

REAL-TIME STREAMING AND ANALYTICS AT SCALE WITH APACHE KAFKA AND APACHE IGNITE

JUNE, 2019 / NEW YORK, NY

@DENISMAGDA | @GAMUSSA | #NYCINMEMORY

HELLO 🖐️



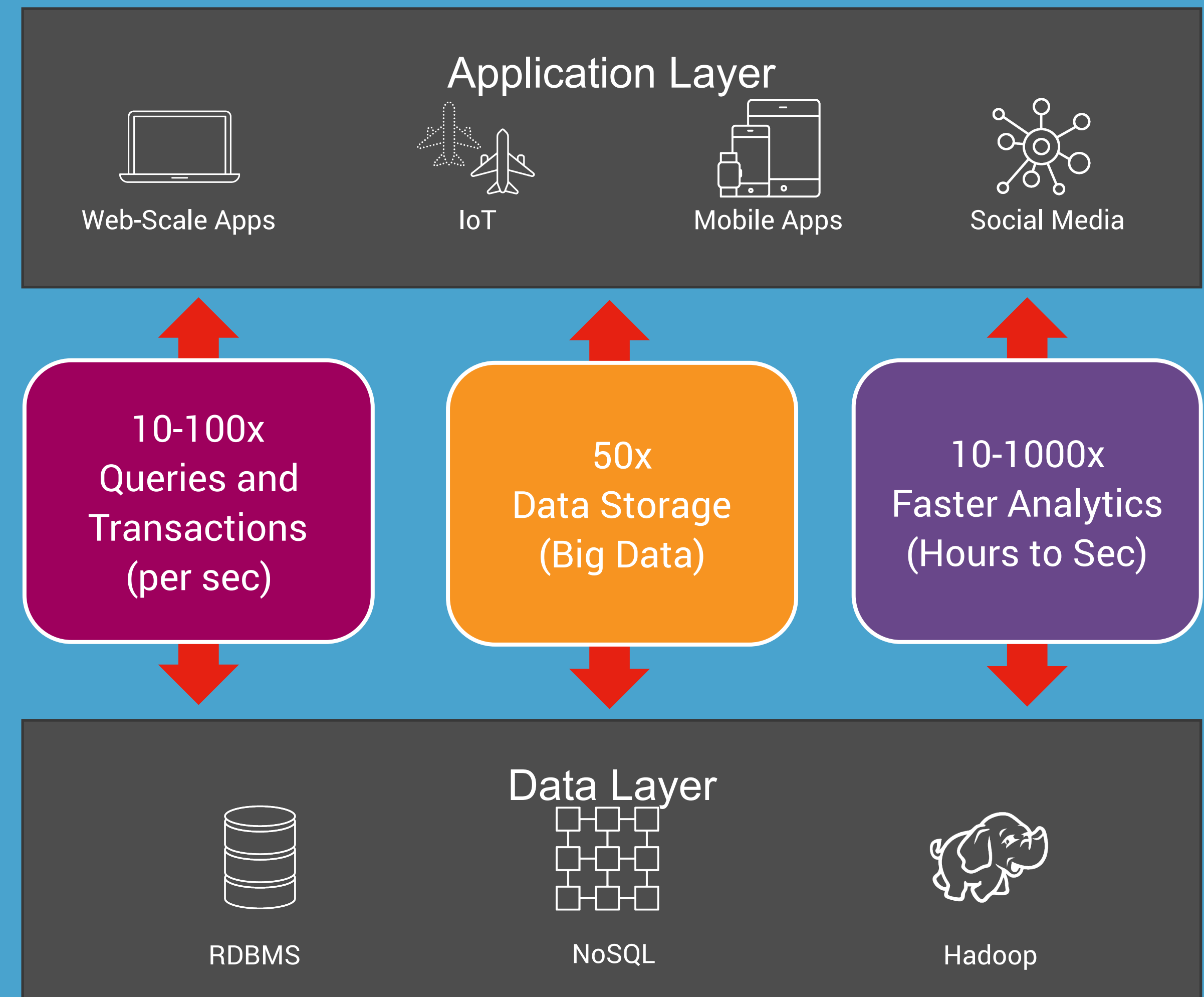
@DENISMAGDA



@GAMUSSA

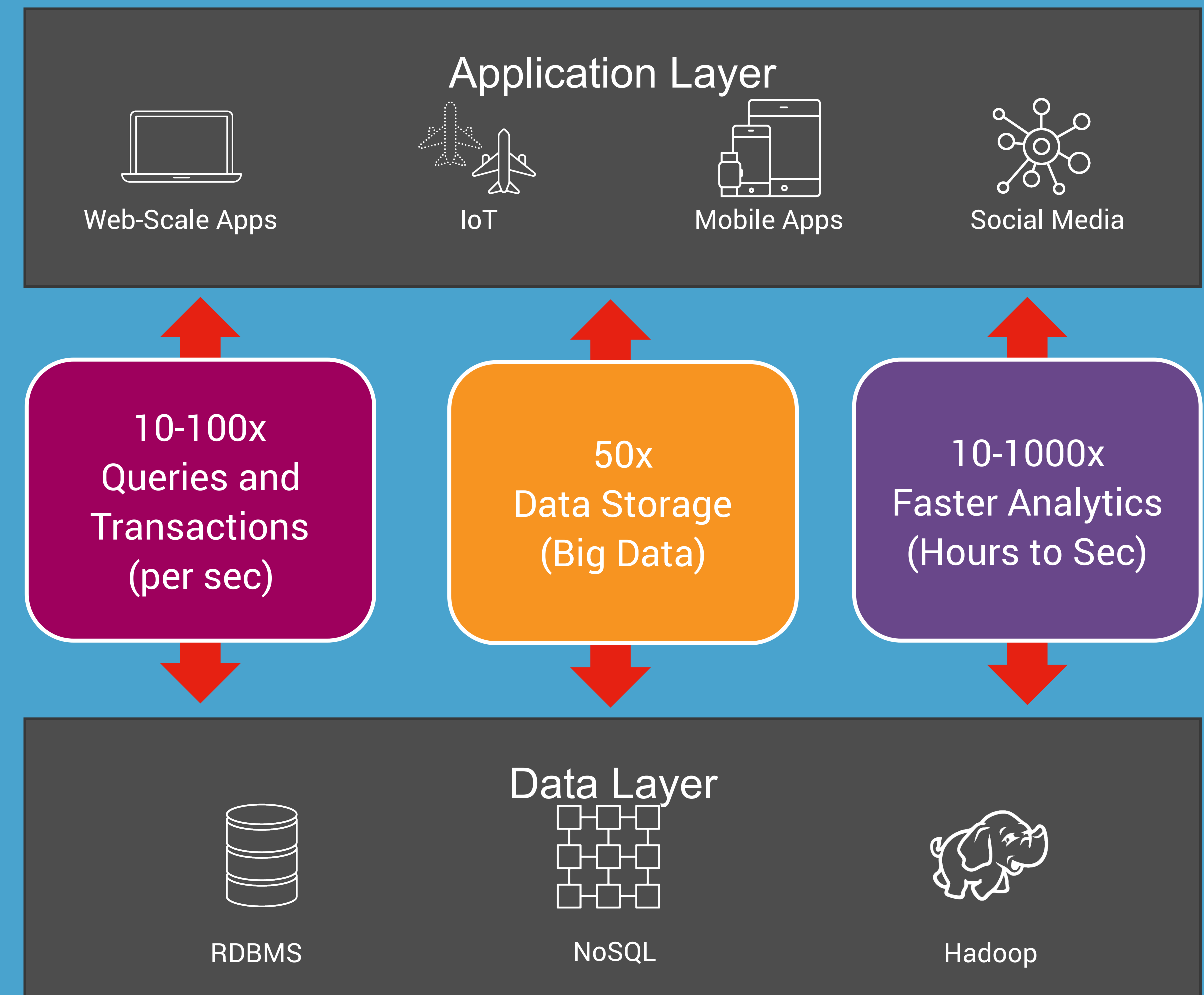
DIGITAL TRANSFORMATION CHALLENGES

DIGITAL TRANSFORMATIONS CHALLENGES



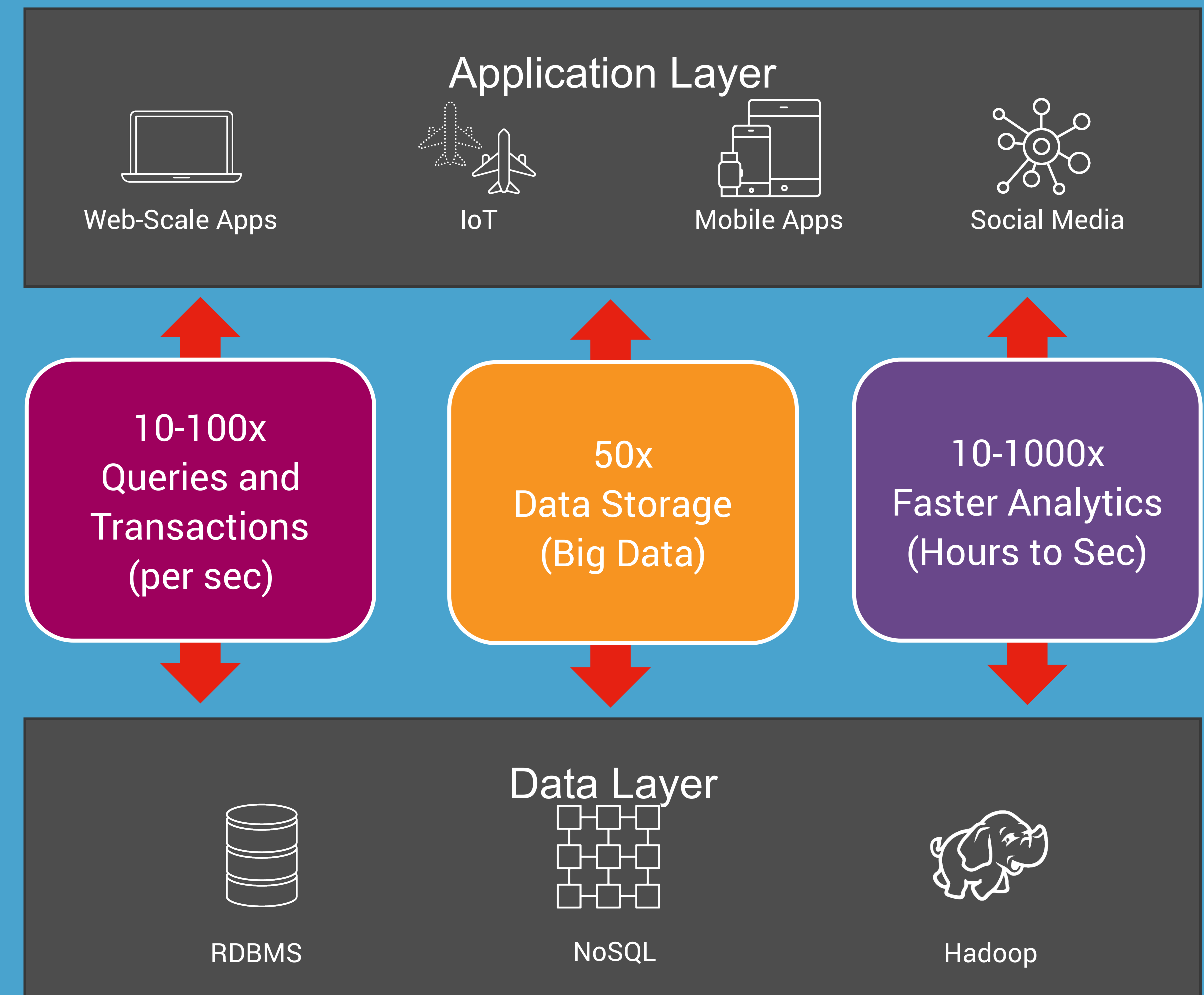
DIGITAL TRANSFORMATIONS CHALLENGES

- **10-100x more queries and transactions**



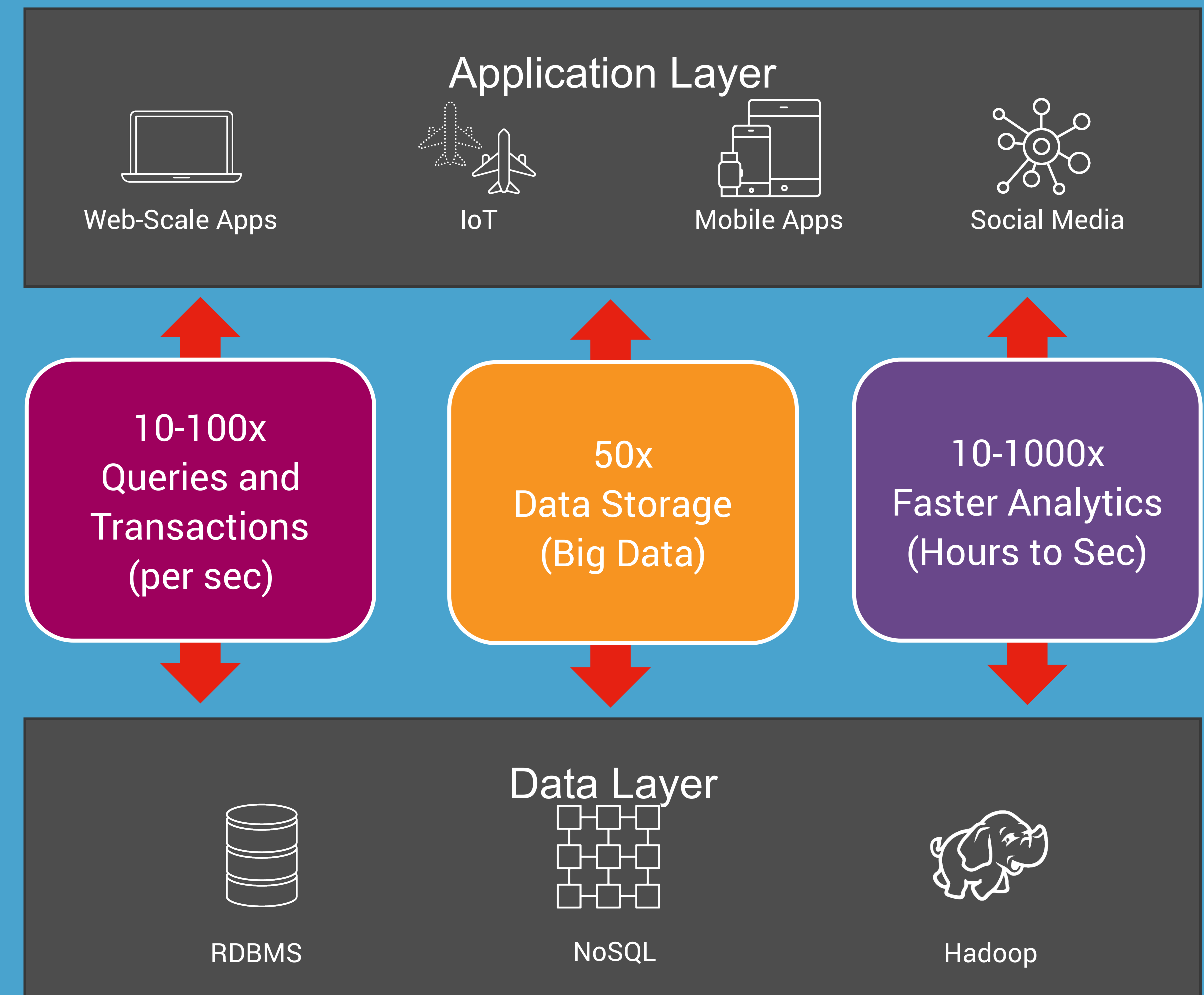
DIGITAL TRANSFORMATIONS CHALLENGES

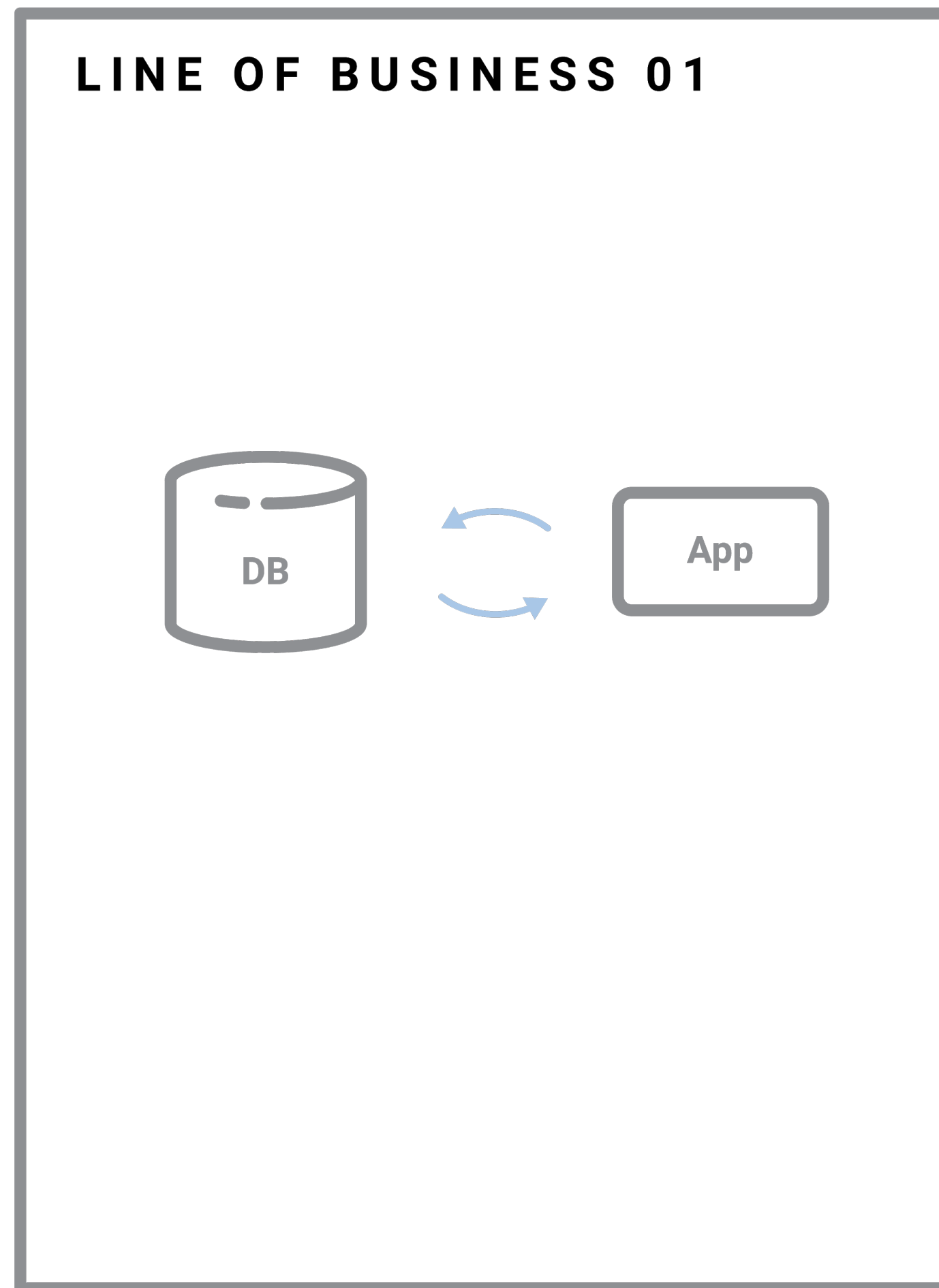
- 10-100x more queries and transactions
- 50x more data today as a decade ago

