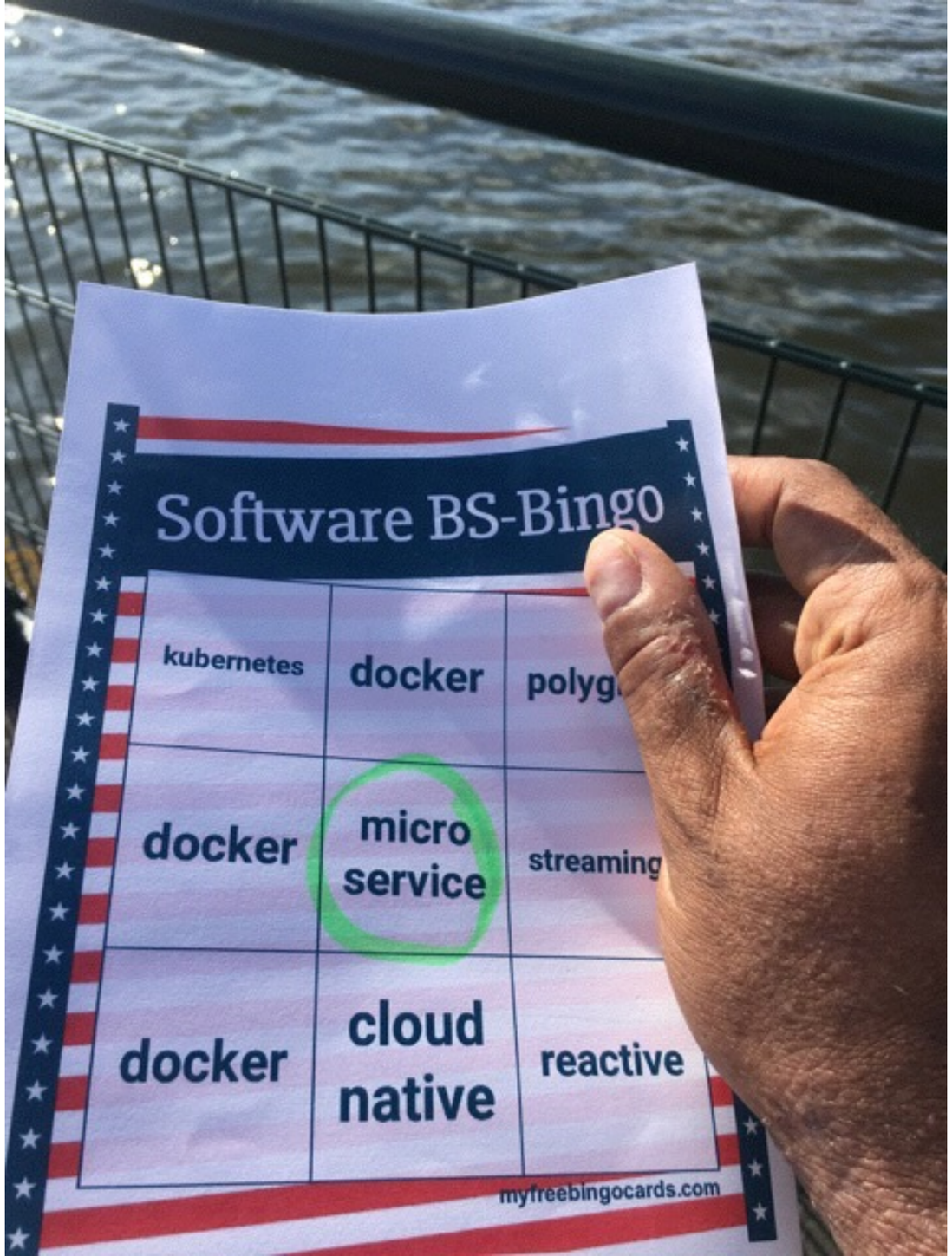
CTO Dexels

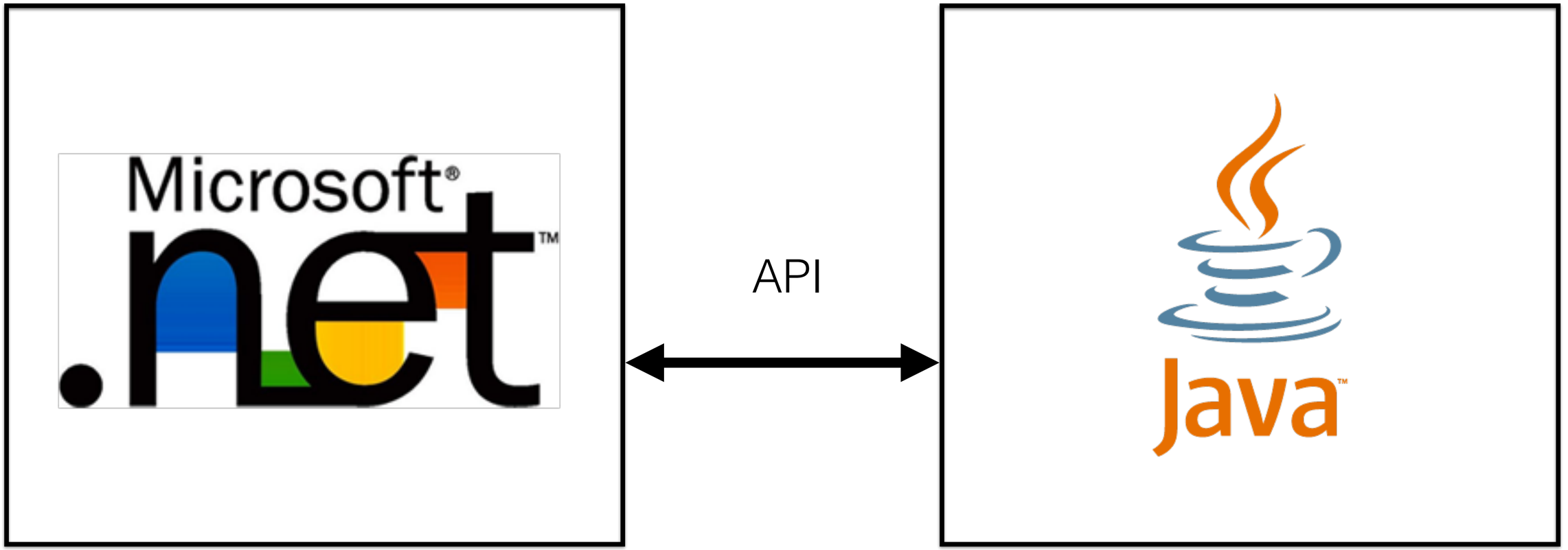
Technology Hipster

Amsterdam

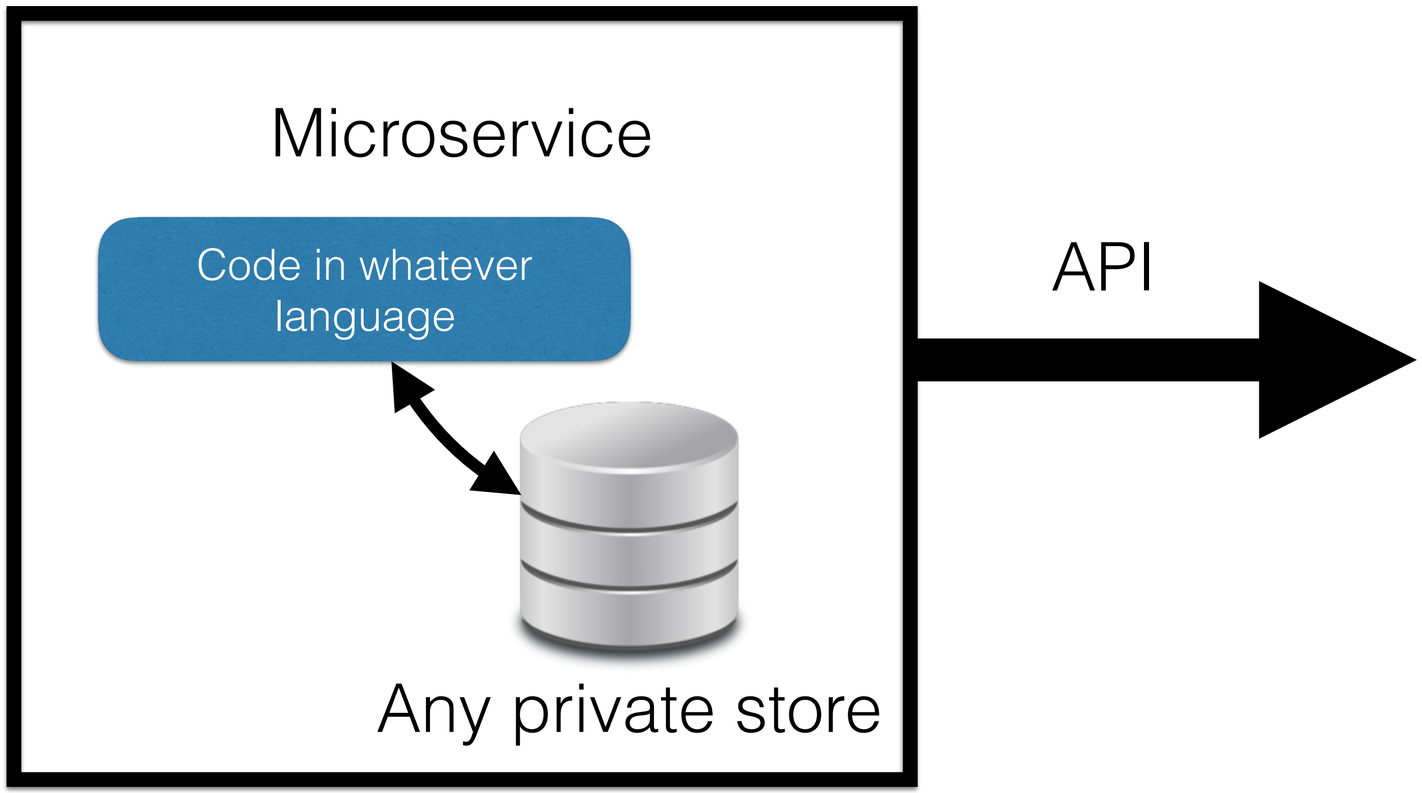
