



MONGODB.LOCAL
LONDON

Hans-Peter Grahsl, NETCONOMY

Streaming Data on the Shoulders of Giants

@hpgrahsl

#MDBlocal



> whoami

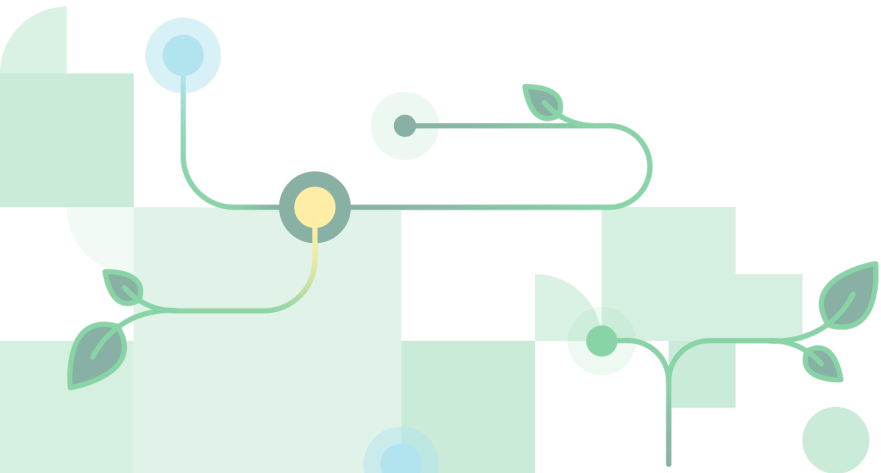
- ✓ working & living in Graz 🇦🇹
- ✓ technical trainer at NETCONOMY
- ✓ independent engineer / consultant
- ✓ associate lecturer
- ✓ 🗣️ occasional conference speaker 📢



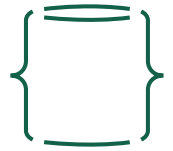
Speed & Agility

Among Top Tech Risks

For businesses to stay relevant they must **deliver value at a breakneck pace** and be constantly seeking new sources of value...