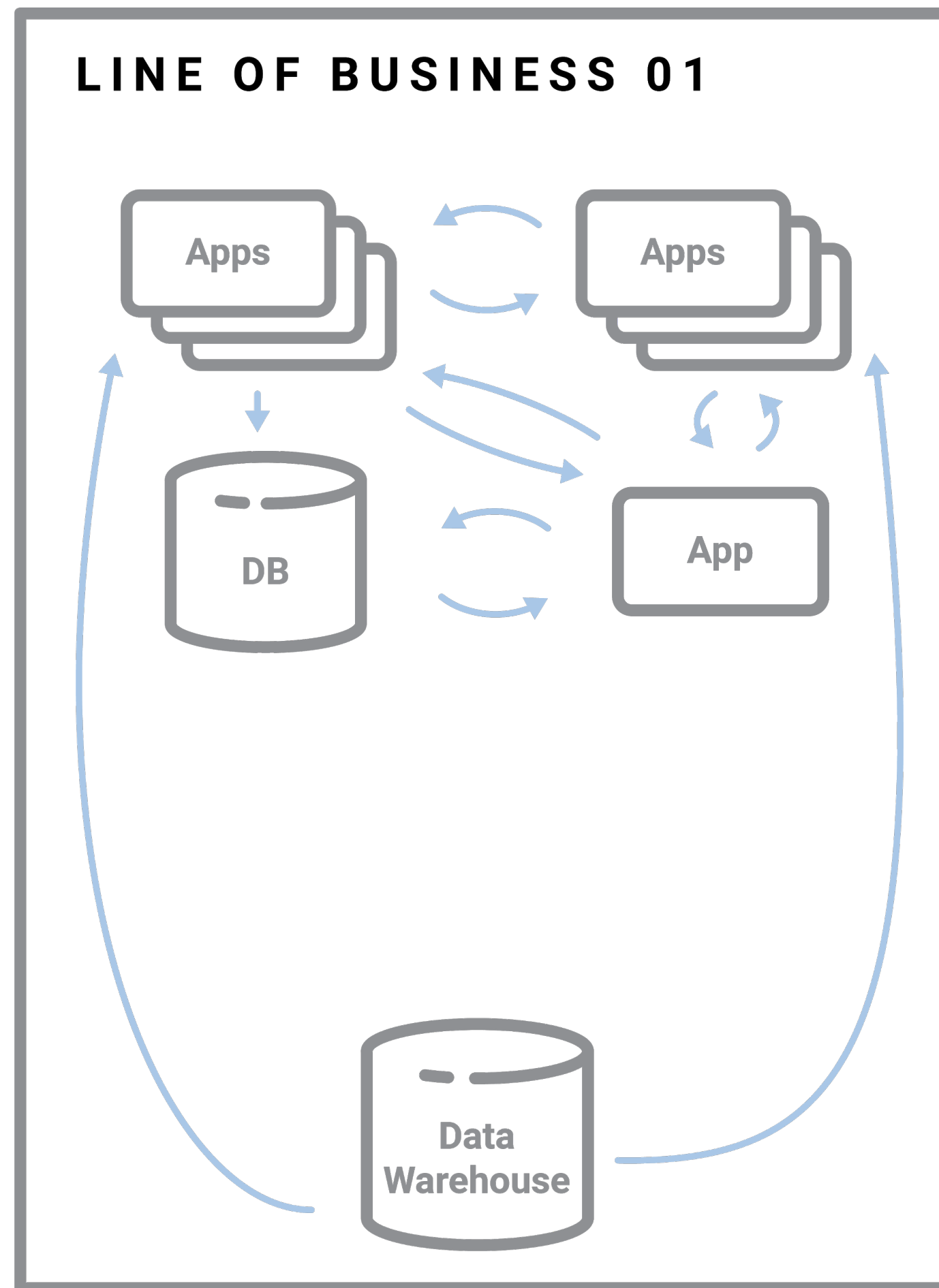


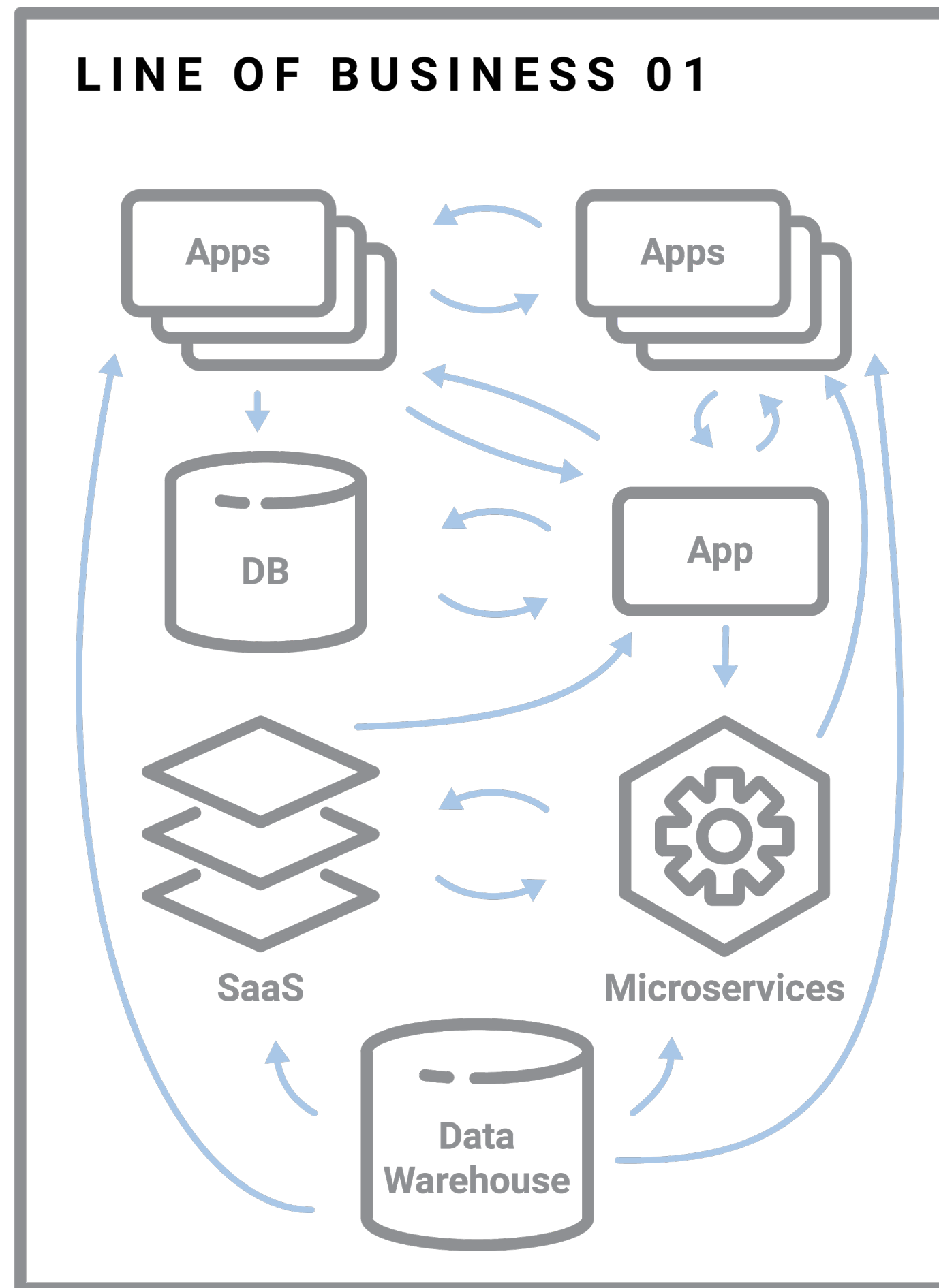
DIGITAL TRANSFORMATIONS CHALLENGES

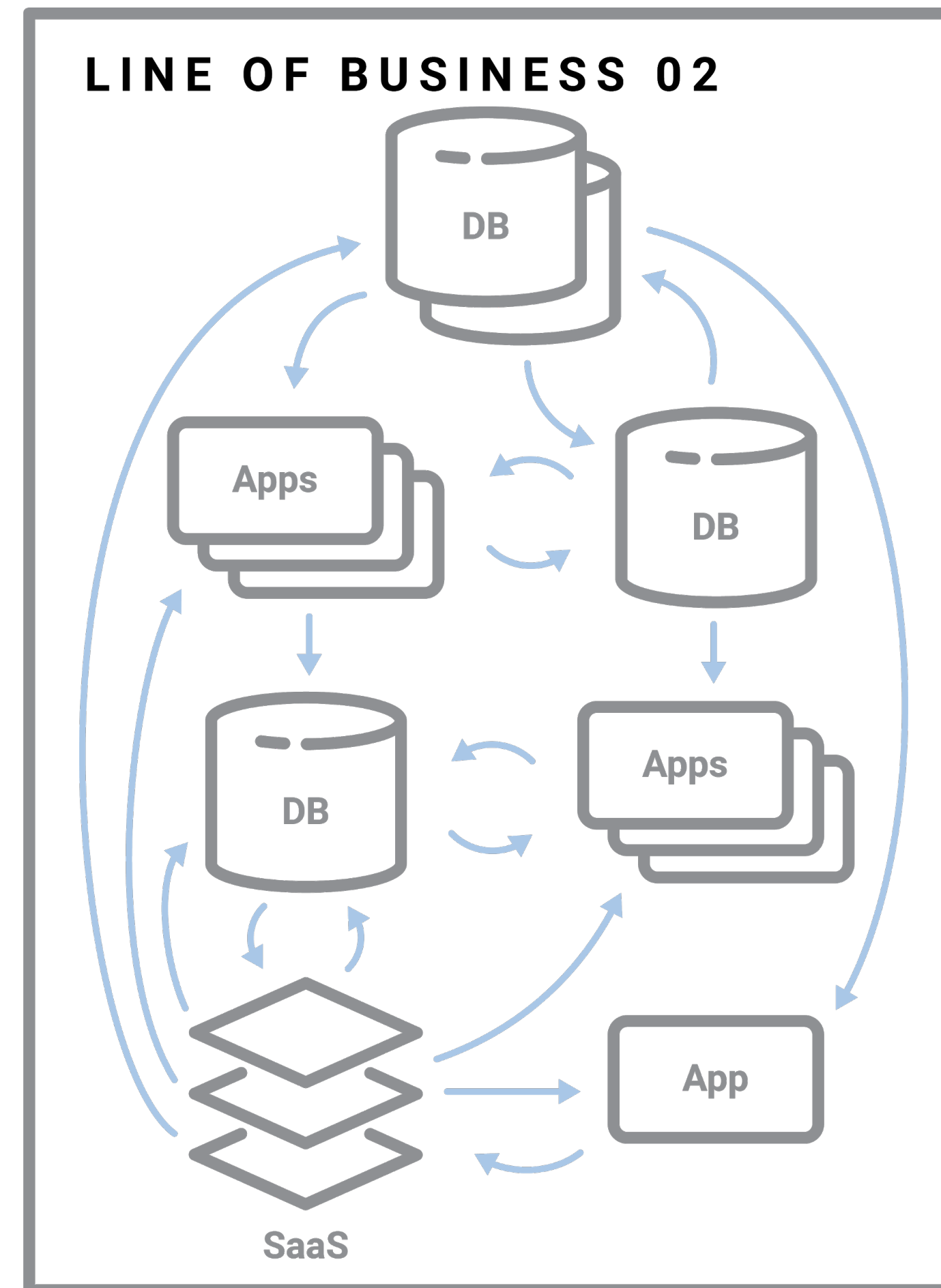
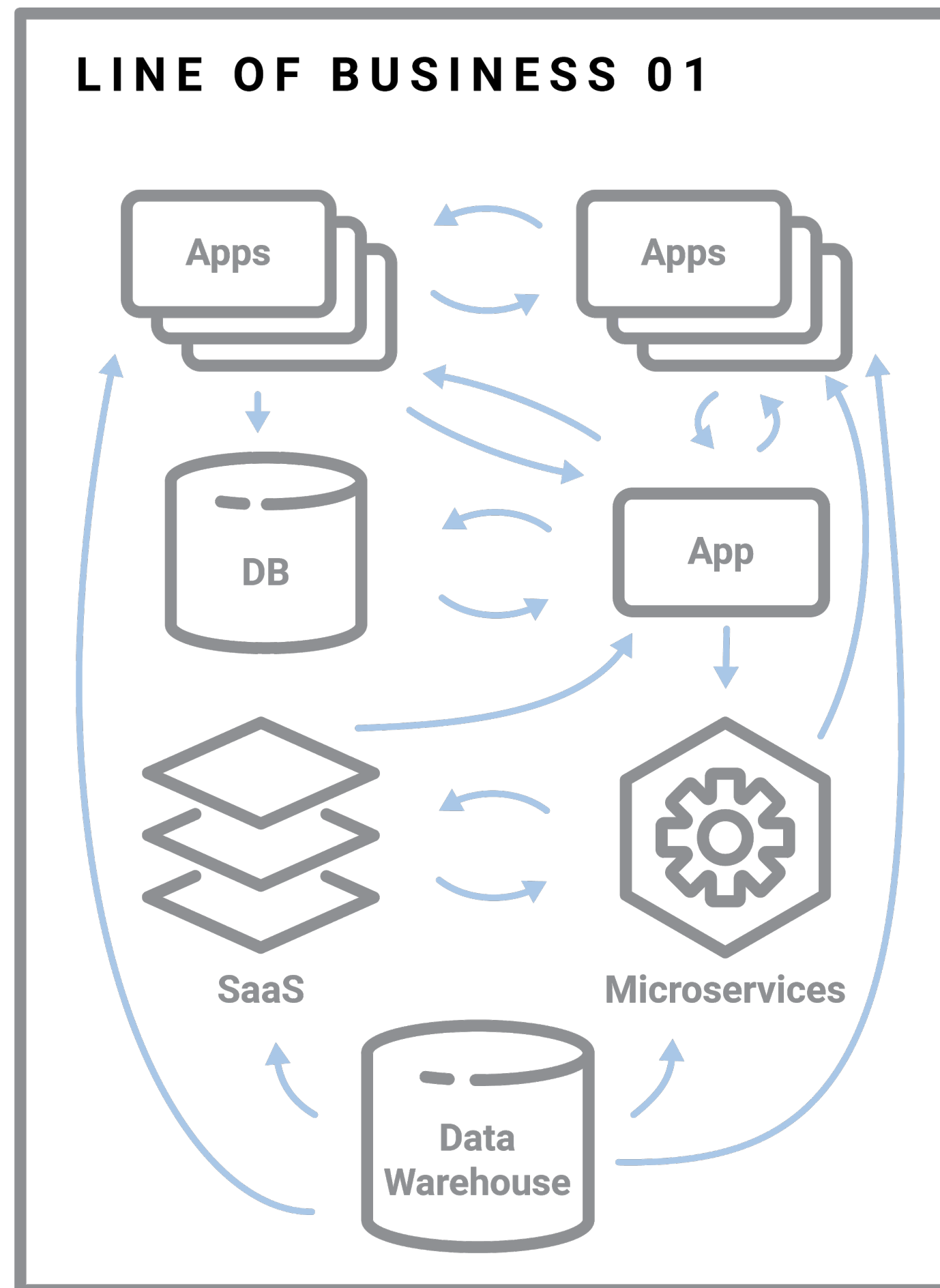
- 10-100x more queries and transactions