BERLIN BUZZWORDS 2017 JUNE 11-13

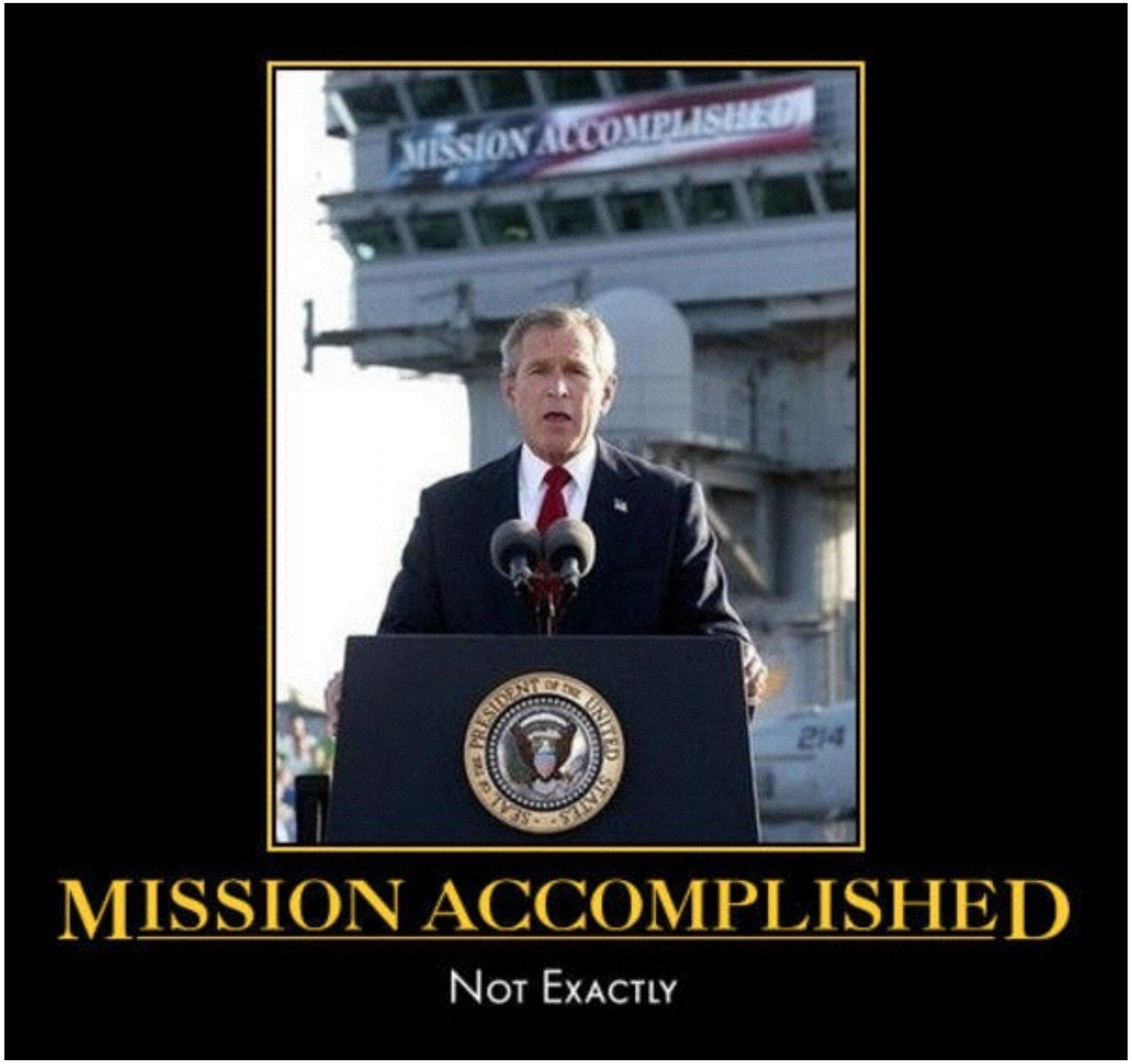
How can we reduce the cost of diversity in software?