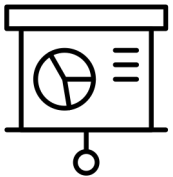
Managing, Processing & Analyzing Data



We all use Data



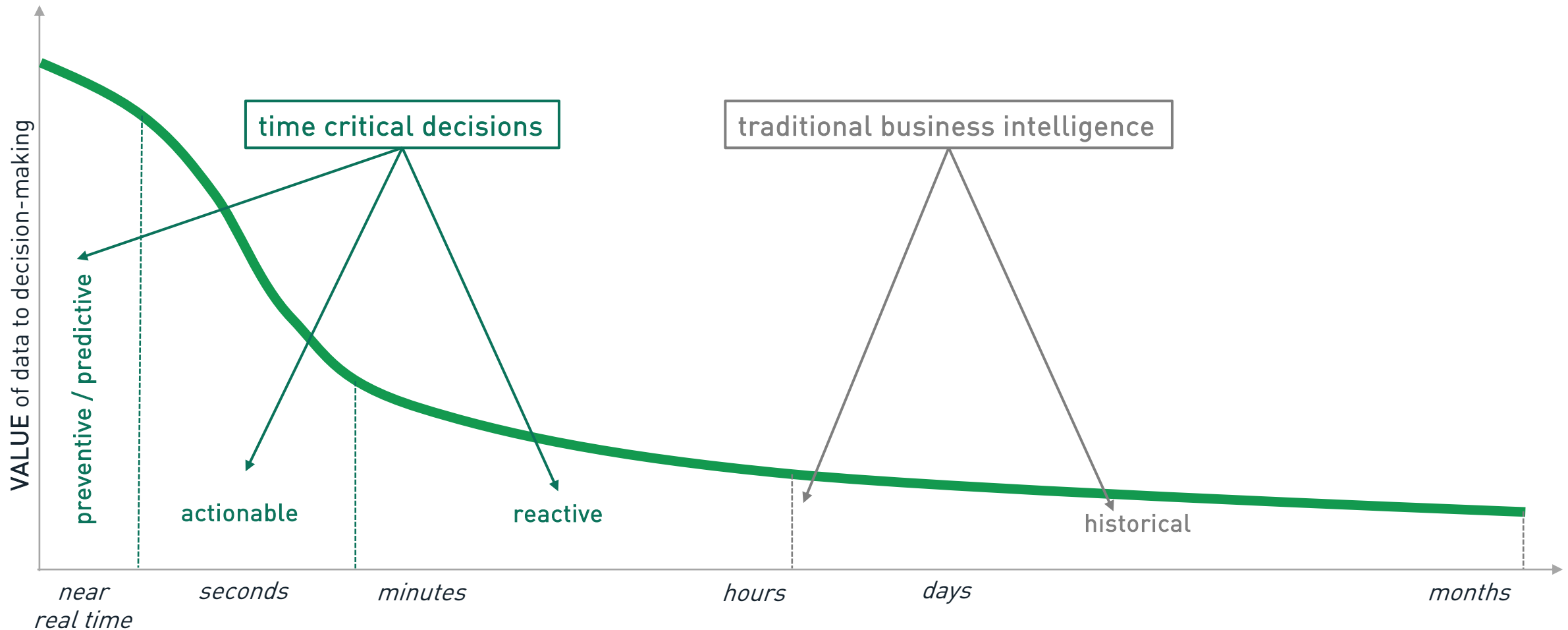
to unlock insights



and drive value



Diminishing Value of Data



Source: Perishable Insights, Mike Gualtieri, Forrester

Historic ETL can be rather painful

- batch-driven
- brittle & error prone
- (too) late answers

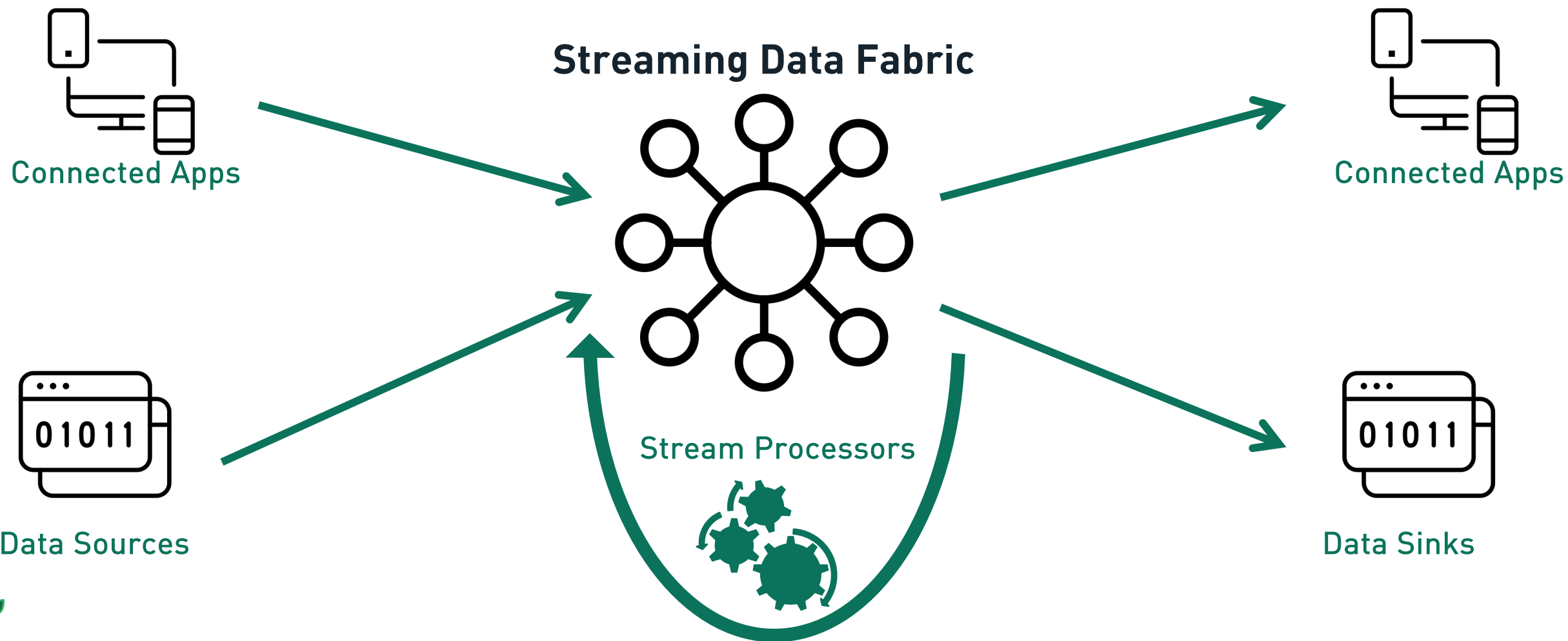
Speed & Agility Antipattern

Alleviate some ETL pain with streaming

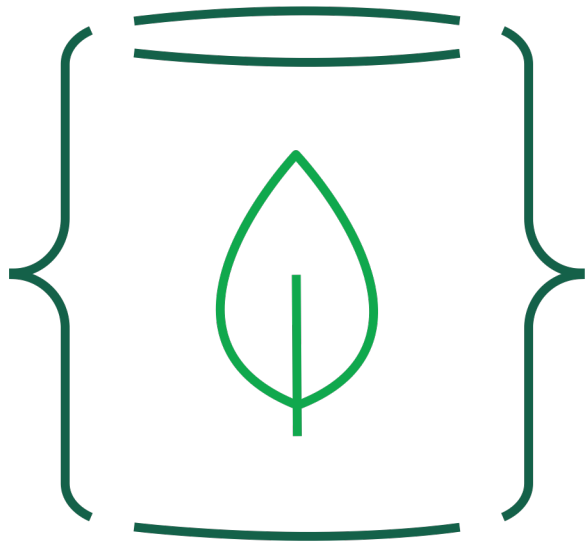
- event-centric & stream-oriented
- quick insights & fast answers

Speed & Agility
Enabler

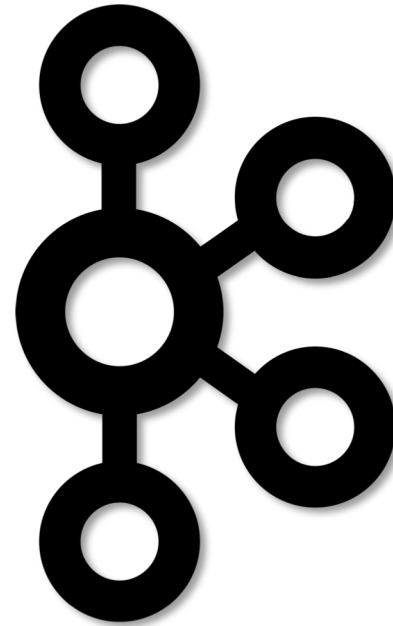
Architecture of a Modern Data Platform?