- 50x more data today as a decade ago
- Overnight analytics become real-time

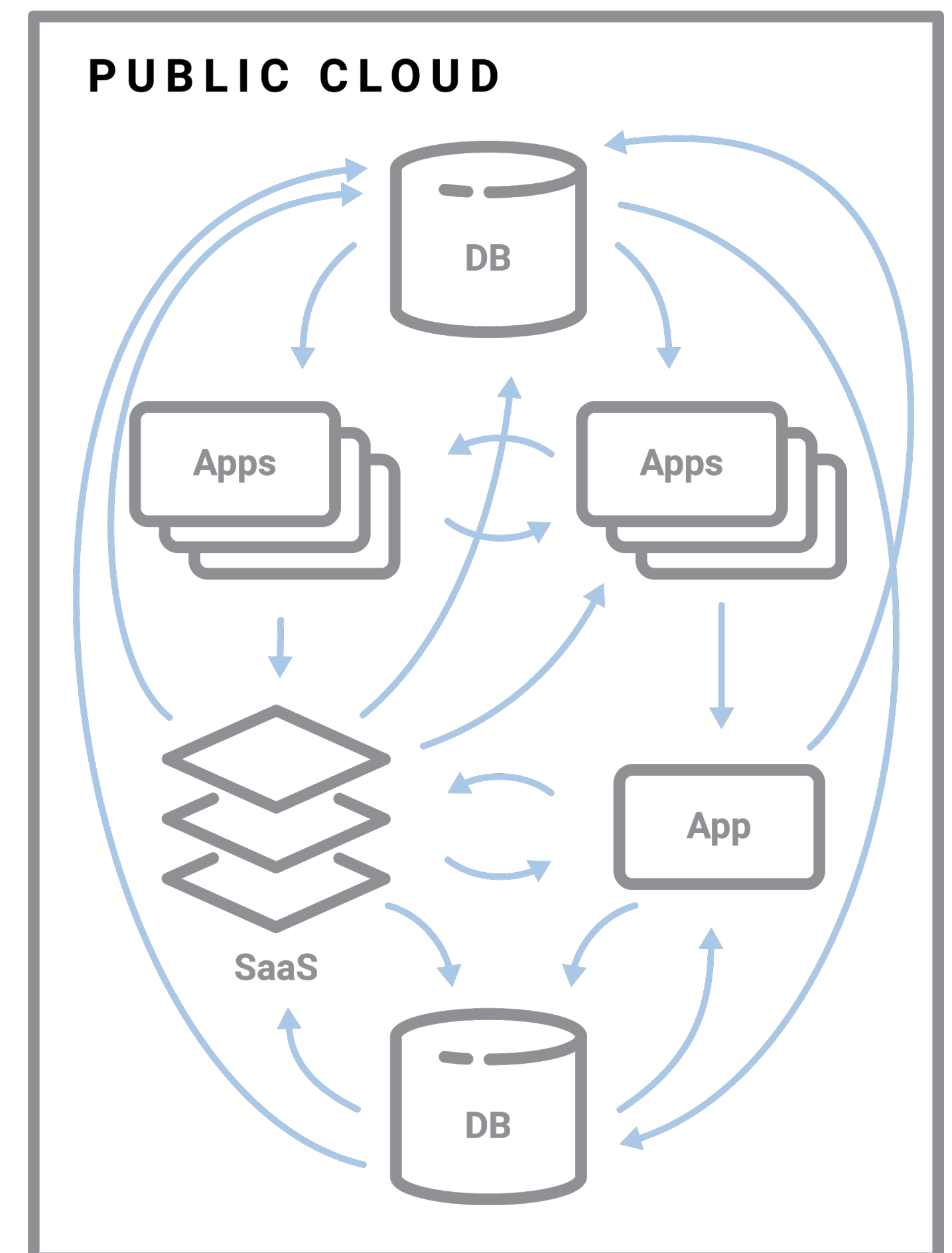
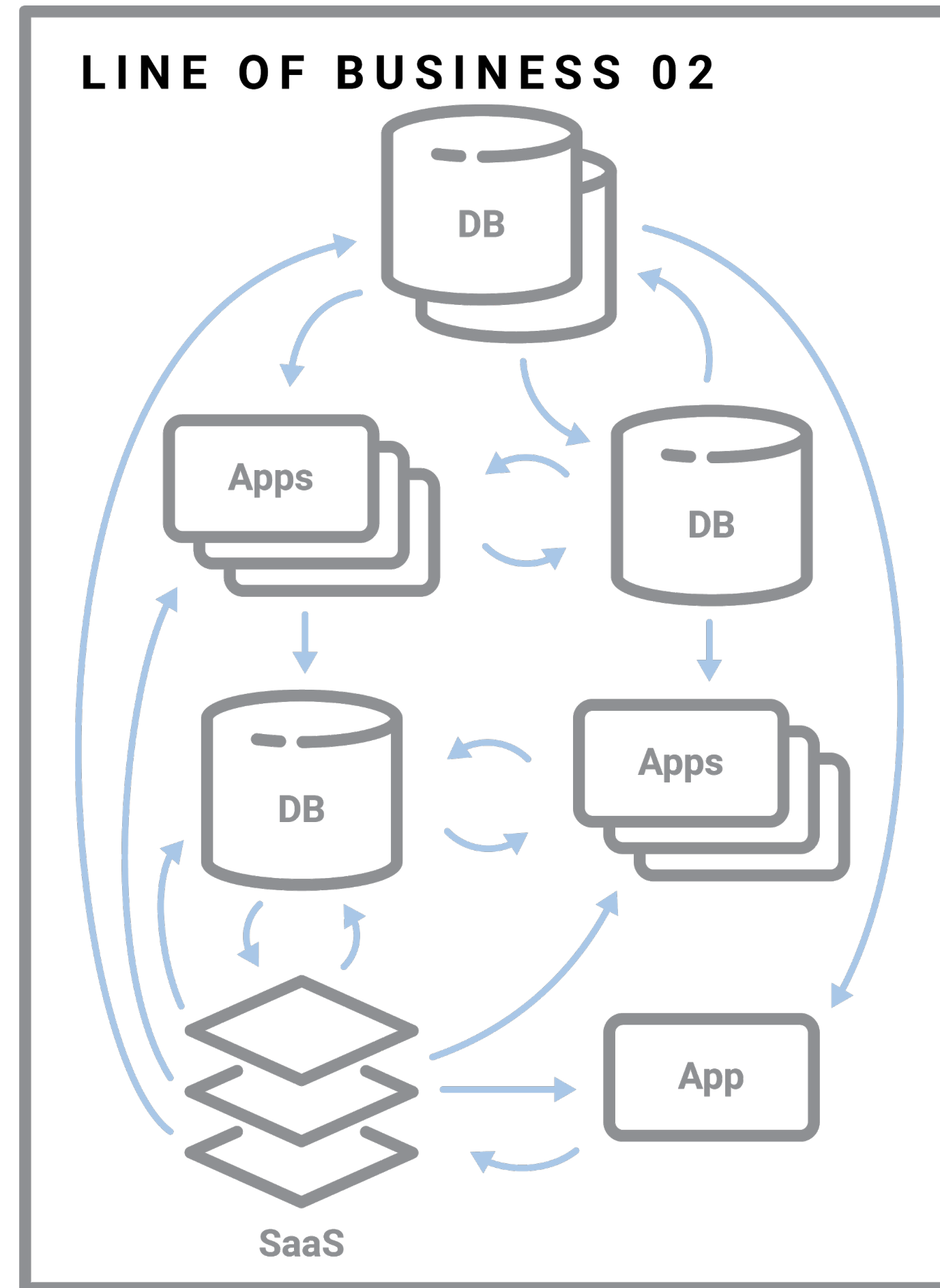
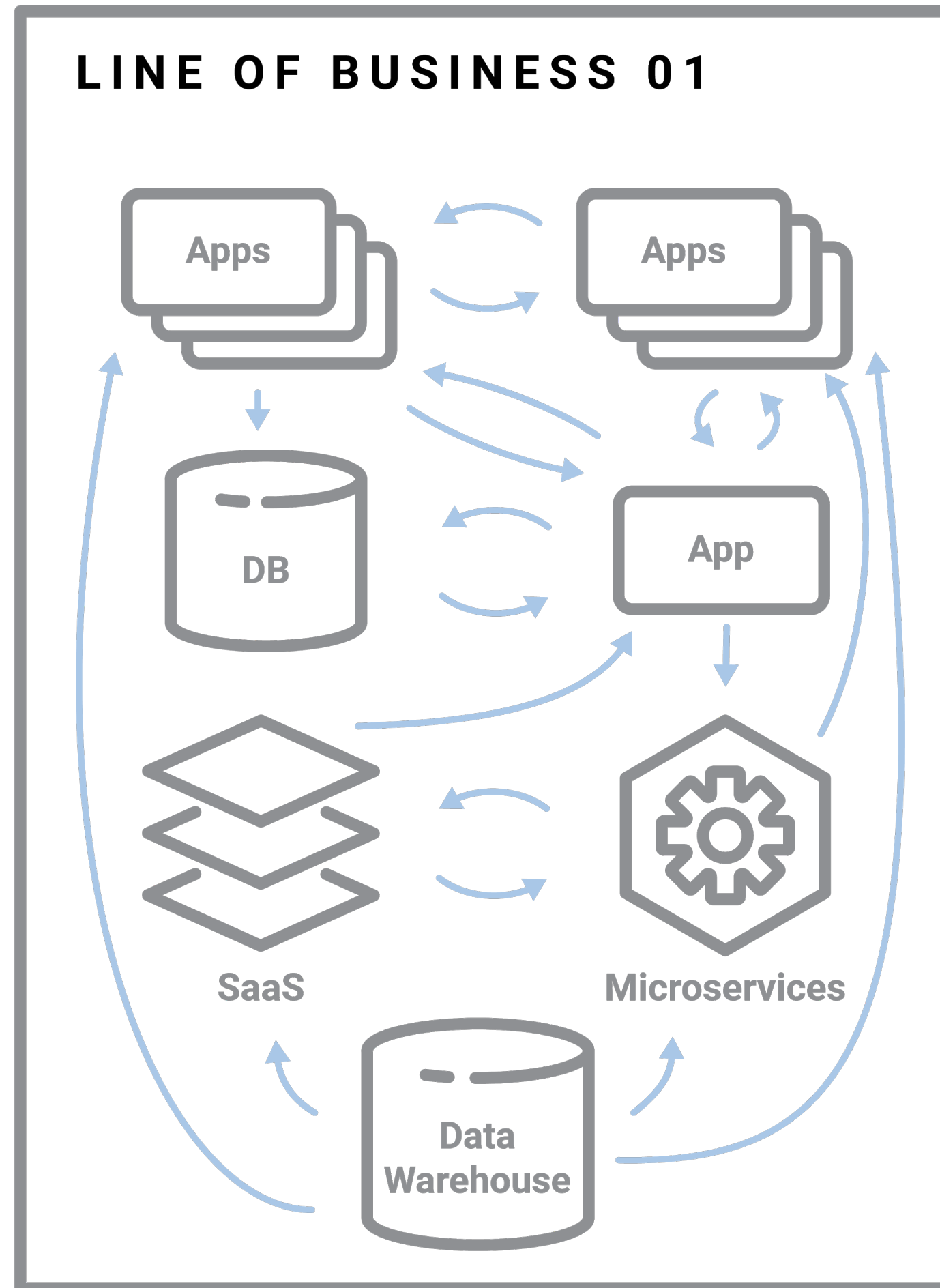