Software BS-Bingo
kubernetes
docker
polyg
docker
micro service
streaming
docker
cloud native
reactive
myfreebingocards.com


Microsoft®
.net™
API
Java™

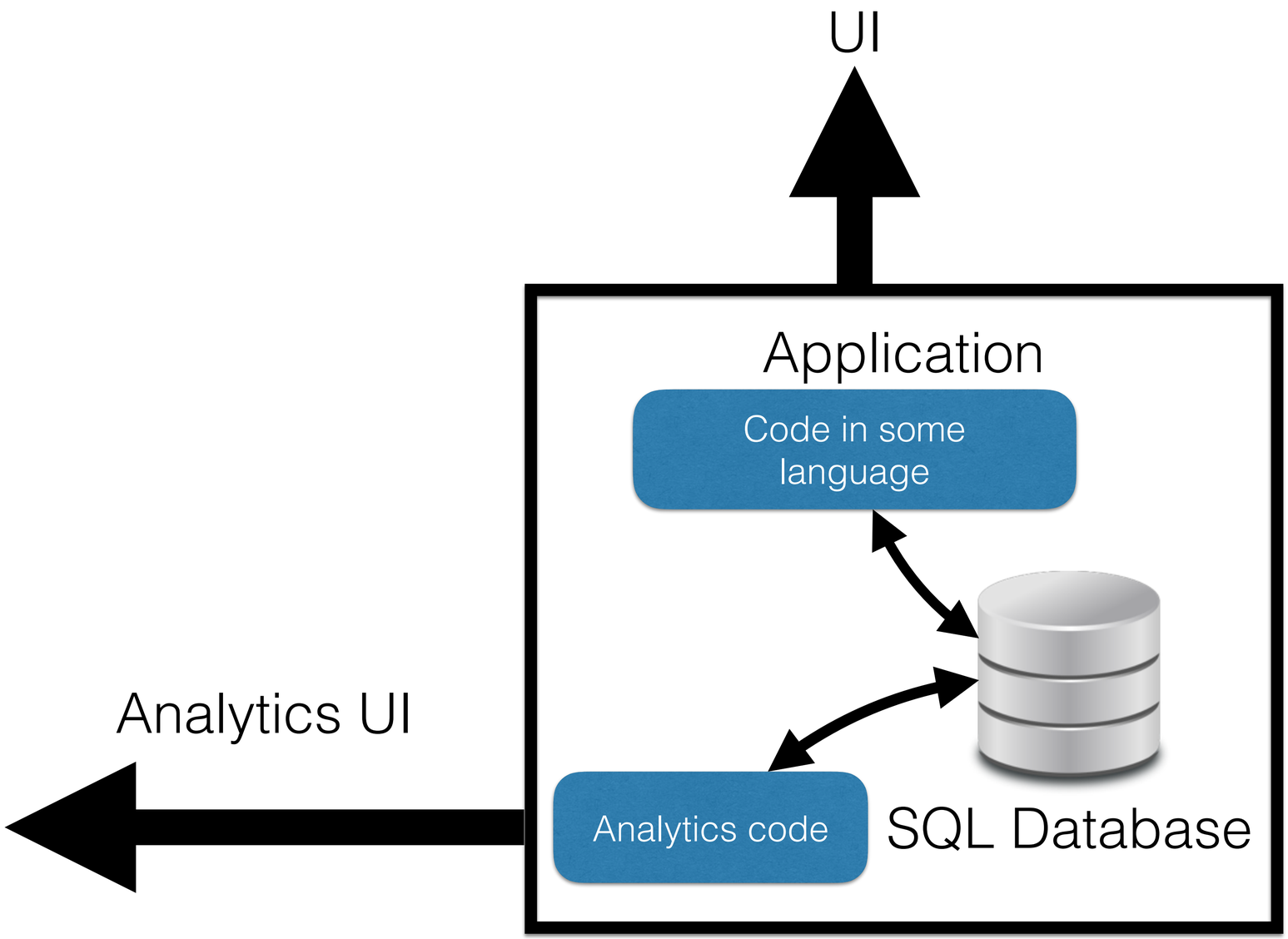
How can we reduce the cost of diversity in ~~software~~ databases?