On the Shoulders of Giants

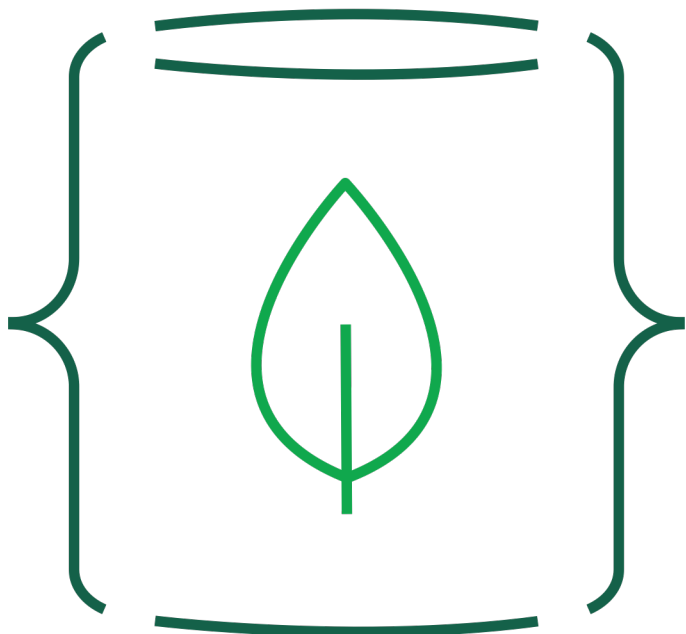


MongoDB®



Kafka®

Modern Database



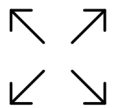
Modern Database



Document Model



Run Anywhere



Distributed & Scalable

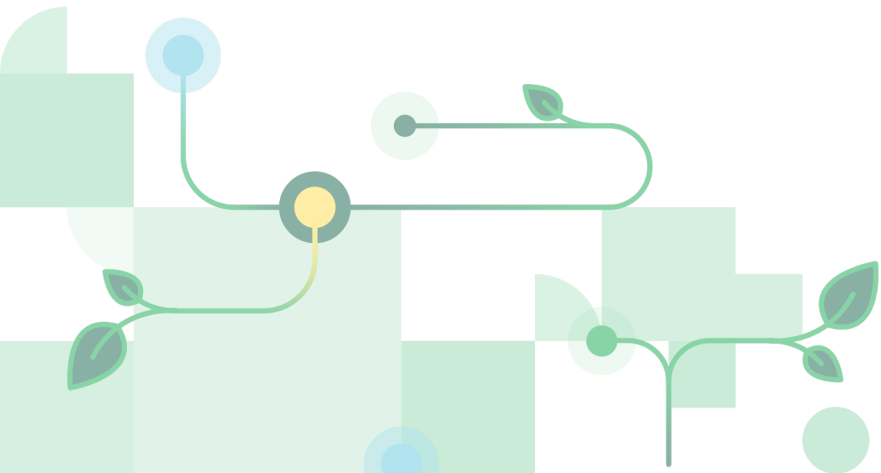


Resilient & Performant

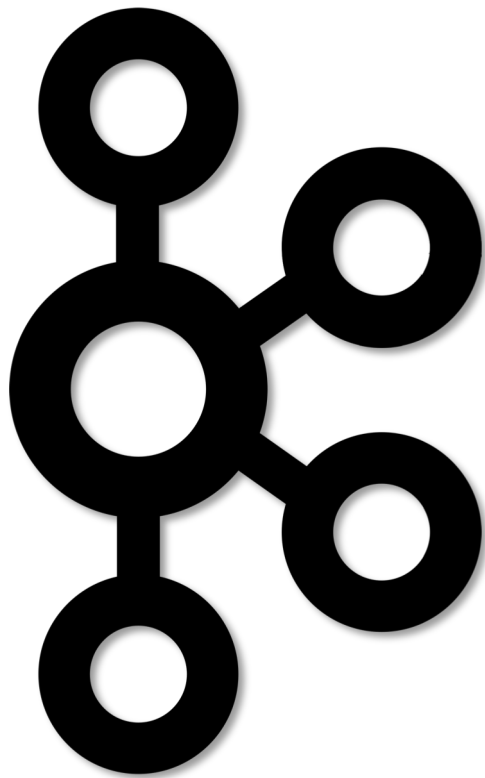


Apache Kafka®

Minimum Viable Introduction



Streaming Platform



Streaming Platform

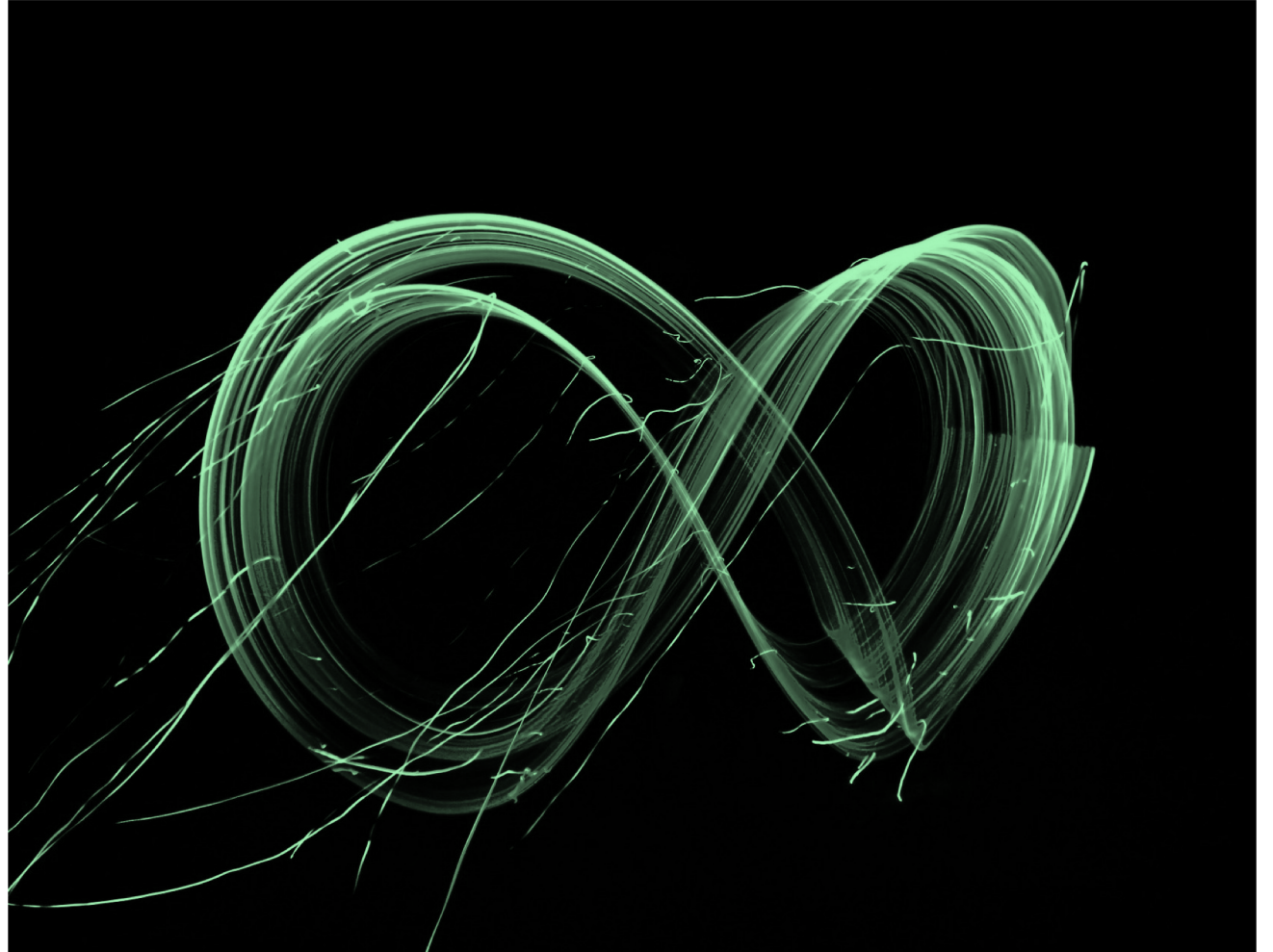
- ✓ distributed
- ✓ horizontally scalable
- ✓ highly fault-tolerant