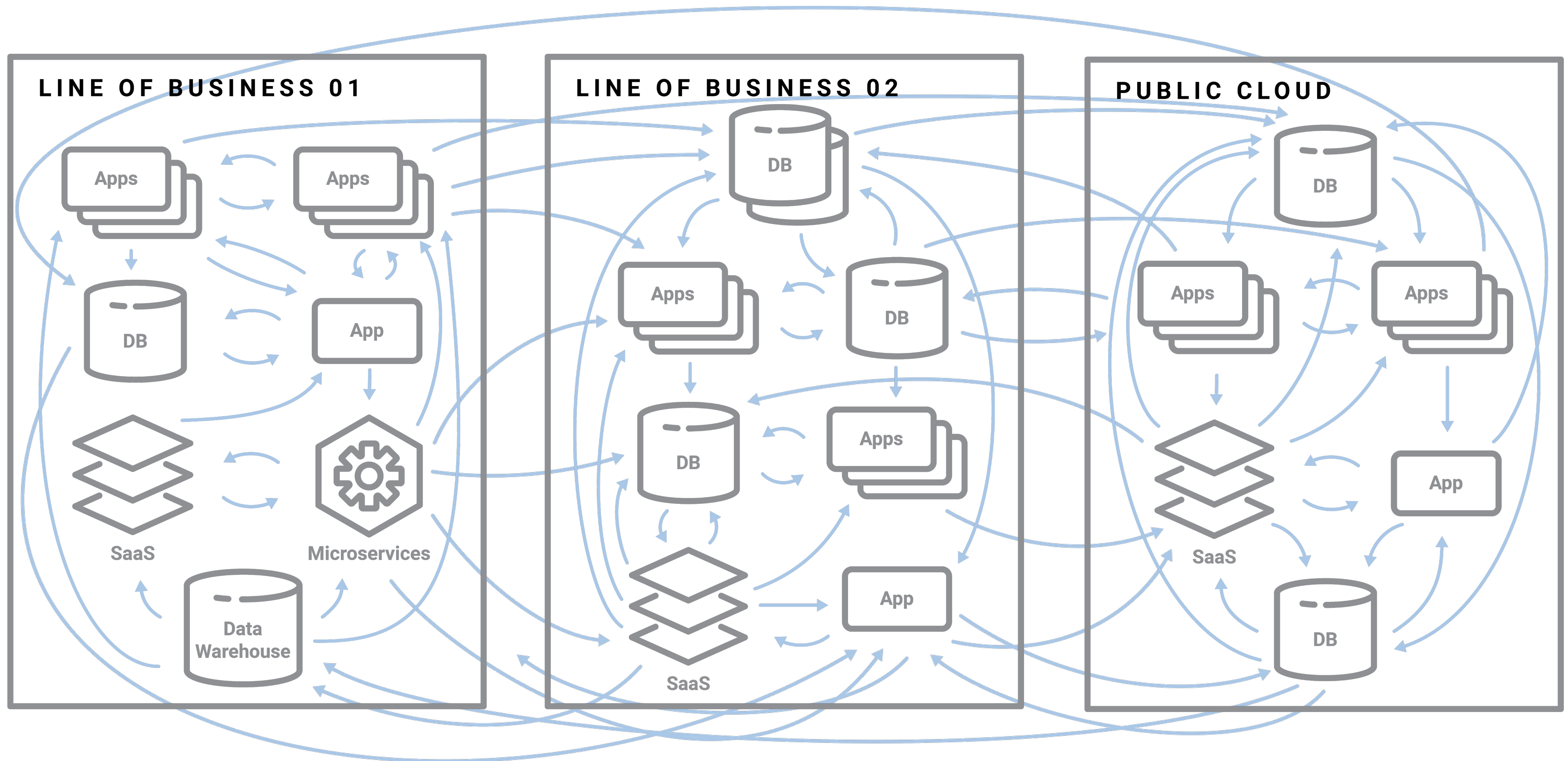


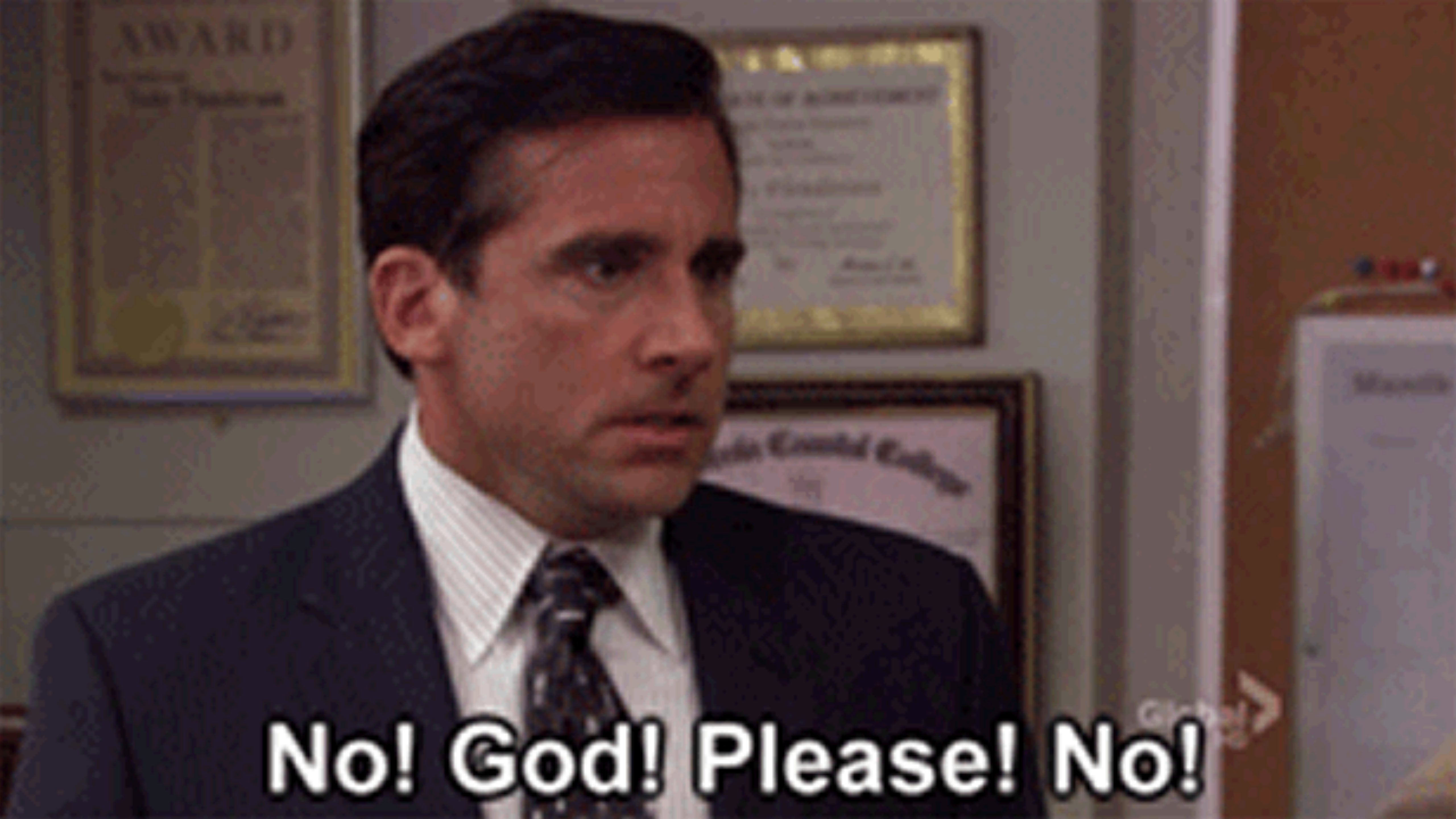




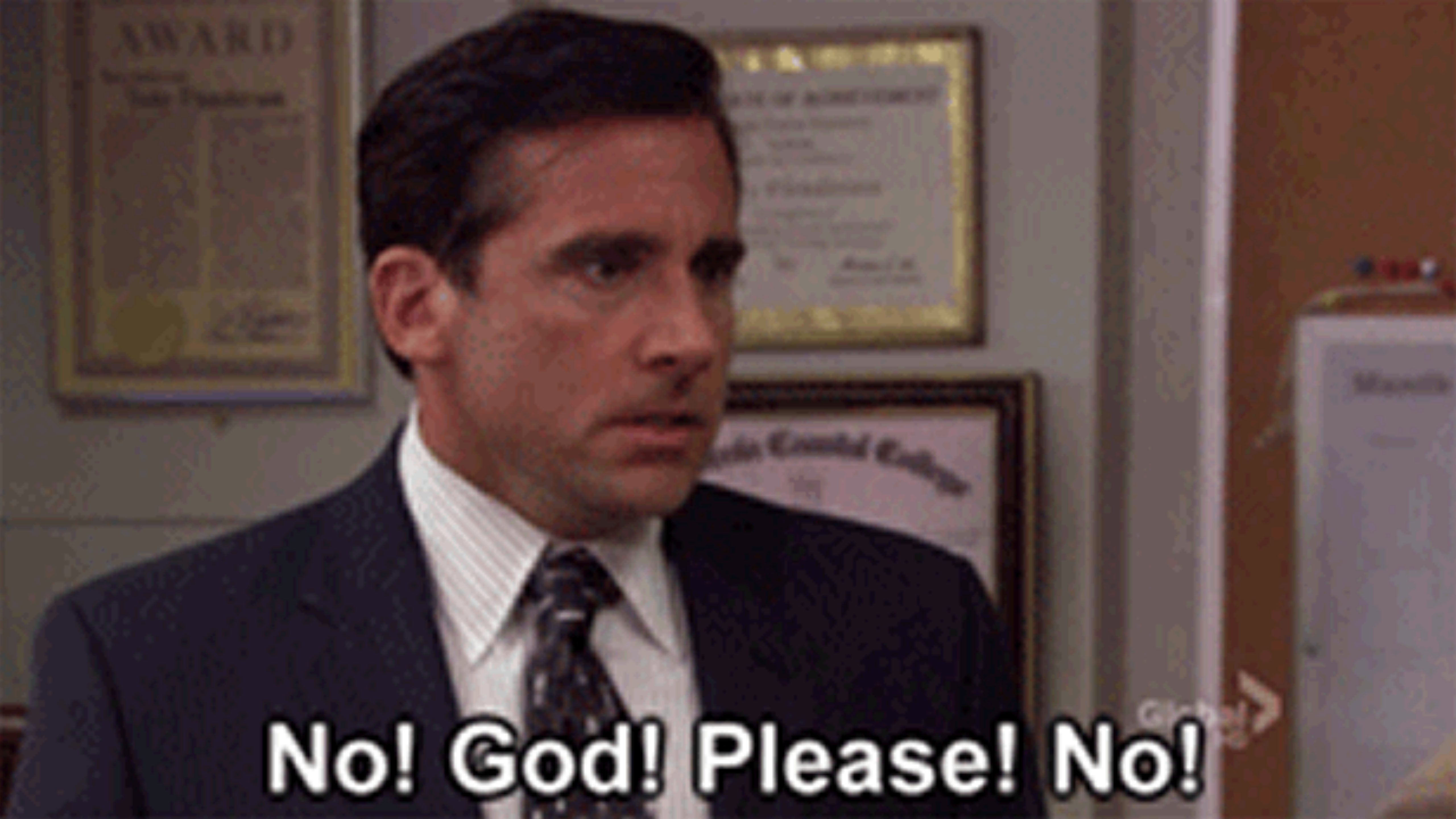






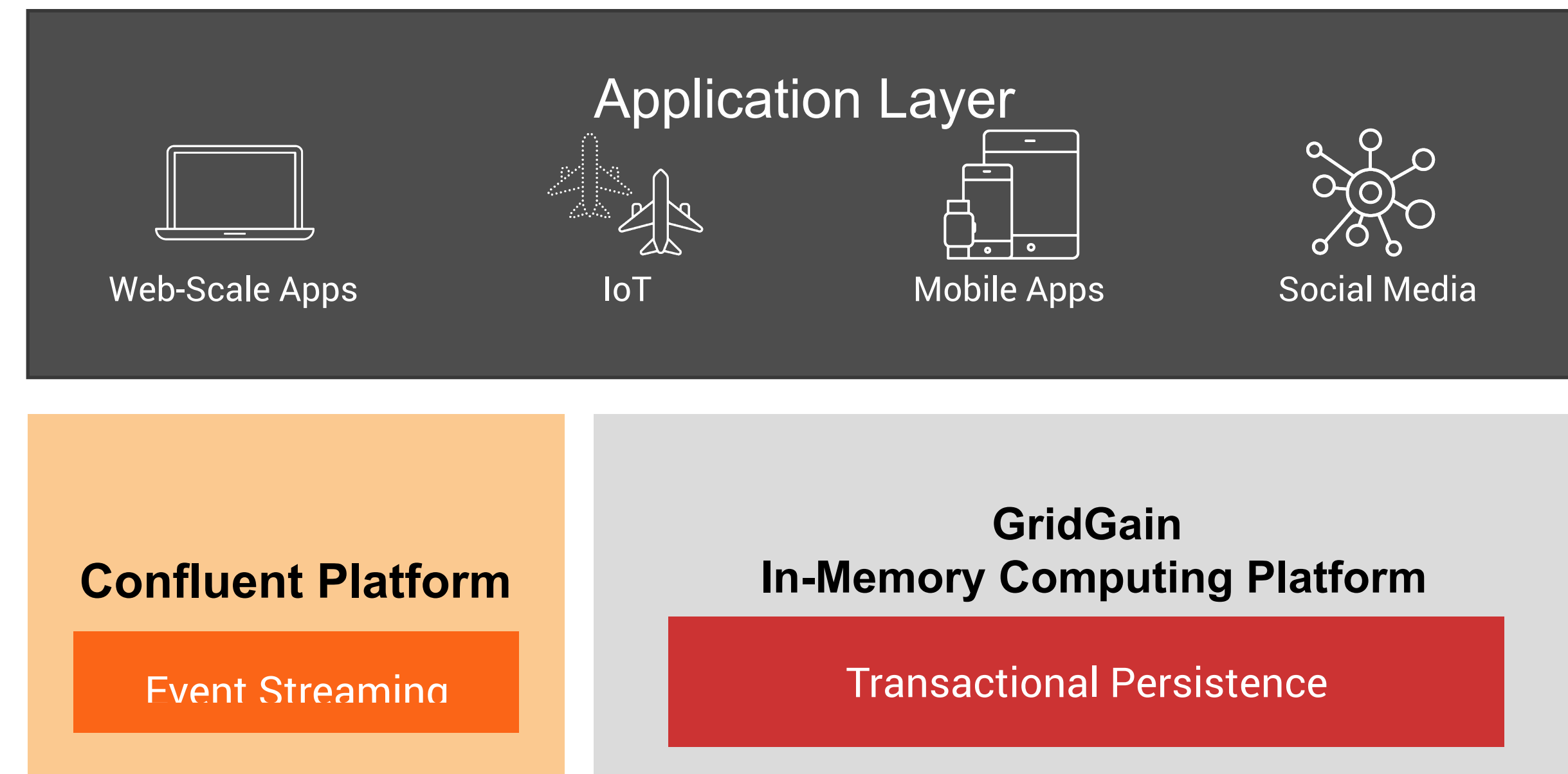


No! God! Please! No!