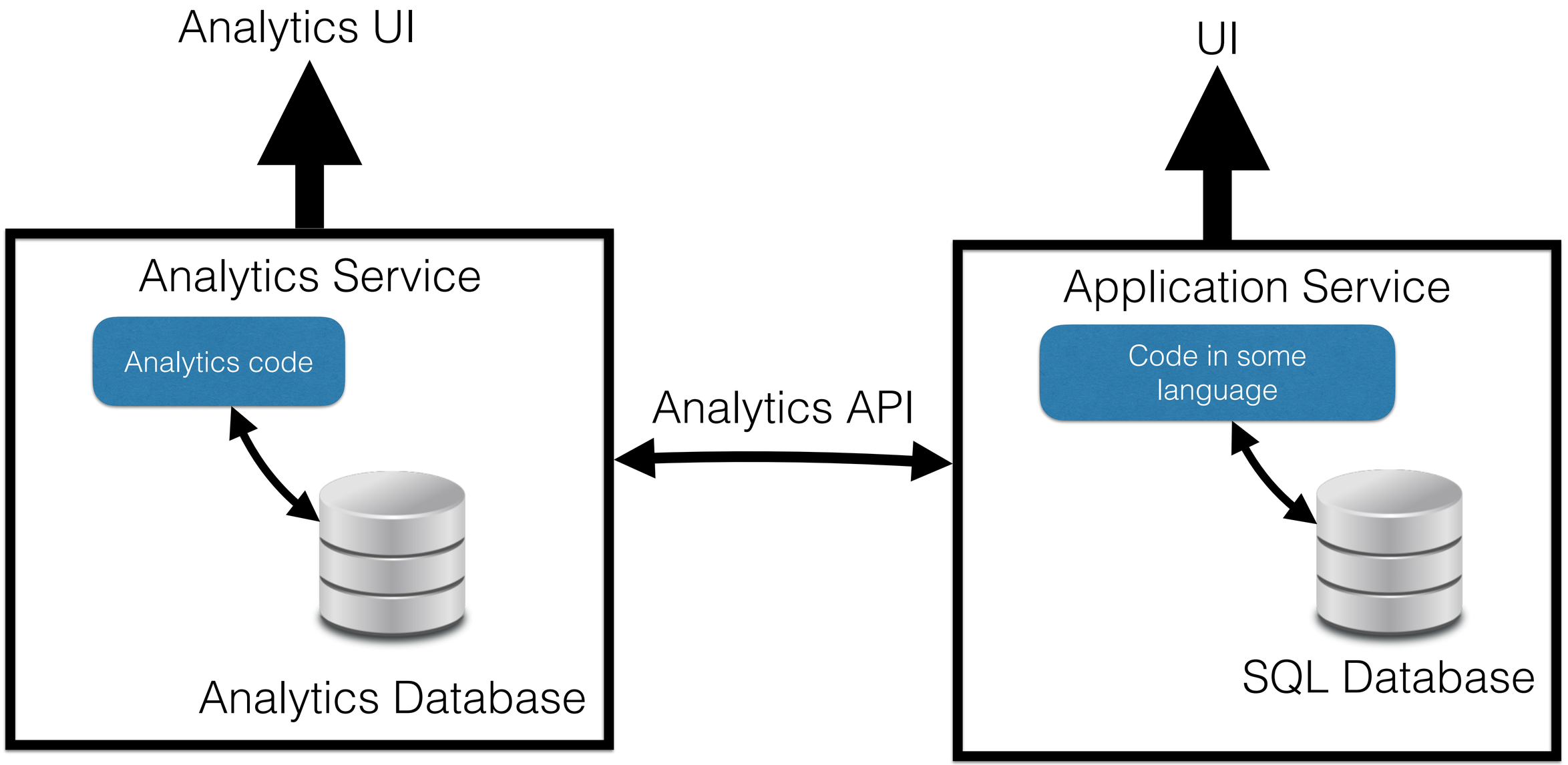

Microservice
Code in whatever language
API
Any private store


MISSION ACCOMPLISHED
MISSION ACCOMPLISHED
NOT EXACTLY

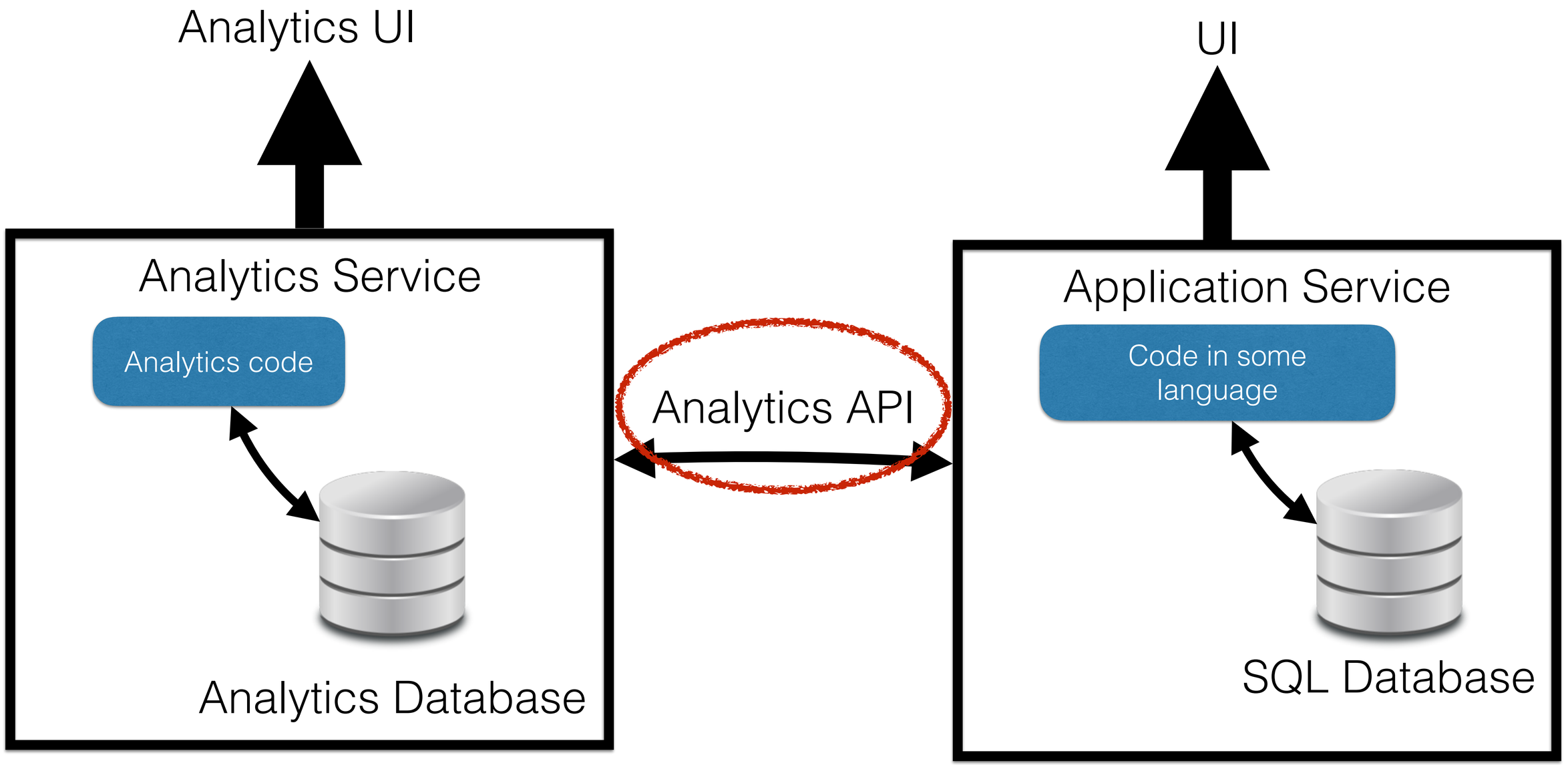
Different parts need the same data

… but in a different way


UI
Application
Code in some language
Analytics UI
Analytics code
SQL Database


Analytics UI
Analytics Service
Analytics code
Analytics Database
Analytics API
UI
Application Service
Code in some language
SQL Database

That was easy!


Analytics UI
UI
Analytics Service
Analytics code
Analytics API
Application Service
Code in some language
Analytics Database
SQL Database

What does that even mean?

How would that work?

GET /give-me-everything