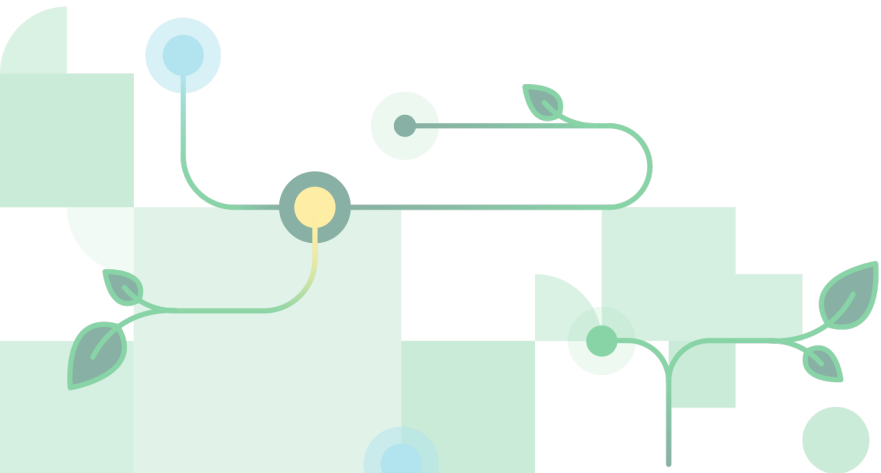
What is Streaming?

“... a type of **data processing** that is designed with **infinite data sets** in mind ...”

– *Tyler Akidau*

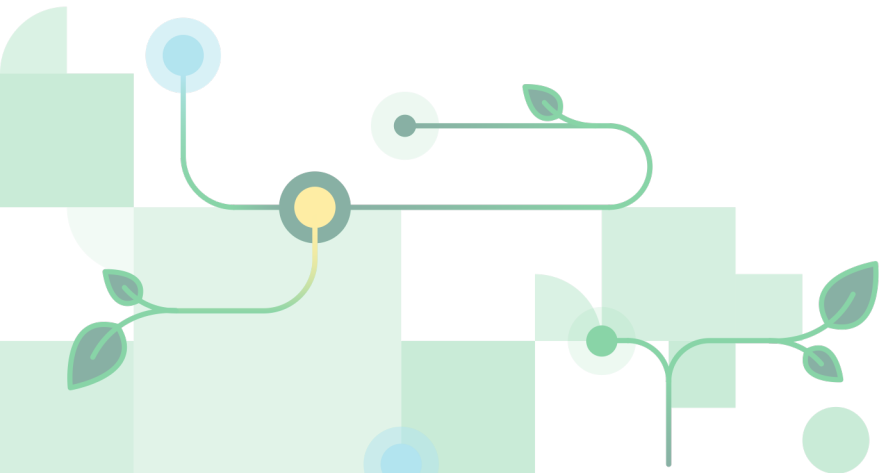


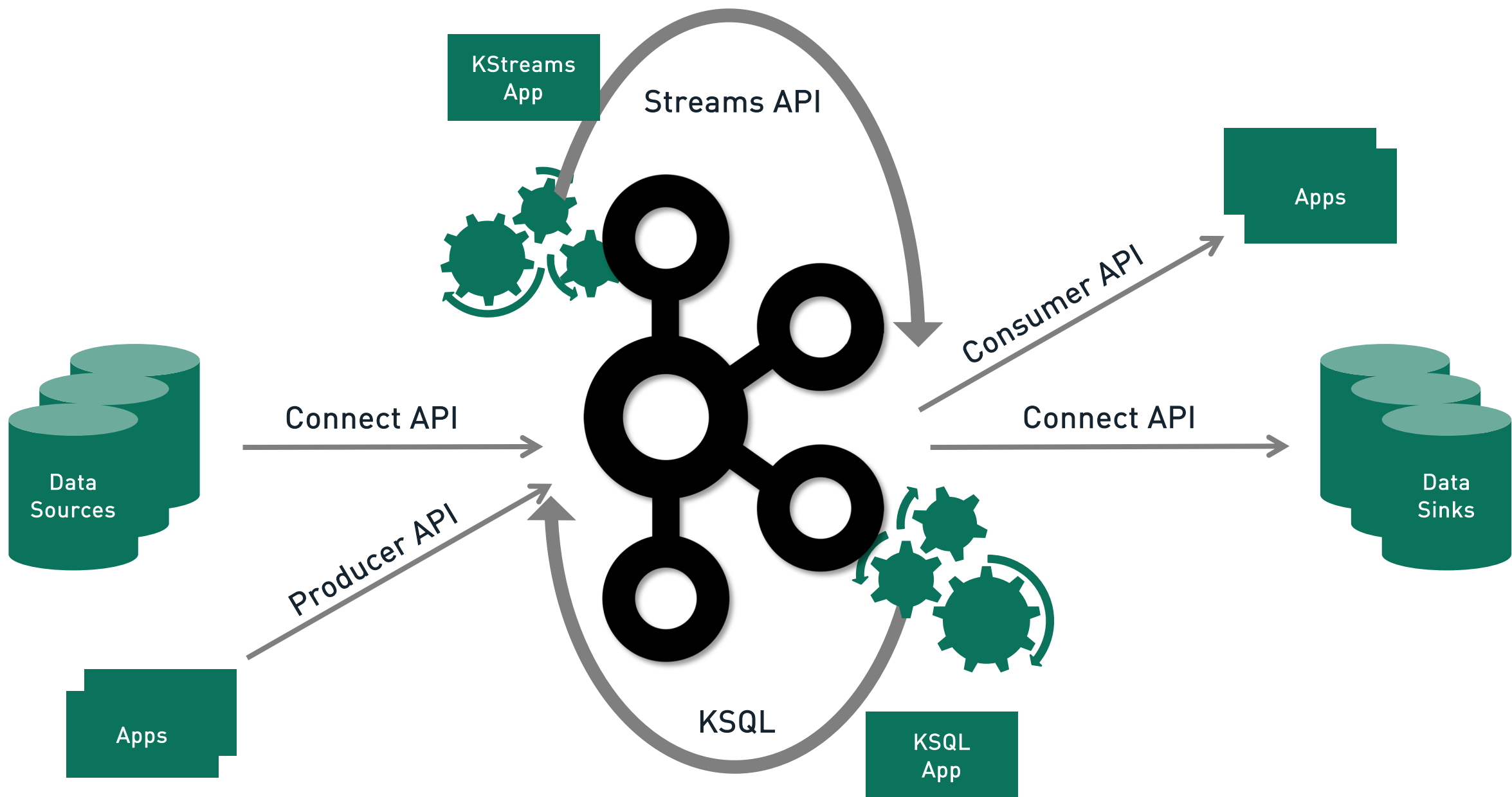
"...everything that happens in a company
– every customer interaction, every API request, every
database change – can be represented as
real-time stream that anything else can tap into,
process or react to."