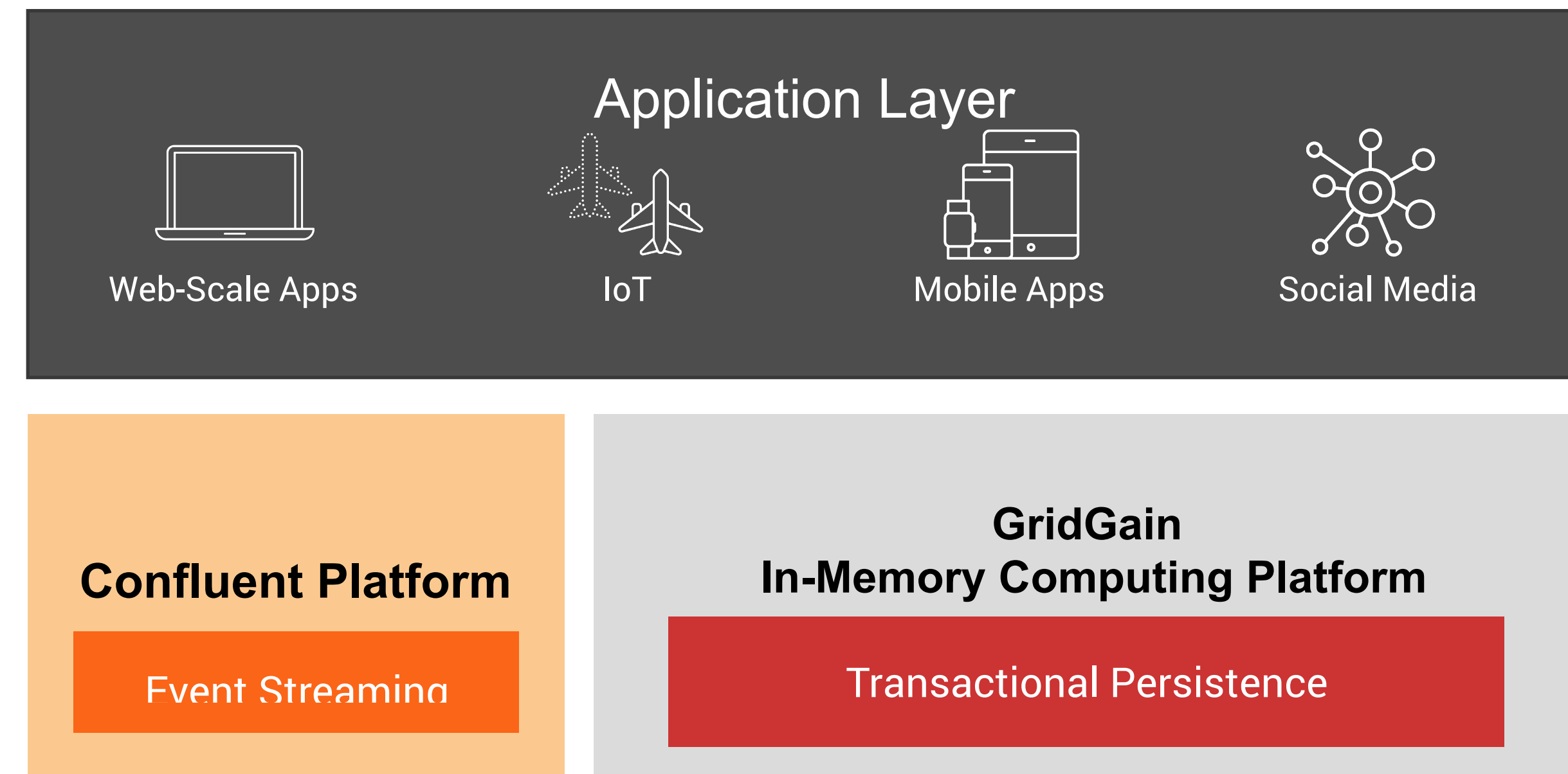
No! God! Please! No!

IN-MEMORY COMPUTING AND STREAM PROCESSING



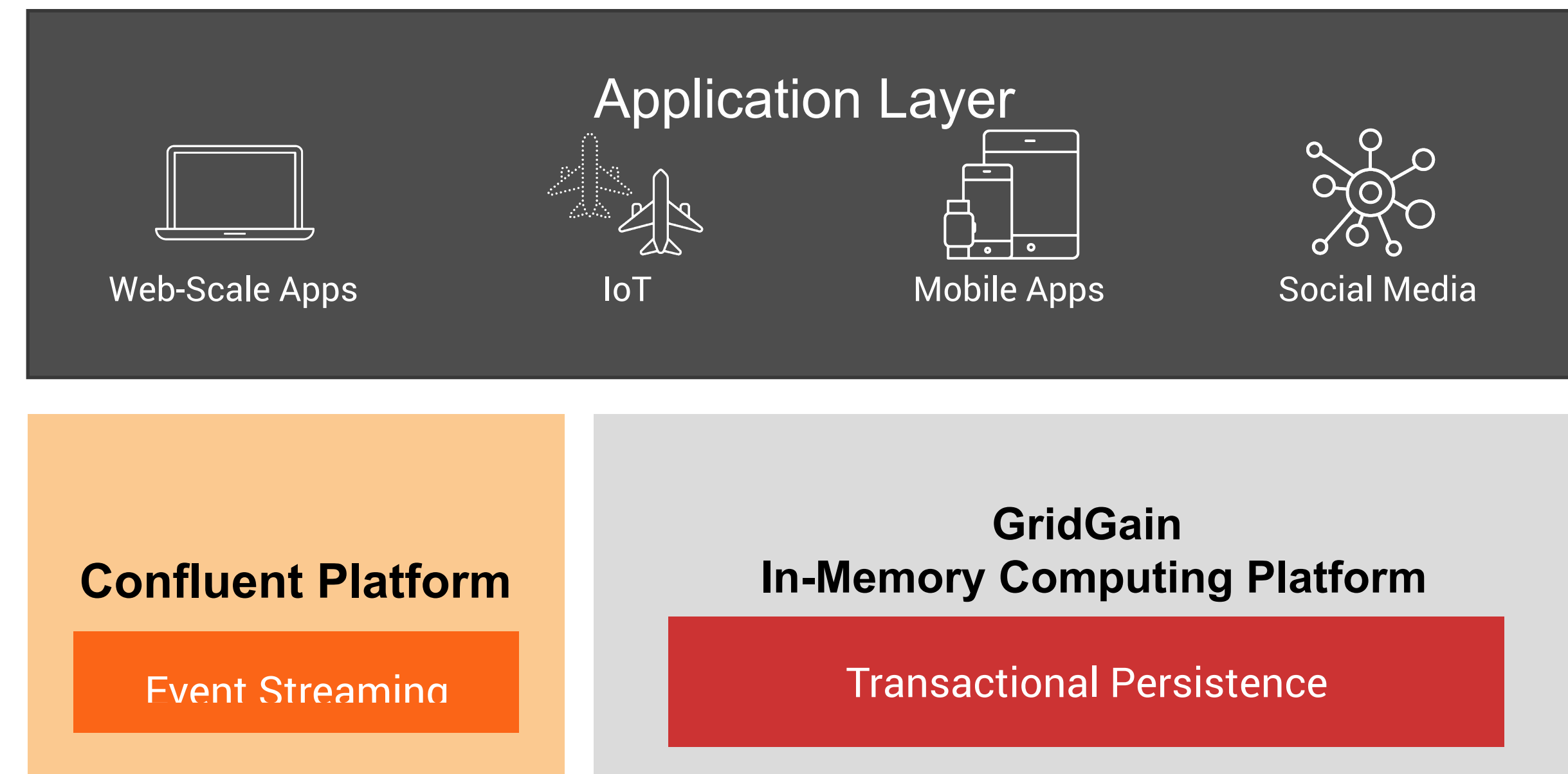
IN-MEMORY COMPUTING AND STREAM PROCESSING

- Performance and velocity increases



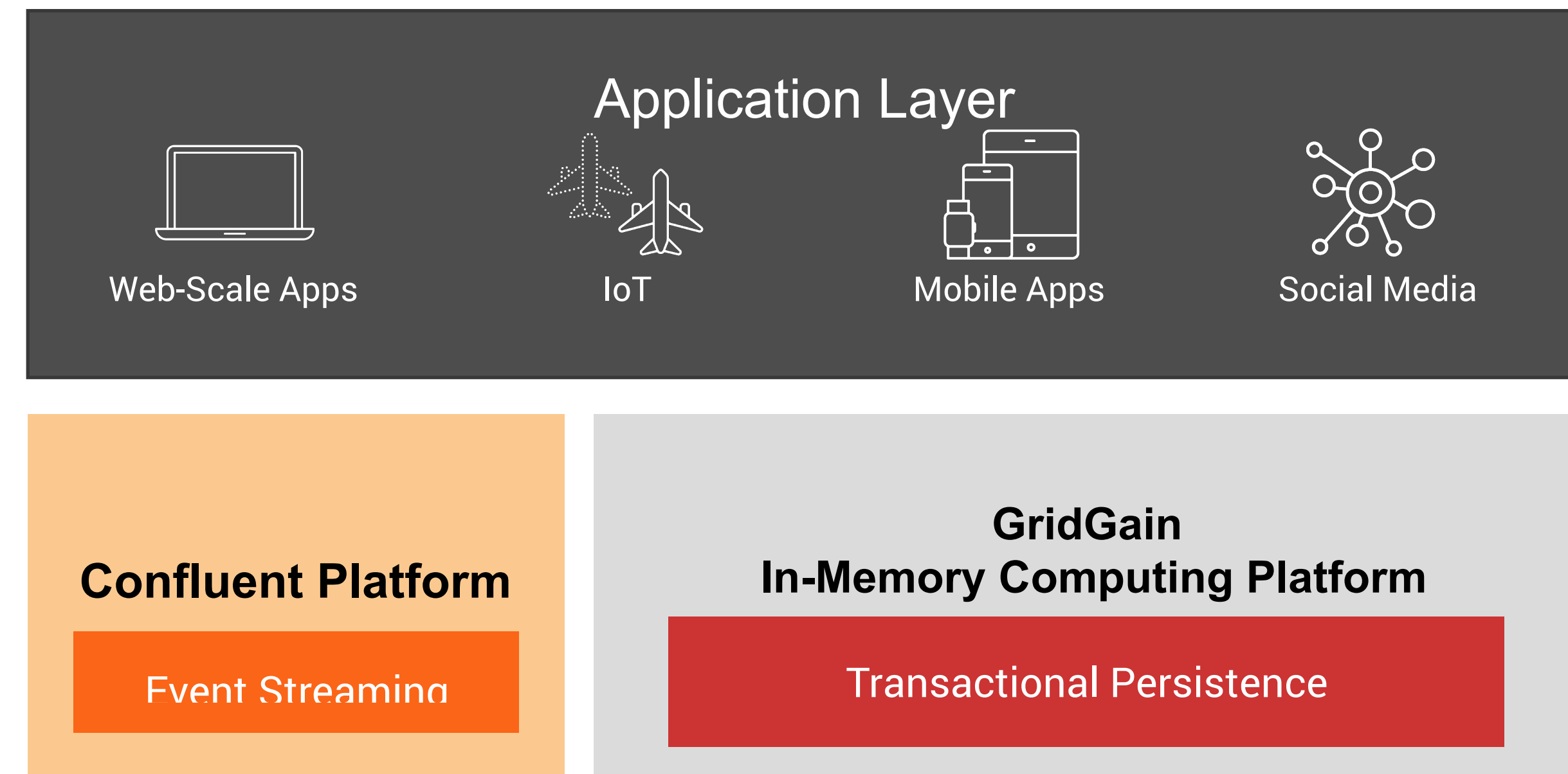
IN-MEMORY COMPUTING AND STREAM PROCESSING