?

GET /get-all-personids

+

GET /person?id=123

?

GET /data?query=“SELECT AVG(age) FROM PERSON”

Analytics UI

UI

Magic* Replication Tool

Analytics Service

Analytics code

SQL Database

Application Service

Code in some language

SQL Database

*Expensive

Issues

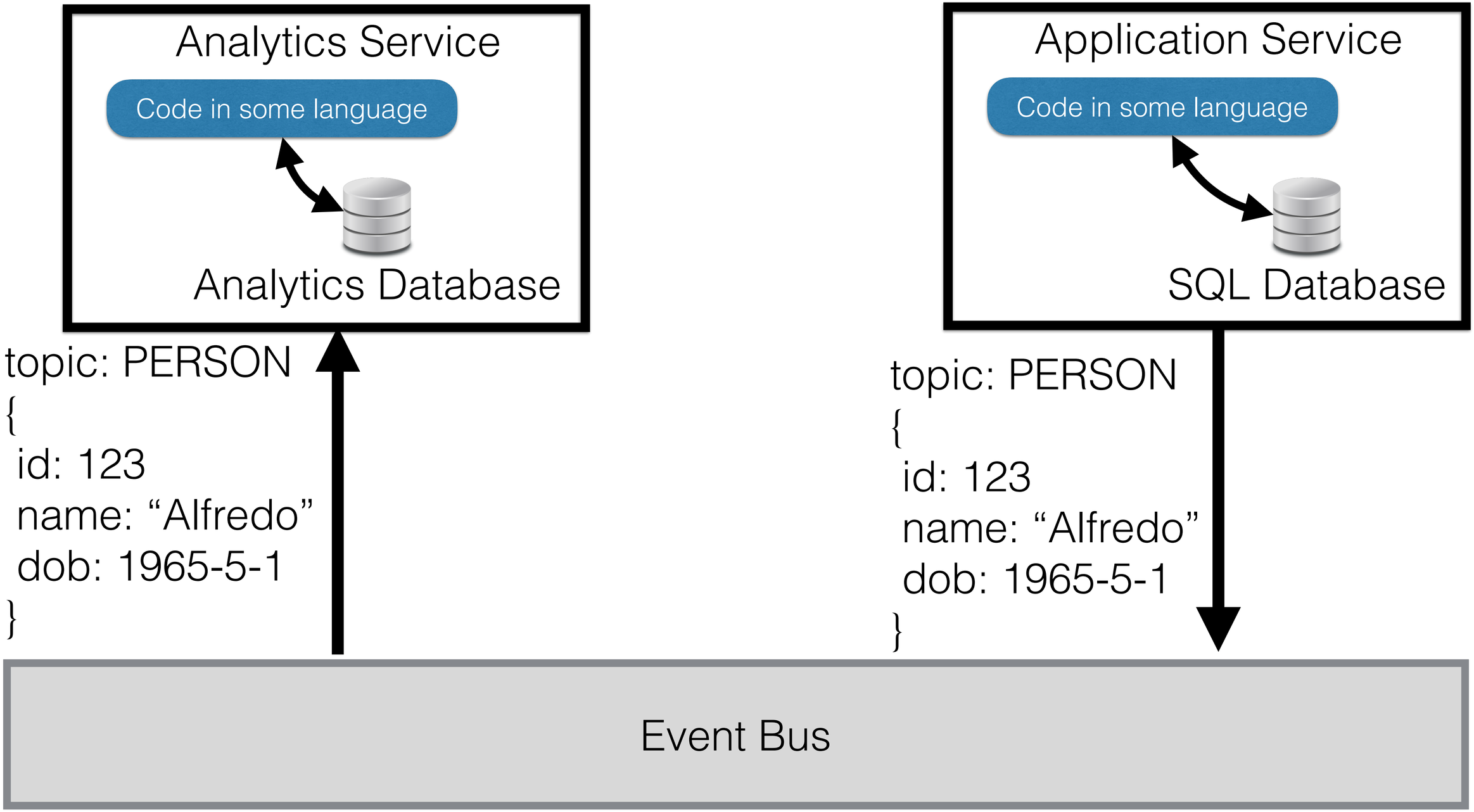
- Data stores are no longer private. We have a tight coupling between databases. It’s not a micro service
- We can’t use any interesting databases :-(
- Difficult to scale out to many services

Event Driven Microservices

Not your father’s micro services

Event Driven Microservices

- Services push events instead of a request/response model
- Usually backed by a publish/subscribe bus