"...Kafka and the whole category of **stream processing** represents a **fundamental paradigm shift** in how the digital part of a company is built, how data is used, and **how applications are built**. This is actually a pretty rare thing..."

– *Jay Kreps*



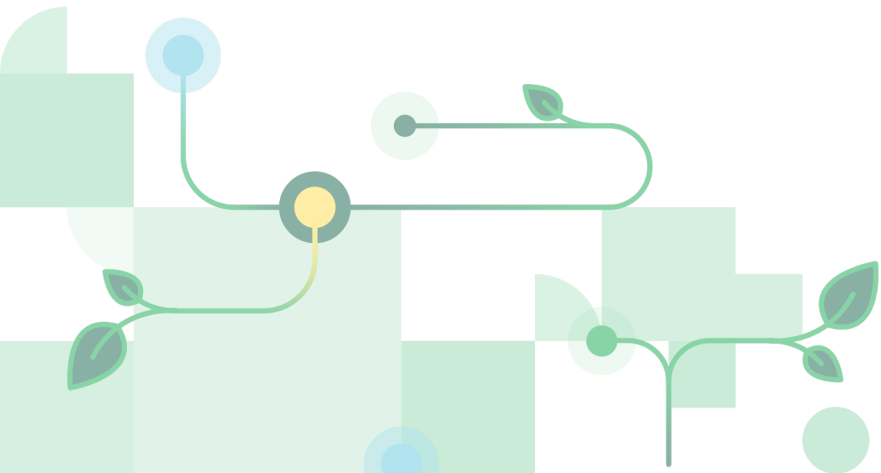


Kafka APIs in a Nutshell...

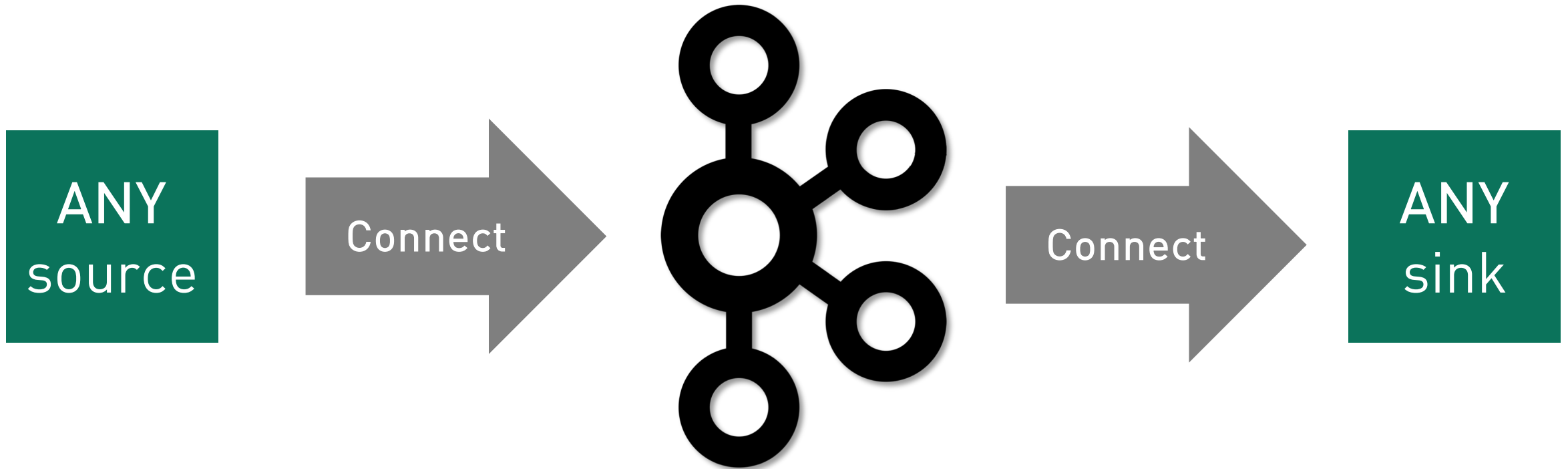
- Producer & Consumer API
 - publish-subscribe scenarios
- Connect API
 - streaming data integration scenarios
- Streams API & KSQL
 - code or SQL-based streaming scenarios

Kafka® Connect

What's it about?