- Performance and velocity increases
- Scalability up to **petabytes** of data

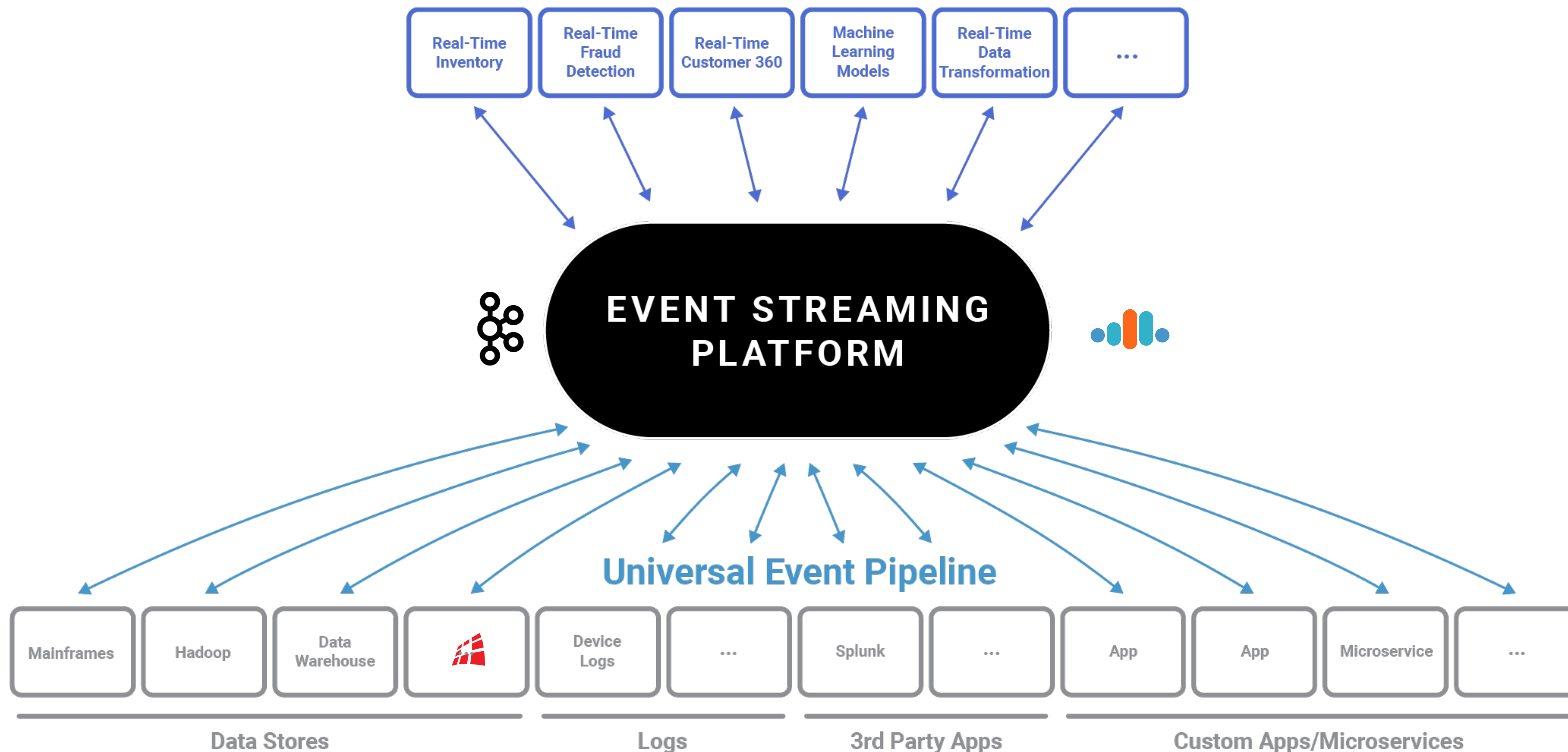


IN-MEMORY COMPUTING AND STREAM PROCESSING

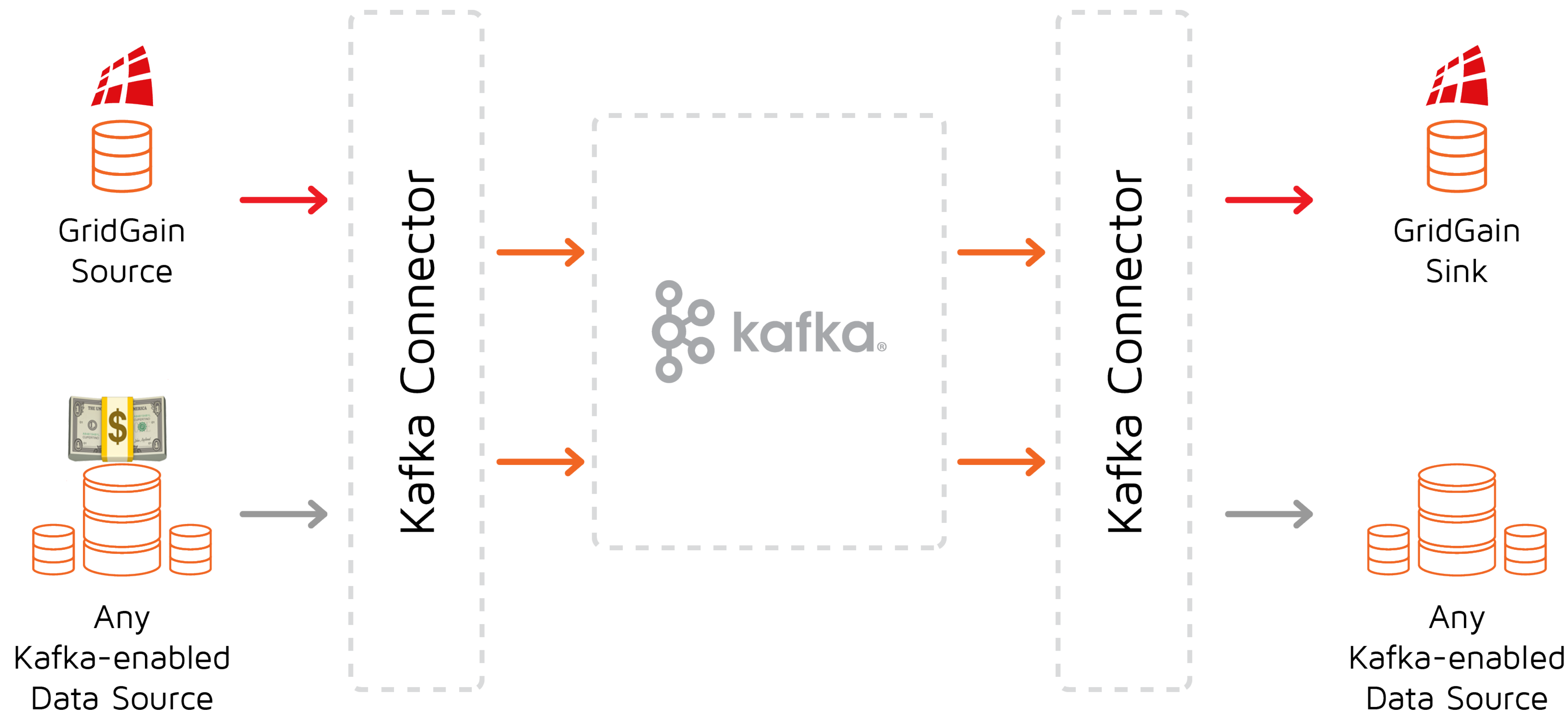
- Performance and velocity increases
- Scalability up to **petabytes** of data
- Act faster by **analyzing streams of data**
using **SQL language**



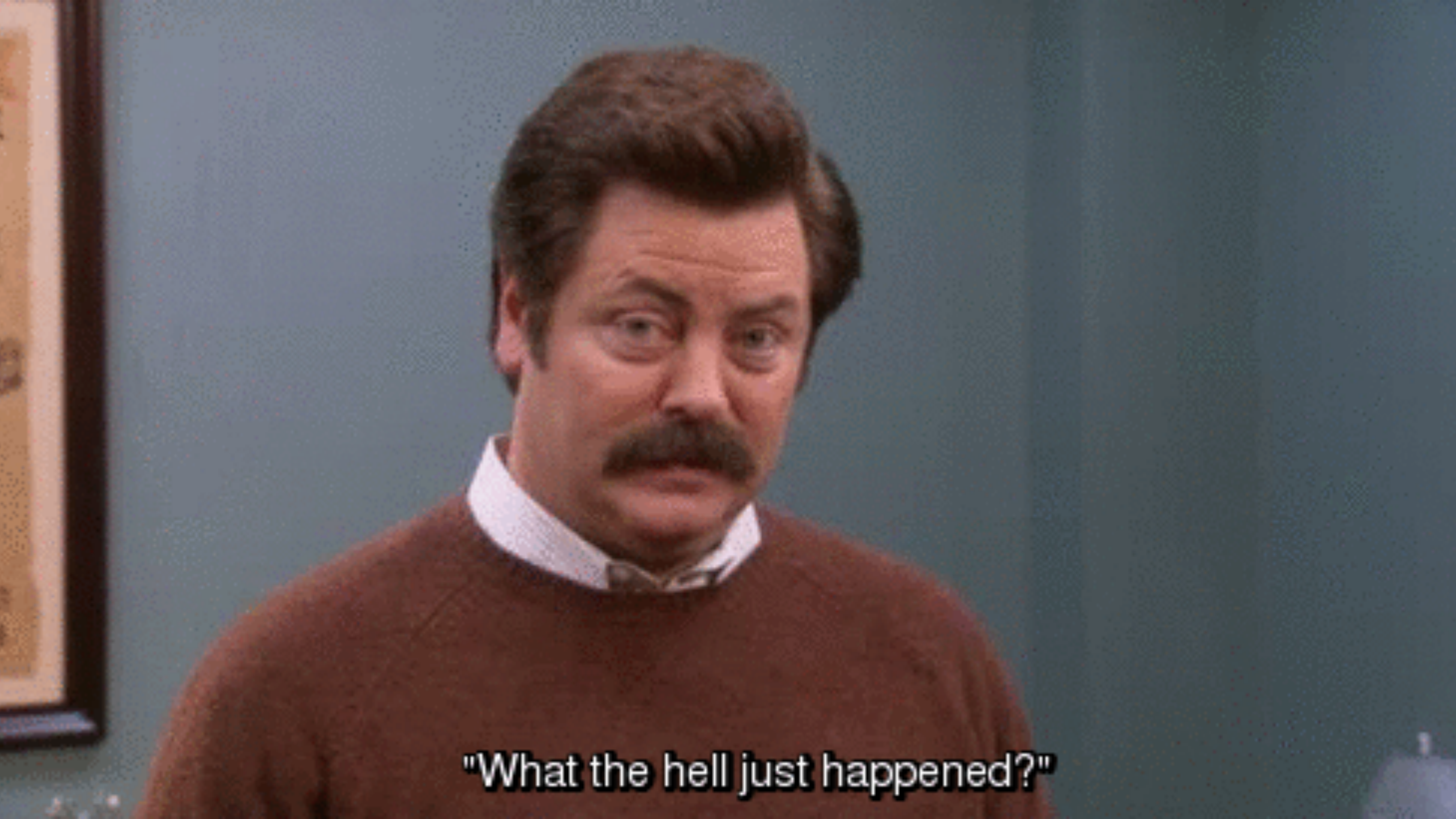
Contextual Event-Driven Apps



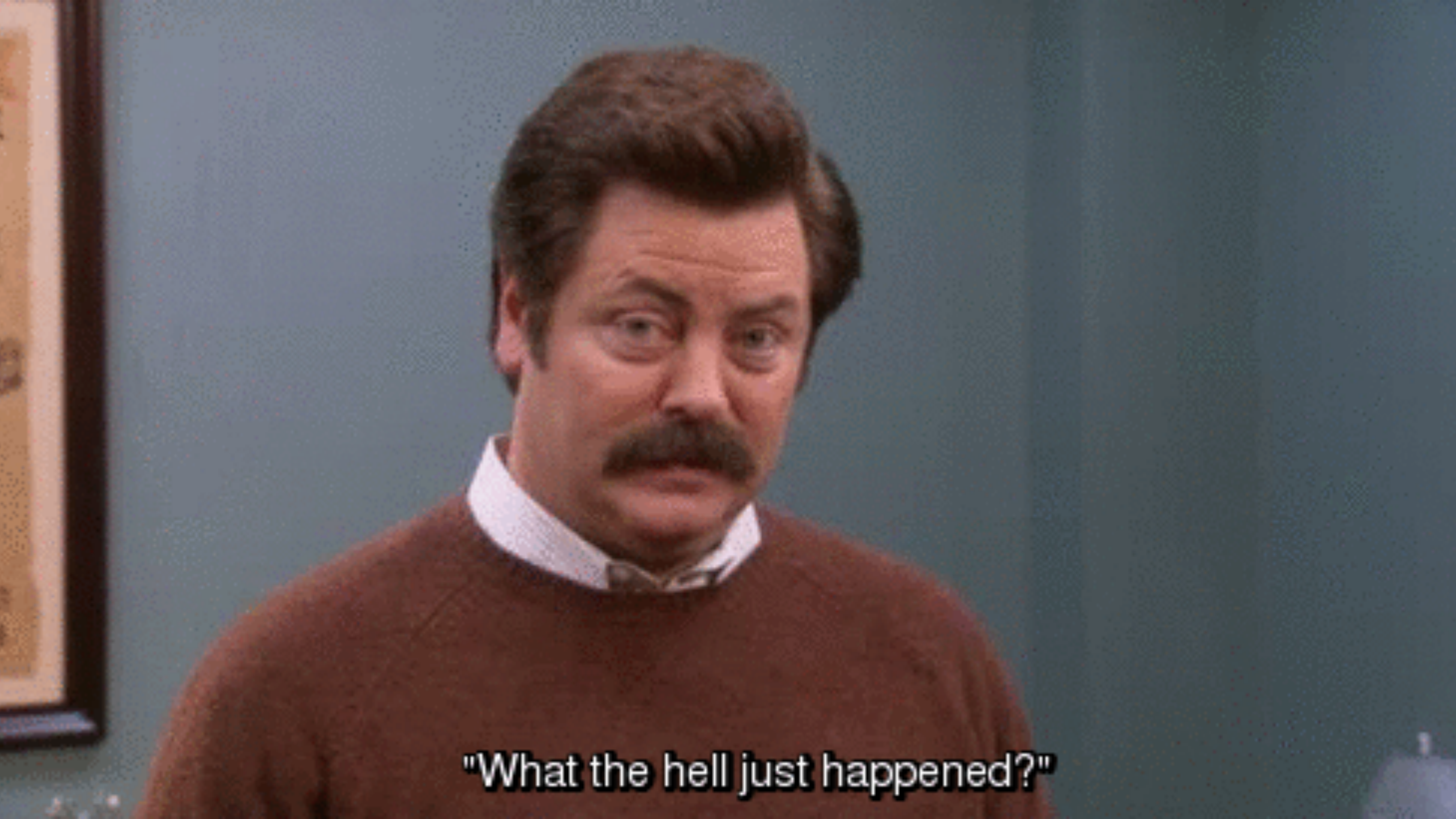
KAPPA ARCHITECTURE: GRIDGAIN AND KAFKA CONNECT



DEMO

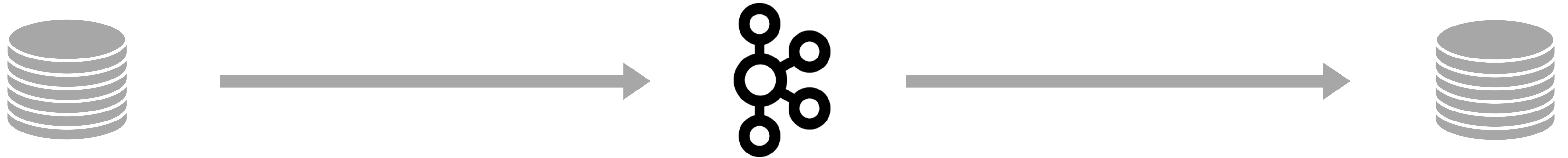


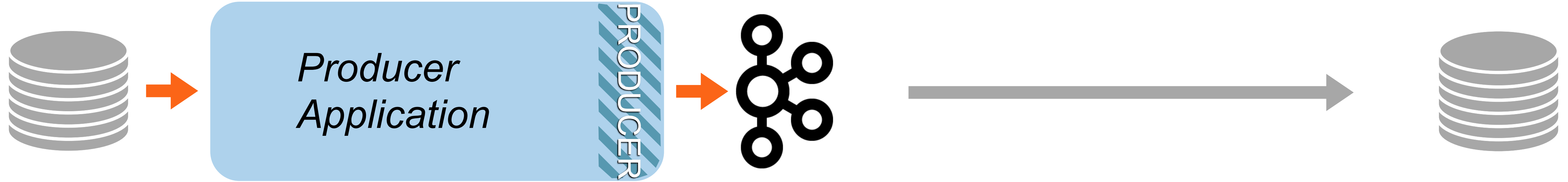
"What the hell just happened?"

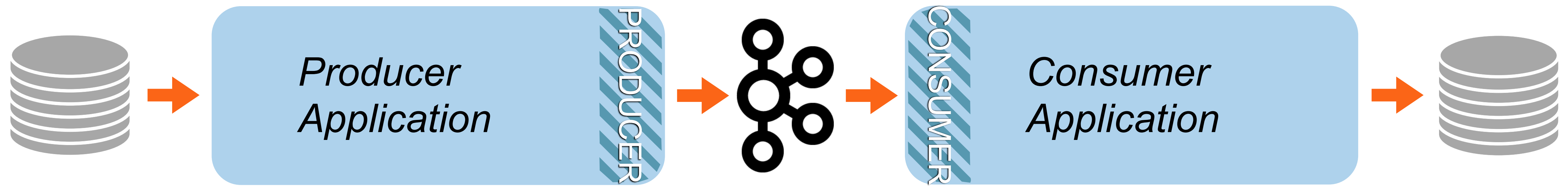


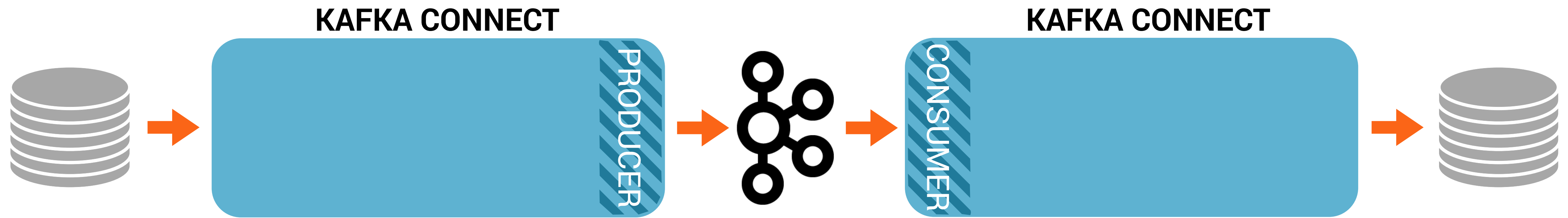
"What the hell just happened?"