Analytics Service
Code in some language
Analytics Database
Application Service
Code in some language
SQL Database
topic: PERSON
{
id: 123
name: “Alfredo”
dob: 1965-5-1
}
topic: PERSON
{
id: 123
name: “Alfredo”
dob: 1965-5-1
}
Event Bus

Kafka

Kafka

- Persistent pub/sub message bus
- High throughput
- Subscribers can consume at their own speed
- Subscribers can request a ‘rewind’ and re-consume a topic
- Has some tricks to keep the data volume down
- Having both fast and slow consumers is not a problem

sport.link

- Service provider for professional and amateur team sports in the Netherlands
- 10+ years old
- Managing personal data, planning competitions, assigning officials, supplying data feeds

sport.link

- 1M+ players
- 4K+ clubs
- 40K matches a week
- Spikey but predictable load

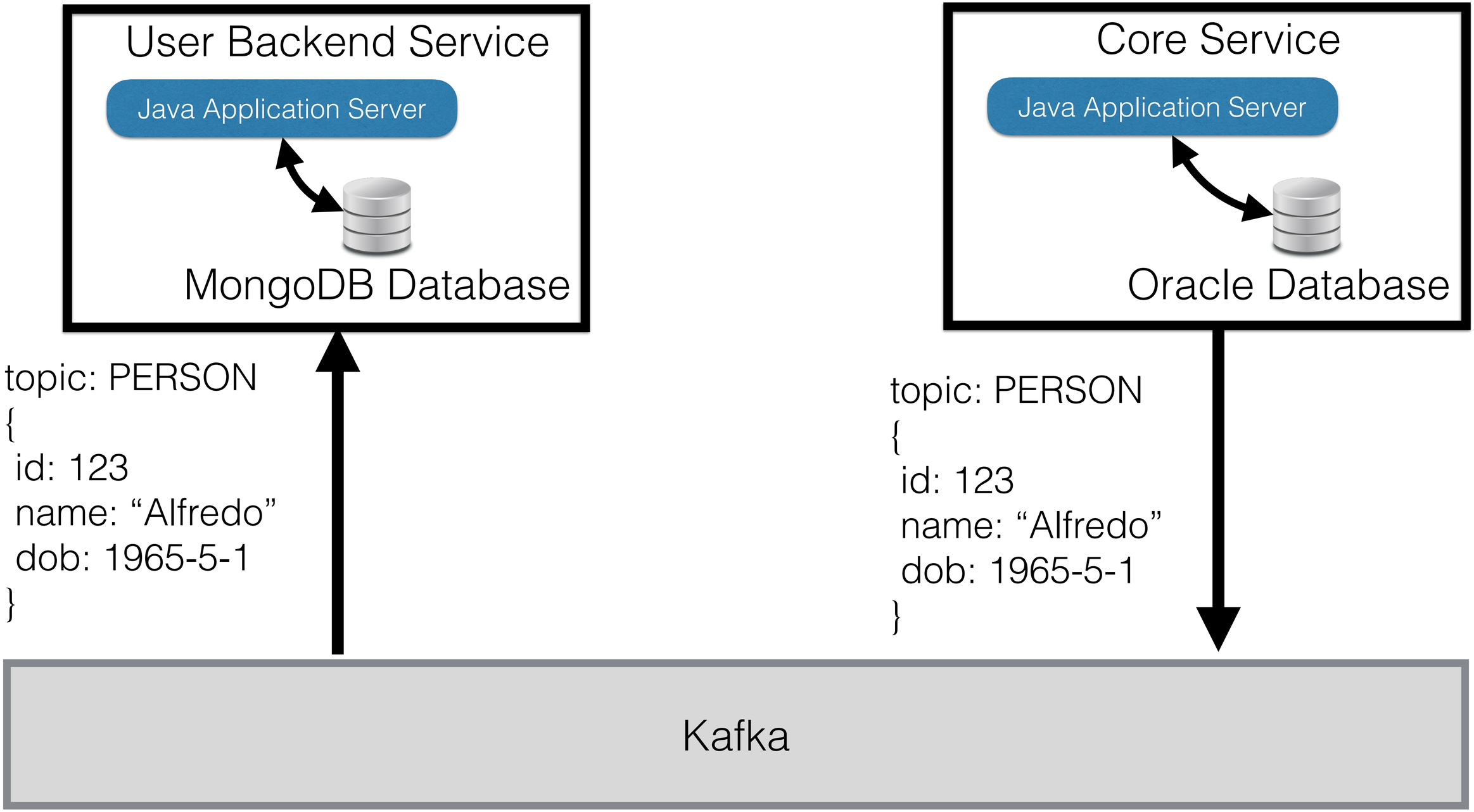
Technology stack

sport.link

- Oracle database
- Cluster of Java based application servers
- Diverse set of clients

Challenge

- Move to a player centric model instead of a club centric
- Order of magnitude more users and load
- Moving away from Oracle is not feasible in the short term
- Scaling Oracle is just too expensive


User Backend Service
Java Application Server
MongoDB Database
Core Service
Java Application Server
Oracle Database
topic: PERSON
{
id: 123
name: “Alfredo”
dob: 1965-5-1
}
topic: PERSON
{
id: 123
name: “Alfredo”
dob: 1965-5-1
}
Kafka


Now what?

Example

TelephoneId	PersonId	Type	Telephone
1	1	Mobile	12345
2	1	Email	alfredo@aol.com
3	1	Twitter	@alfredo