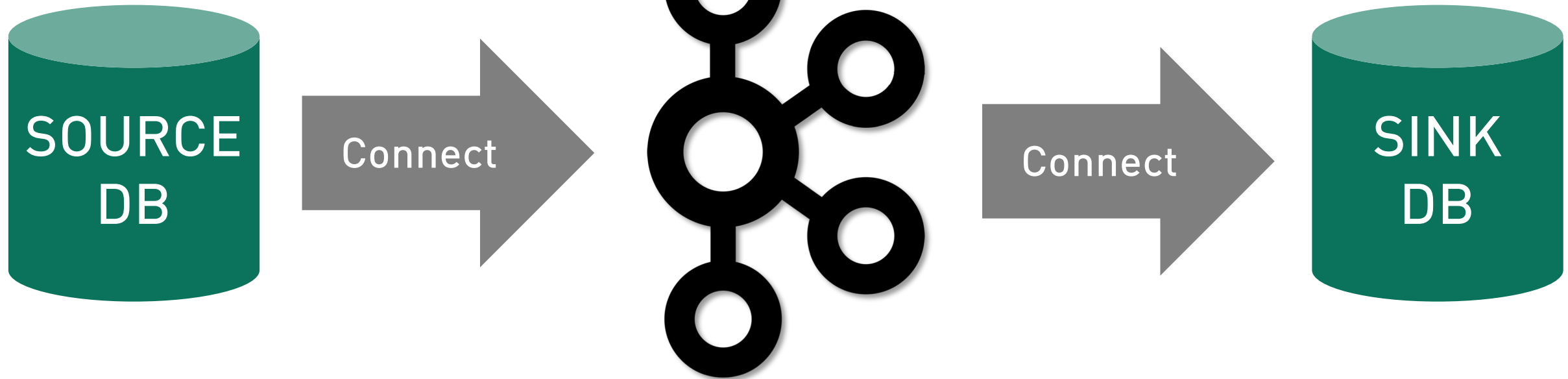
Kafka Connect Basics



ANY → e.g. file systems, data stores, REST endpoints, ...

Kafka Connect Basics

often about **data stores**

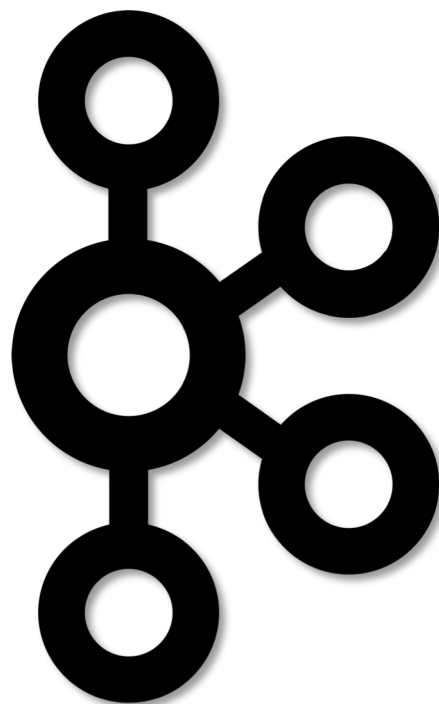


Kafka Connect Basics

or more concretely



Connect



Connect



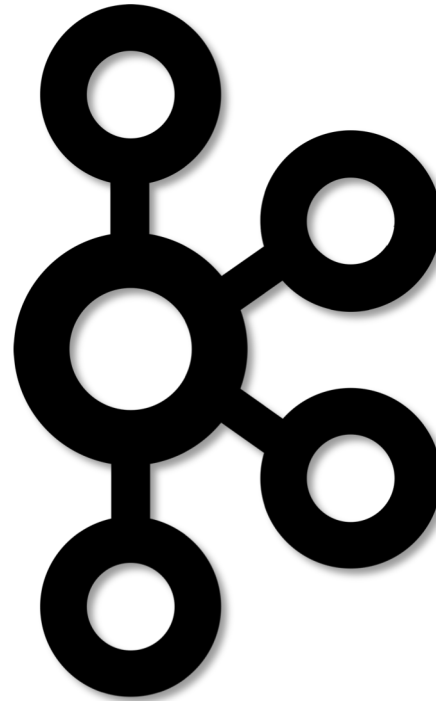
<https://hub.confluent.io> → many many more