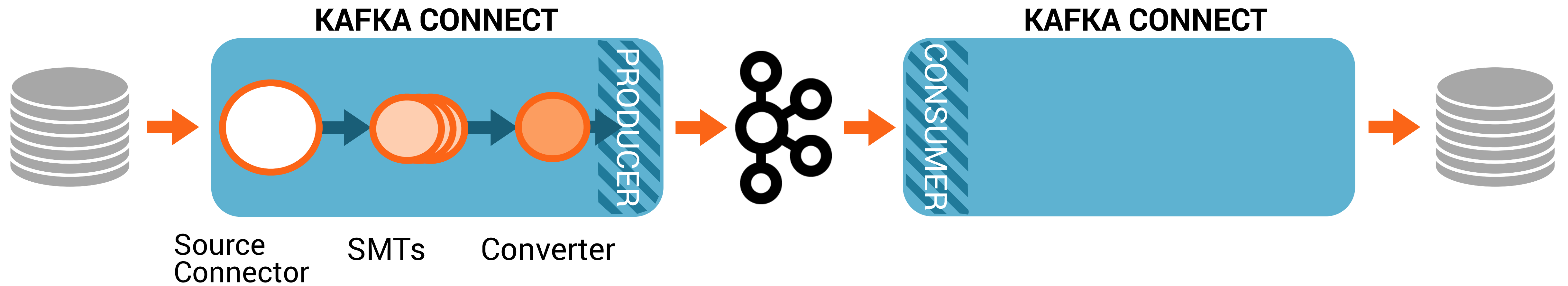
ENTER KAFKA CONNECT

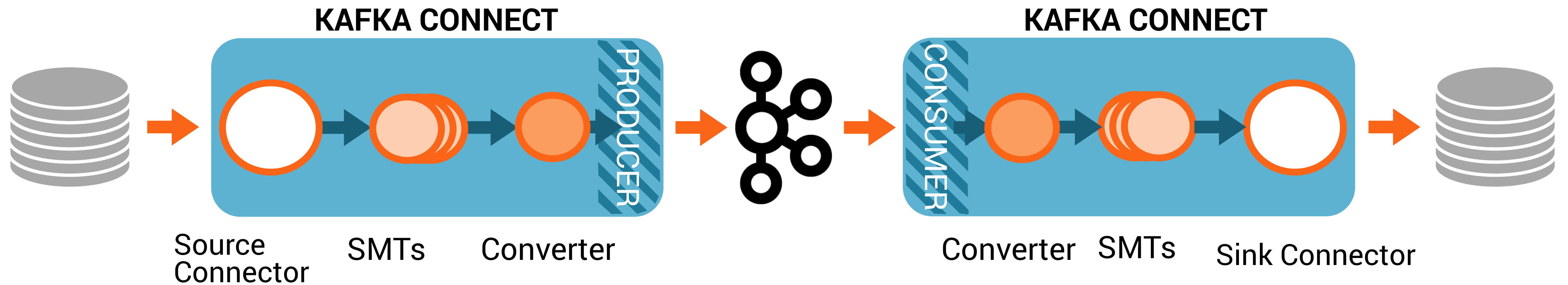




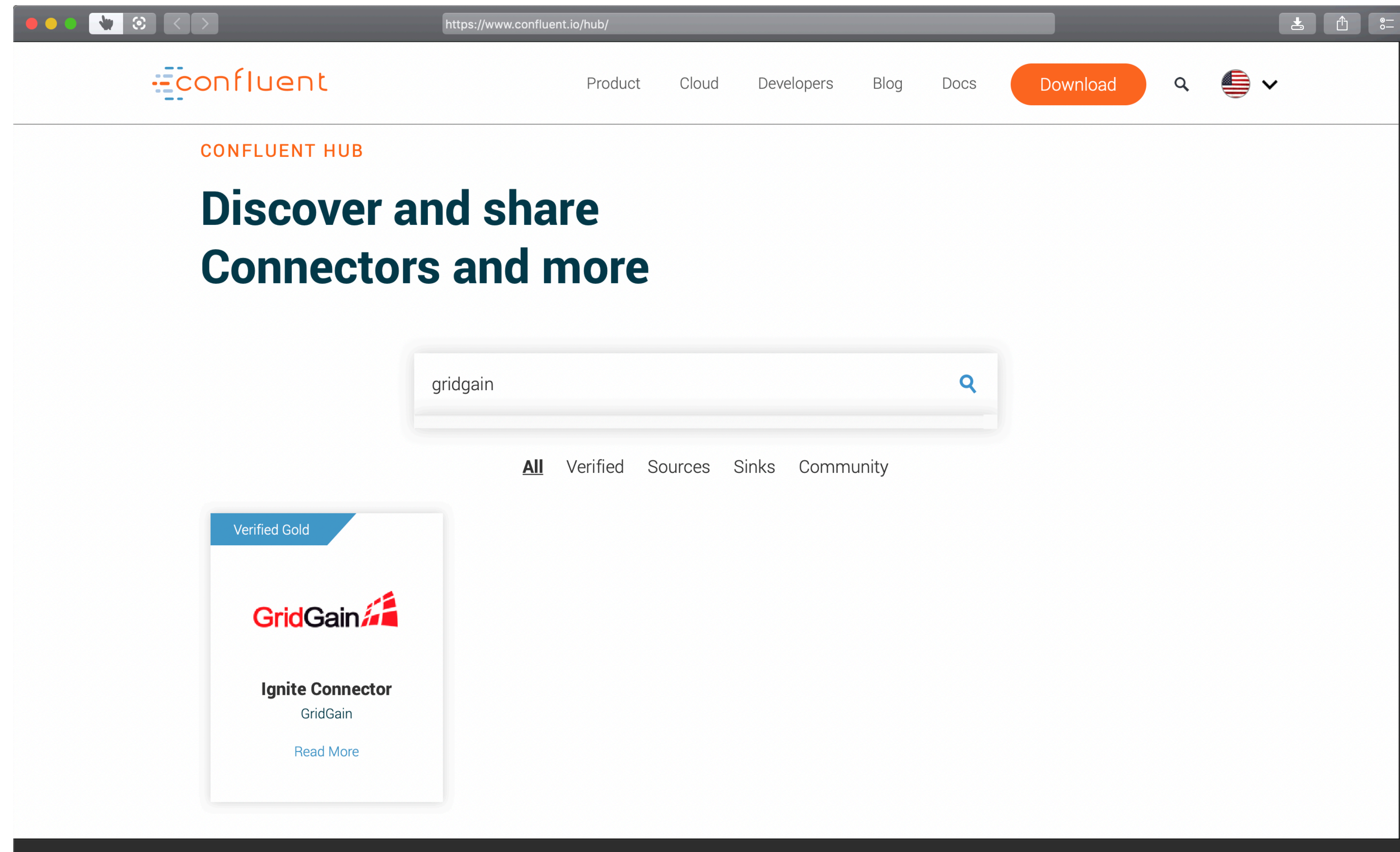




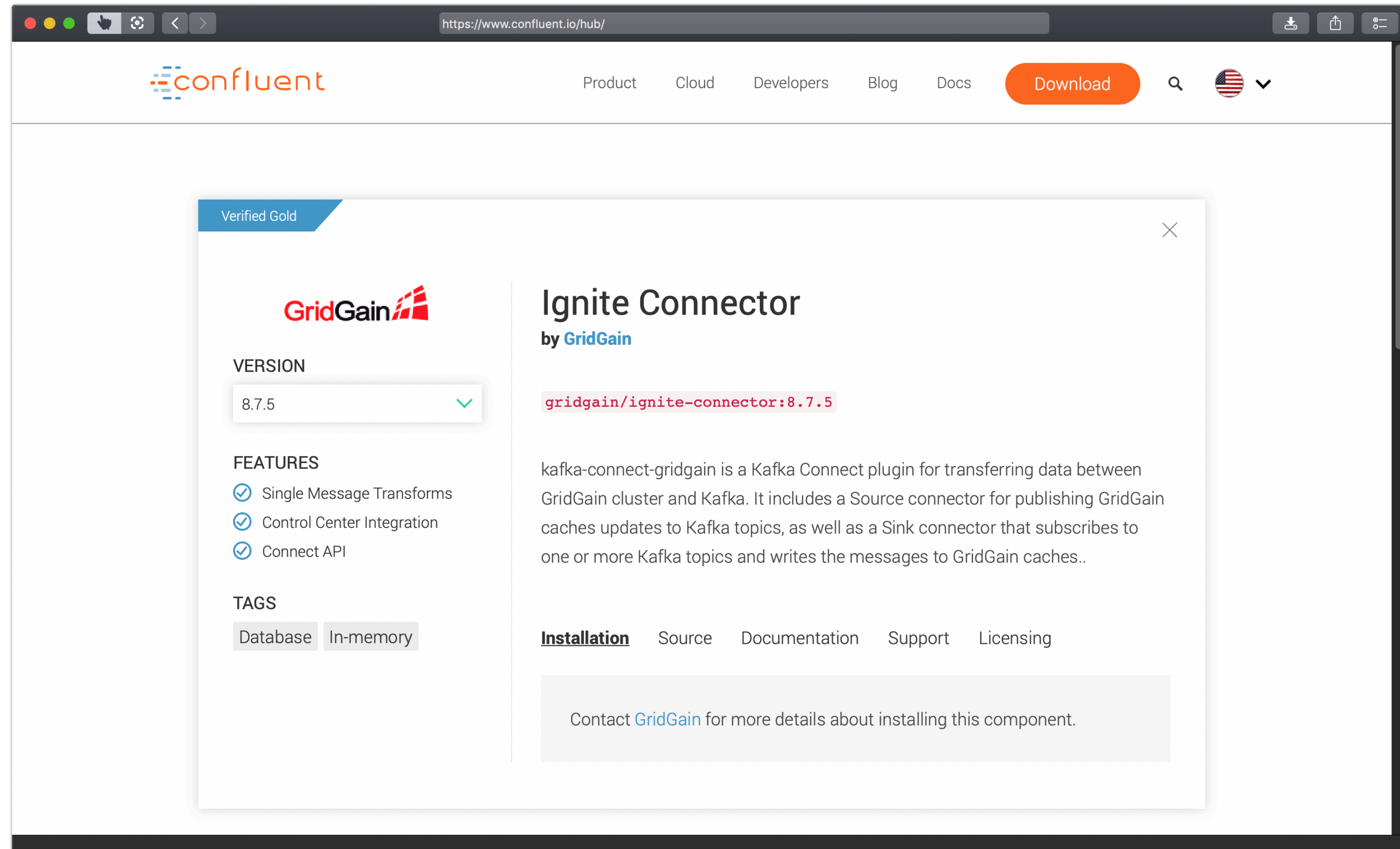


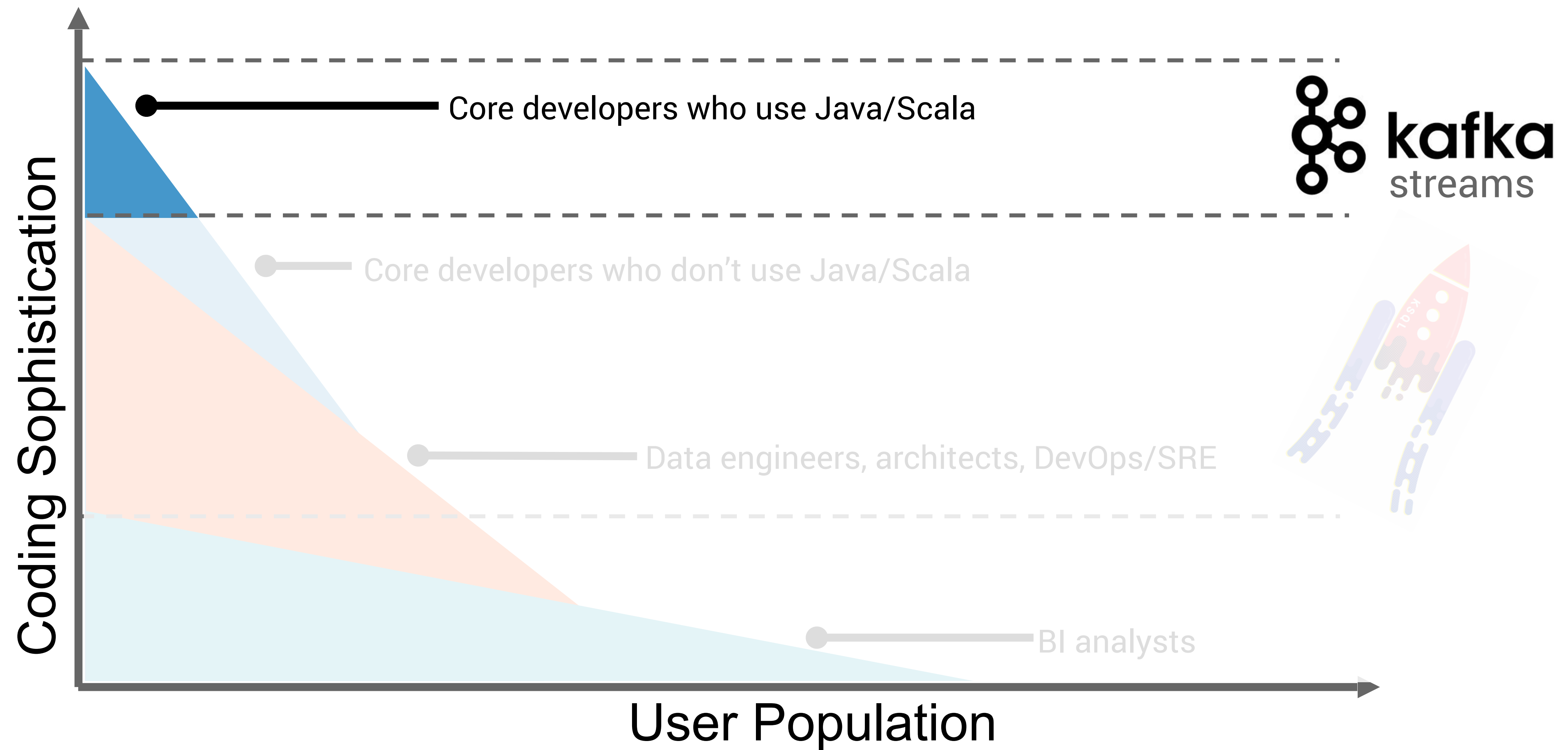


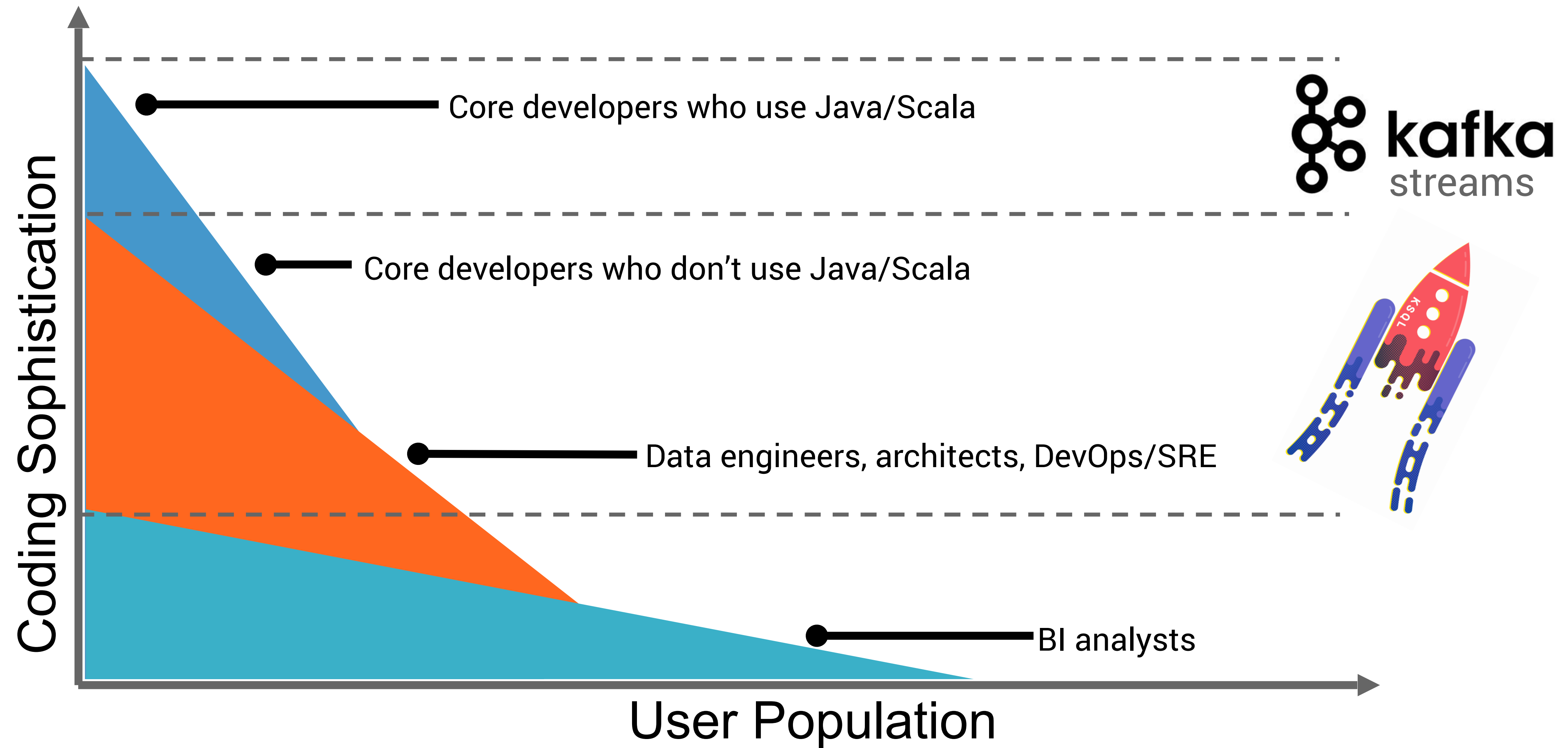
✓ Discover connectors,
SMTs, and converters



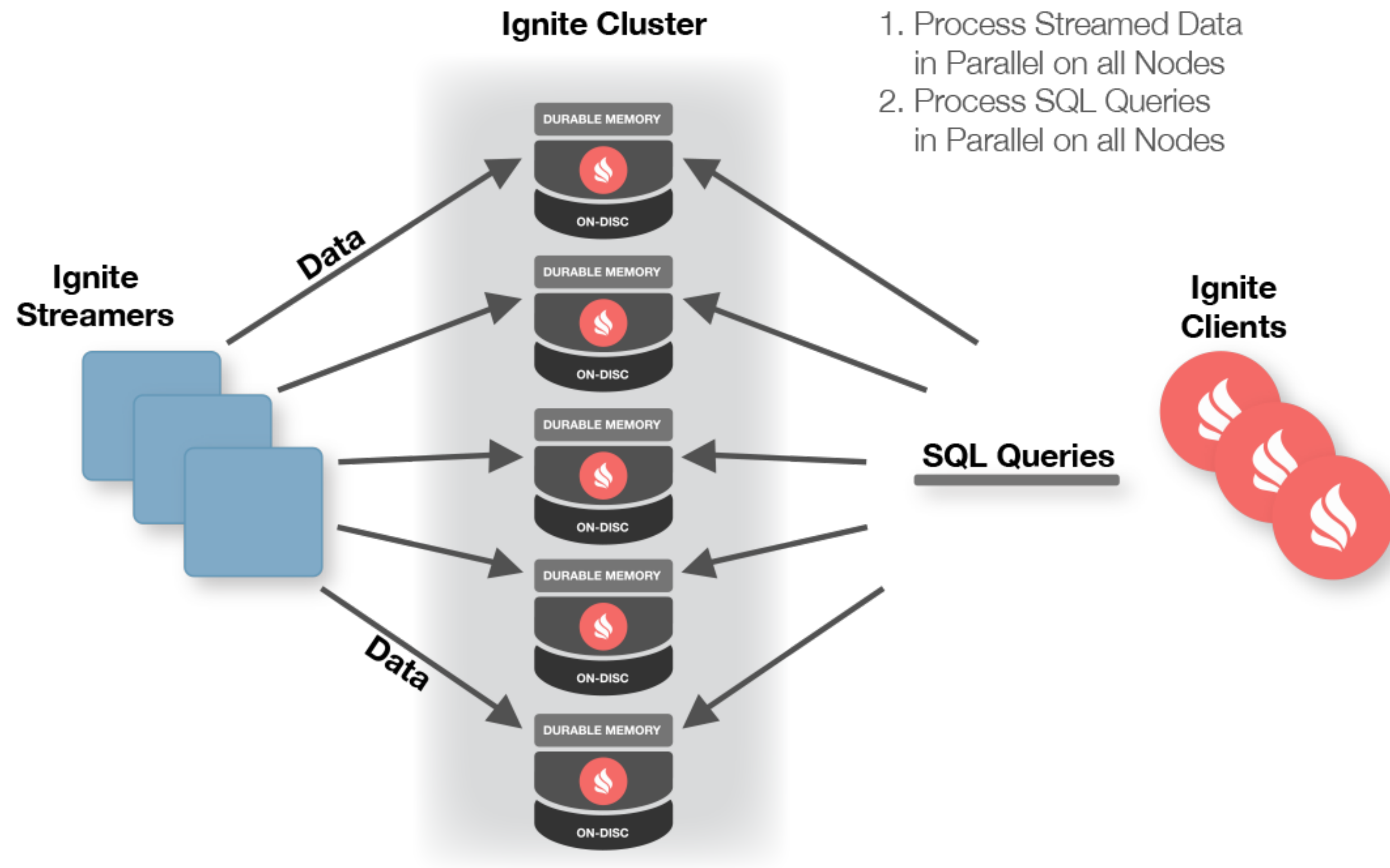
- ✓ Discover connectors, SMTs, and converters
- ✓ Descriptions, licensing, support, and more





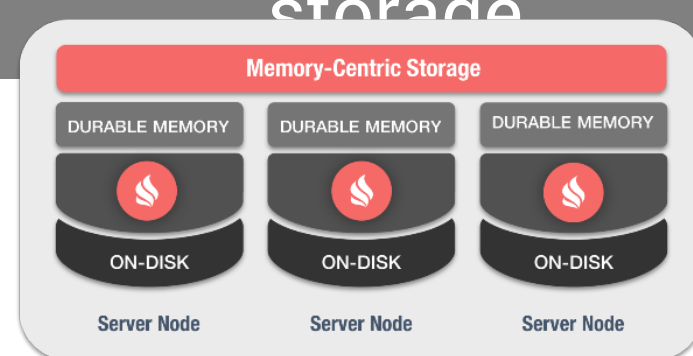


STORE AND PROCESS WITH GRIDGAIN



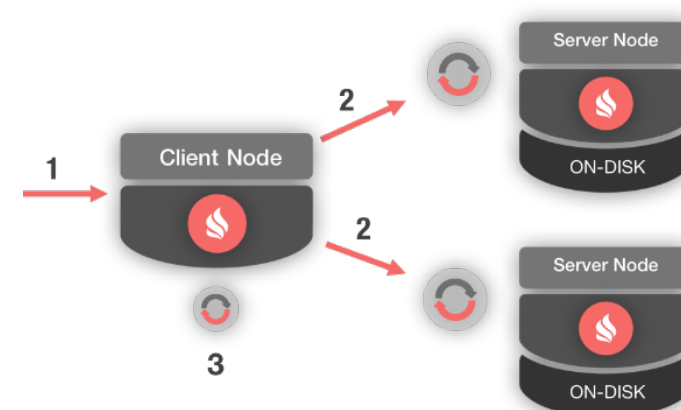
ESSENTIAL GRIDGAIN APIS

Distributed memory-centric storage



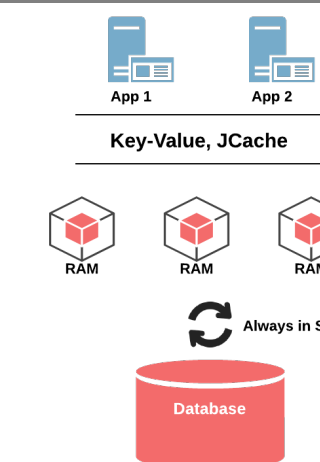
Combines the performance and scale of in-memory computing together with the disk durability and strong consistency in one system

Co-located Computations



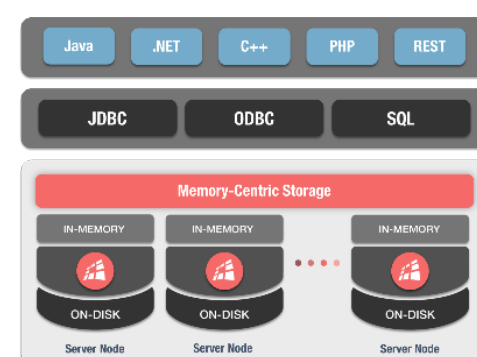