PersonId	Name	Birthdate
1	Alfredo	1990-01-01
2	Ben	1991-01-02

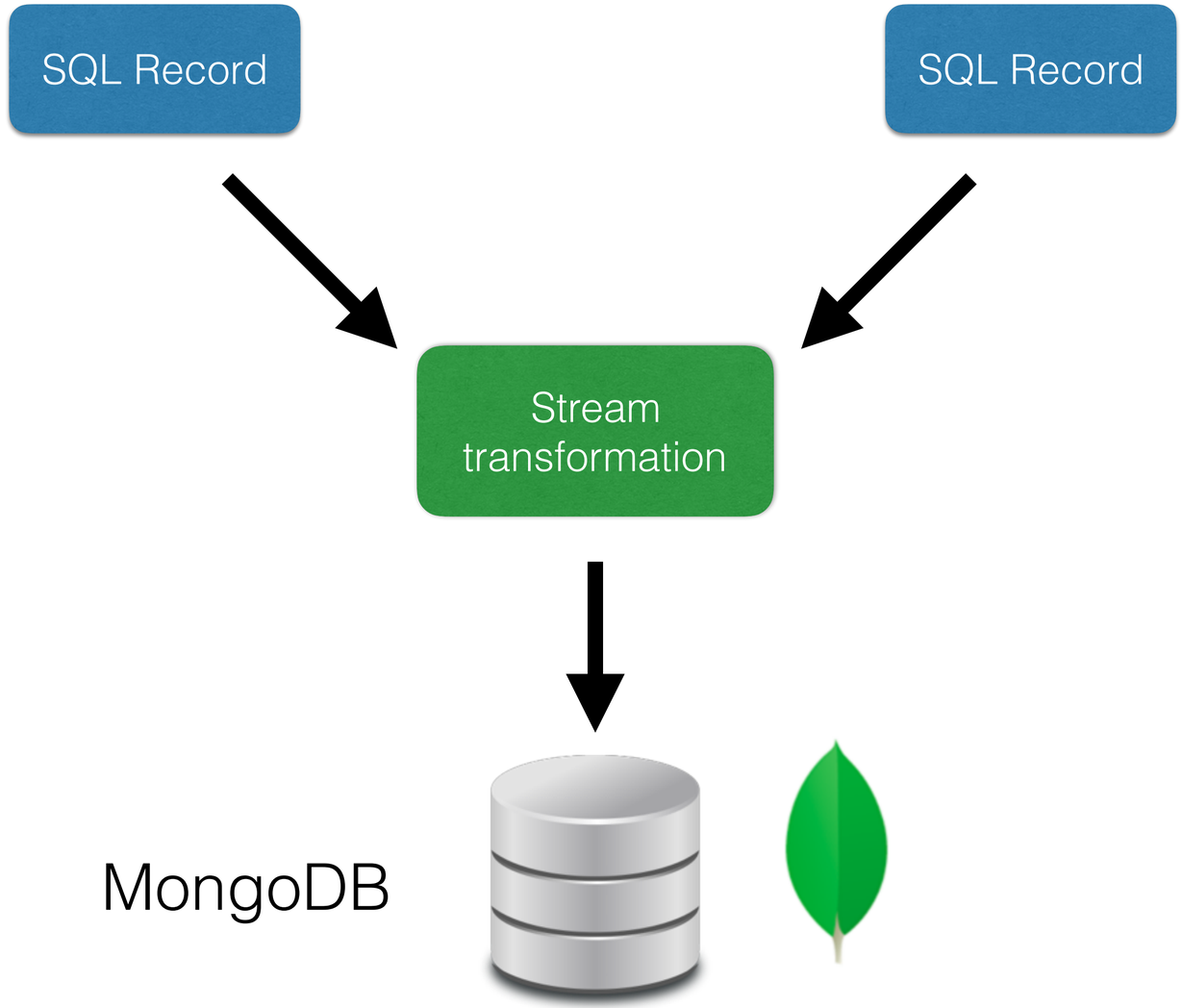
SELECT * FROM Communication C WHERE PersonId = 1

SELECT * FROM Person WHERE PersonId=1

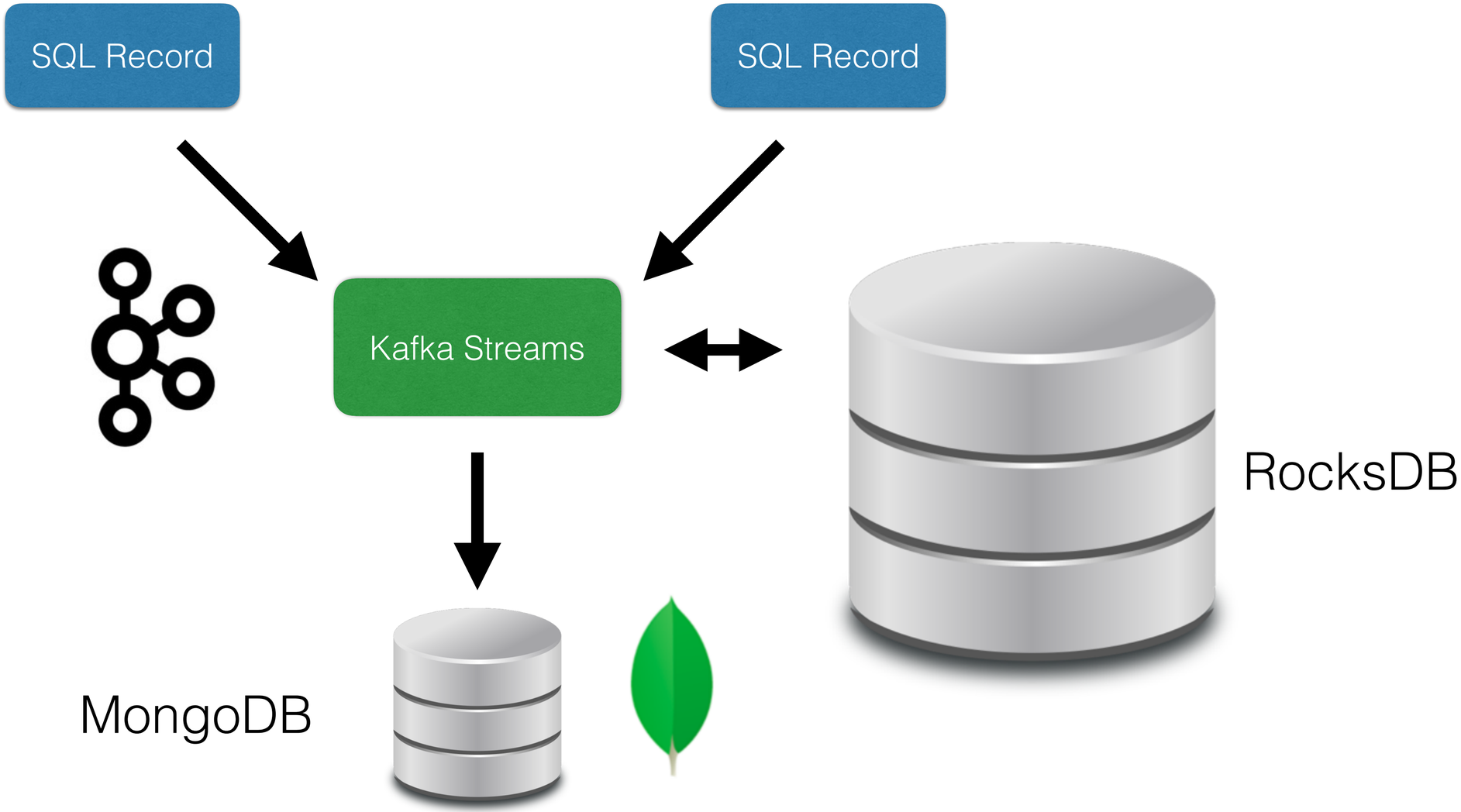
MongoDB

```
{
   “_id":1,
   "Name":"Alfredo",
   "DOB":"1990-1-1",
   "Communication":{
      "Mobile":"12345",
      "Email":"alfredo@aol.com",
      "Twitter":"@alfredo"
   }
}
```