Kafka Connect Basics

or more concretely



Connect



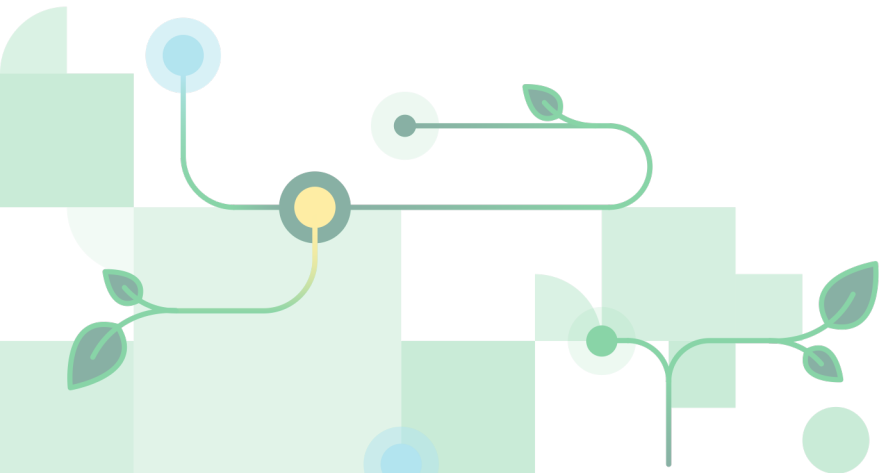
Connect



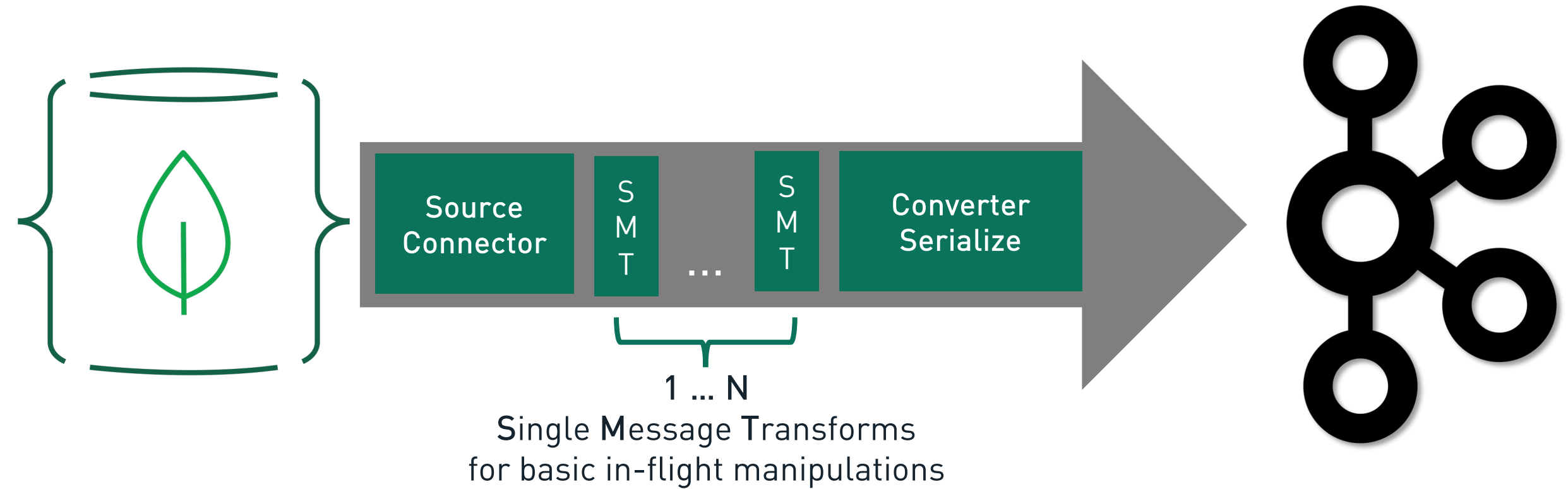
<https://hub.confluent.io> → many many more

Kafka® Connect

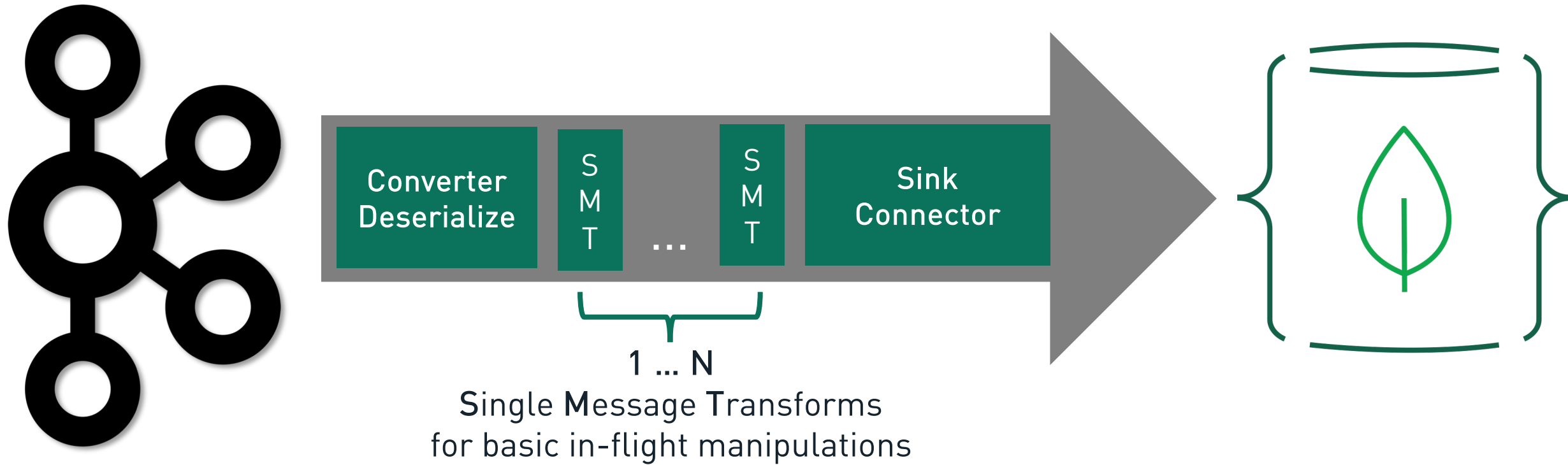
How do connectors operate?