Brings the computations to the servers where the data actually resides, eliminating need to move data over the network

Distributed Key-Value



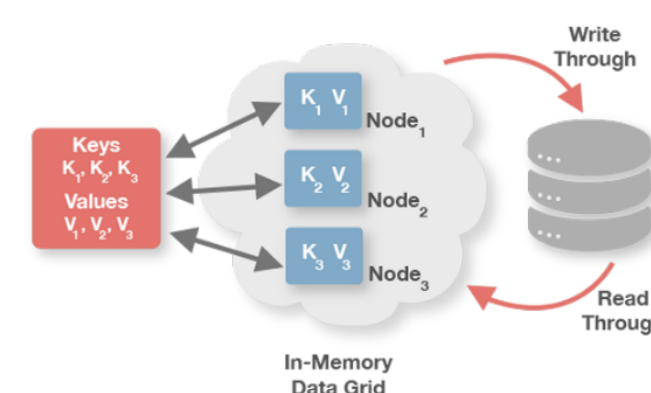
Read, write and transact with fast key-value APIs

Distributed SQL



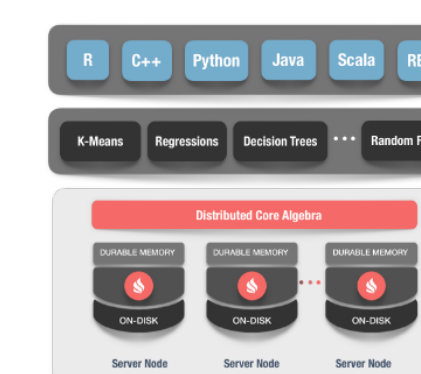
Horizontally, fault-tolerant distributed SQL database that treats memory and disk as active storage tiers

ACID Transactions



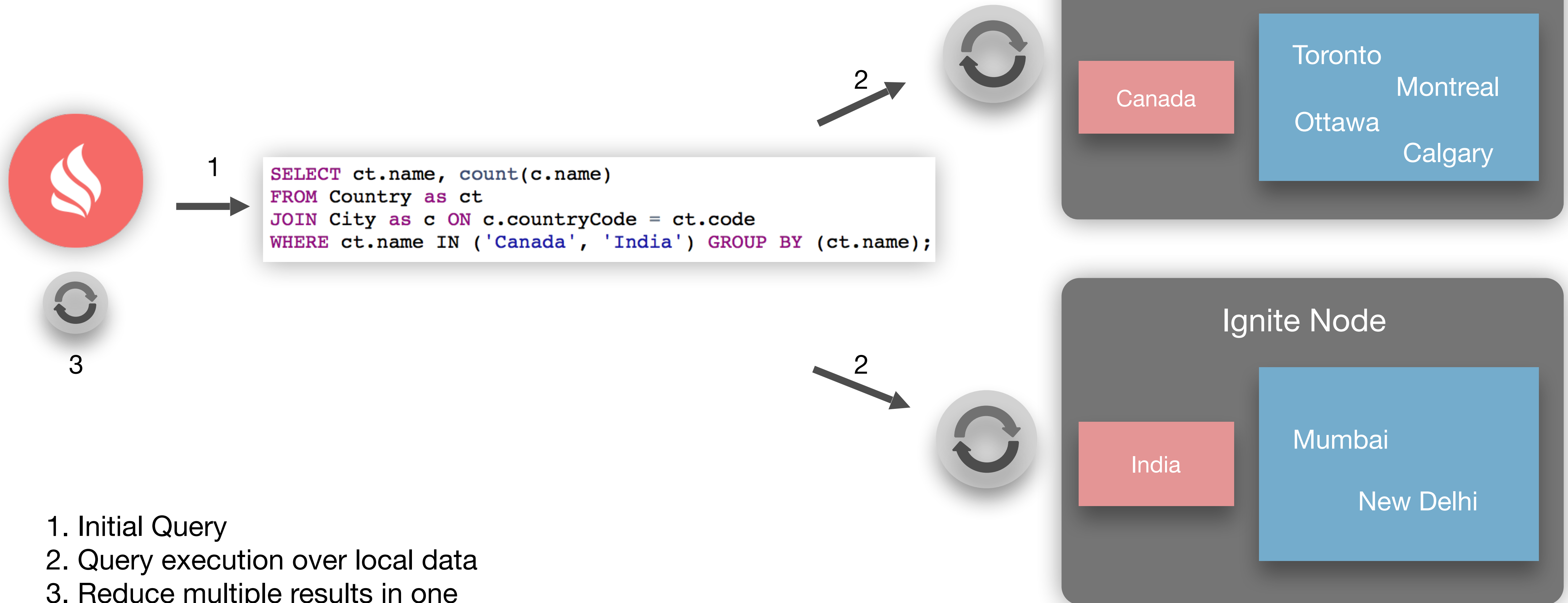
Supports distributed ACID transactions for key-value as well as SQL operations

Machine and Deep Learning



Set of simple, scalable and efficient tools that allow building predictive machine learning models without costly data transfers (ETL)

GRIDGAIN SQL FOR REAL-TIME ANALYTICS



1. Initial Query
2. Query execution over local data
3. Reduce multiple results in one

THANKS!

@denismagda
dmagda@gridgain.com

@gamussa
viktor@confluent.io

Q&A