Stream Processing



Stream Processing



Kafka Streams at Scale

- ± 500M rows of SQL data
- ± 50 joins
- 500 topics
- 400 Gb of Kafka Data
- 300 Gb of RocksDb data
- Building a complete replica from scratch takes many hours
- After that <100ms latency for changes

Development cycle

- Developing and testing is hard for stateful code
- Starting a new ‘generation’ is costly
- Contaminated data might show up

Conclusions

- Went into production early June
- Generally behaves well (aside from some glitches)
- Kafka Streams is in a lot better state than a few months ago


NOW THE FUN BEGINS
quickmeme.com

Elasticsearch

- Add unstructured search to our application
- Reduces load on our source databases
- Users expect google-like interfaces

Neo4J

- Graph Database
- Some analytics are much easier to express in terms of graphs

Firebase Realtime

- Real-time database
- ‘Backend As a Service’
- Essentially one big JSON document
- Very easy to use client libraries for web and mobile
- Safe to develop

Caches

- We can use our streaming engine to update caches

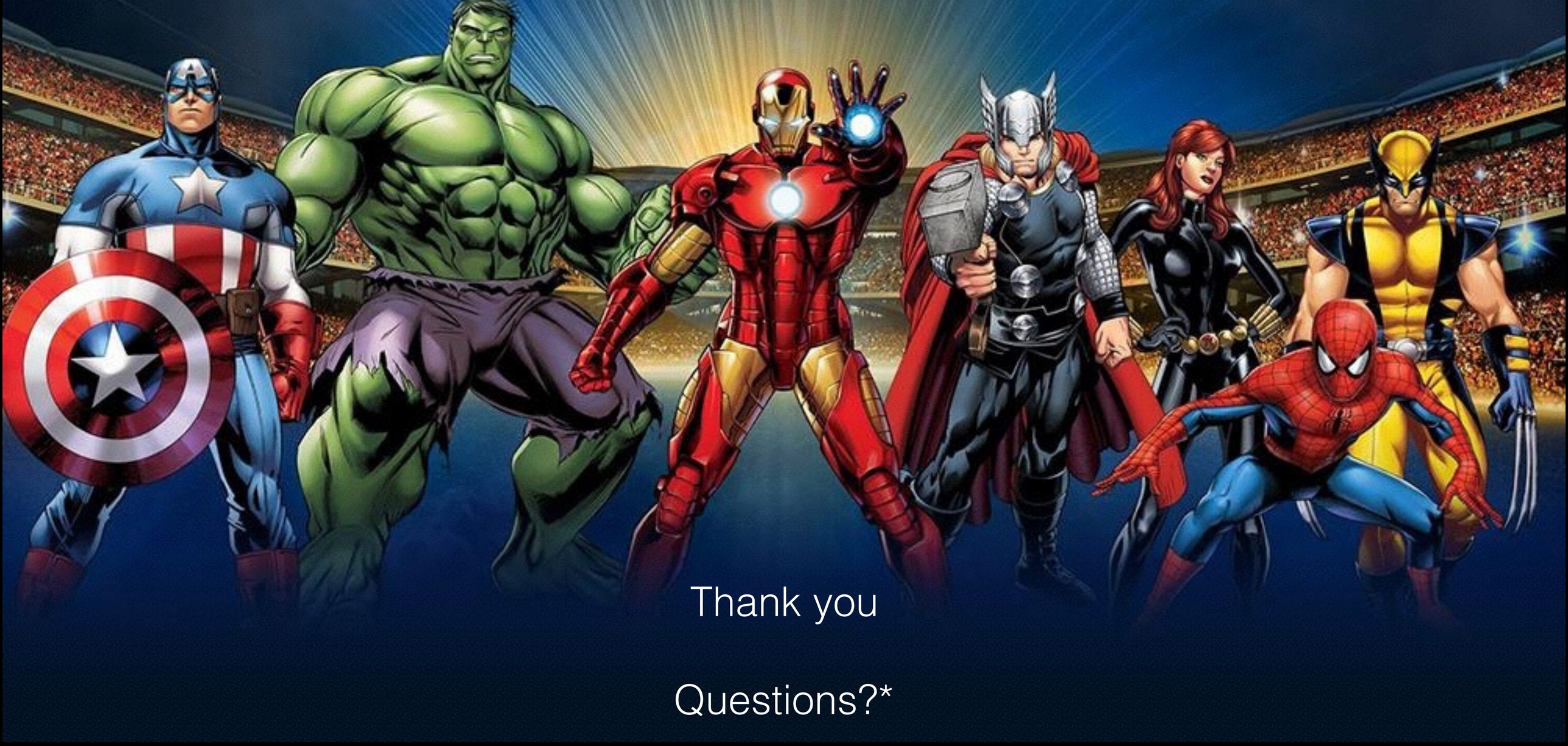
Push to clients

- We can push data to clients in real time

Things I ignored

- Change Capture (“Change Data Capture: The Magic Wand We Forgot”)
- Eventual consistency
- Kafka compaction / partitioning


DATABASES ARE
PEOPLE
Recite
made with ♥ @ Recite.com



Thank you

Questions?*

* Any question that is not: “Why don’t you use Postgres? Postgres can do anything”