Kafka Source Connectors

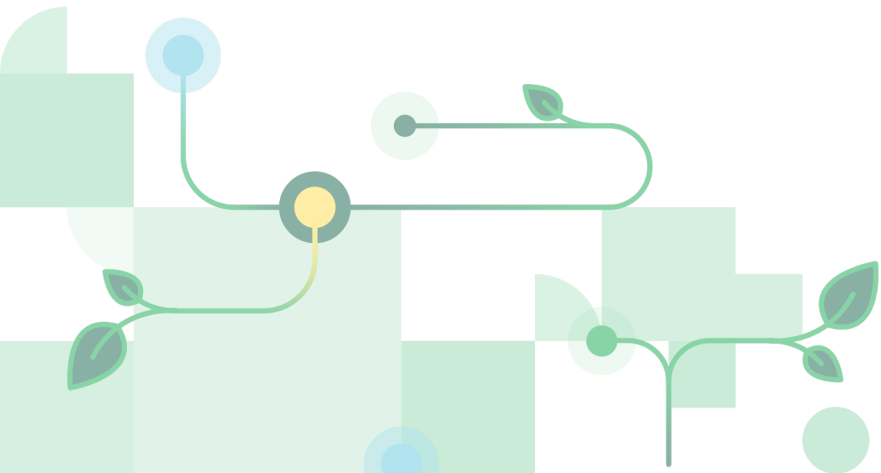


Kafka Sink Connectors

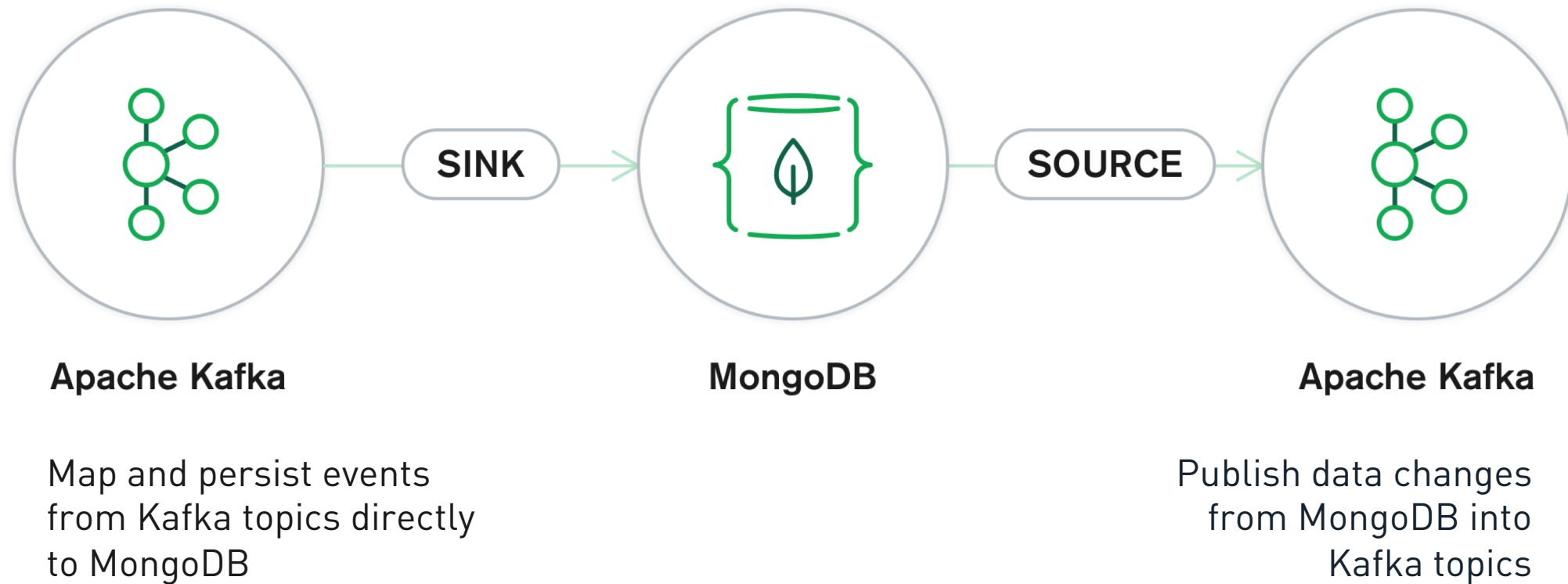


Official Connector

Announced at #MDBW19

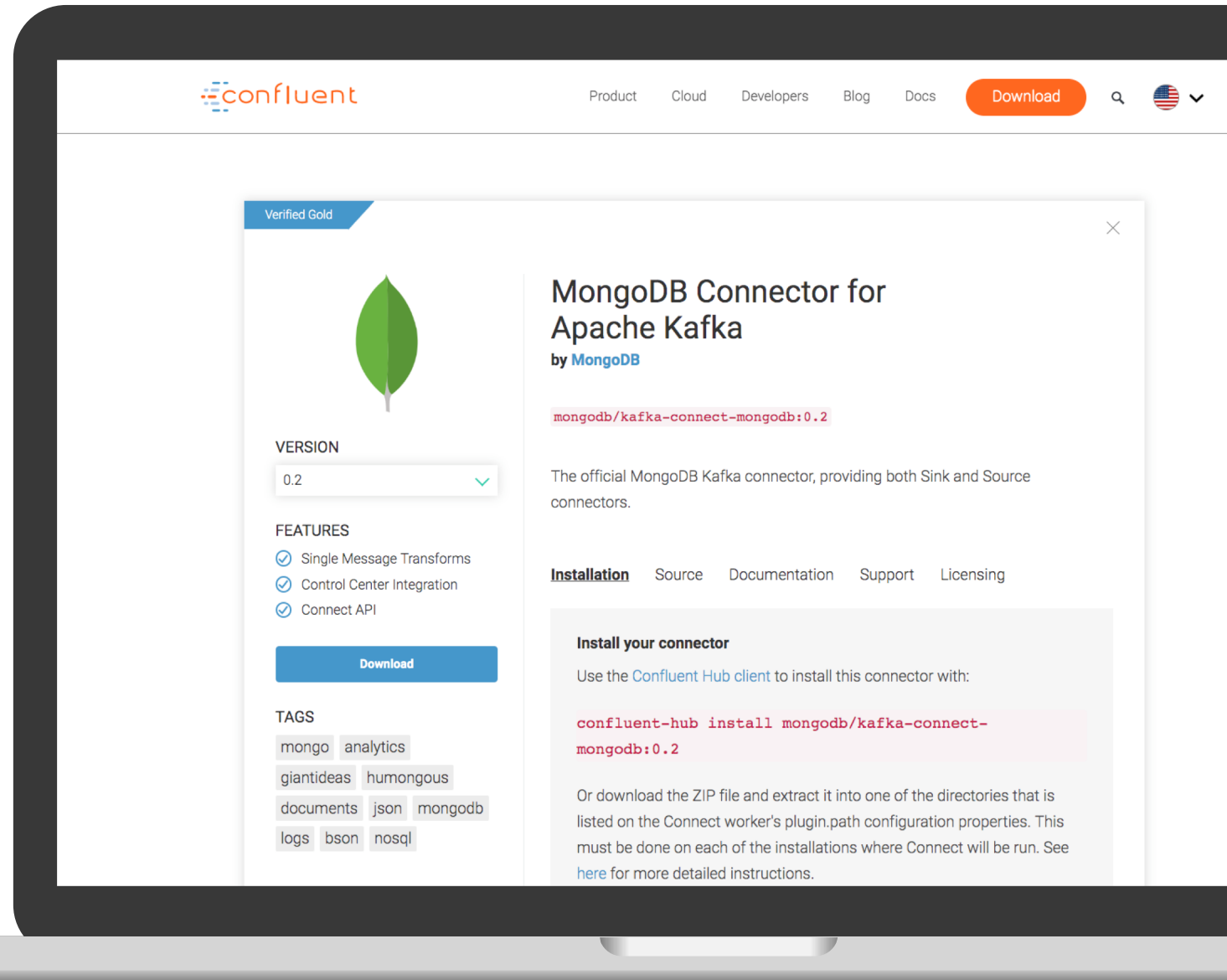


MongoDB Connector for Apache Kafka



MongoDB Connector for Apache Kafka

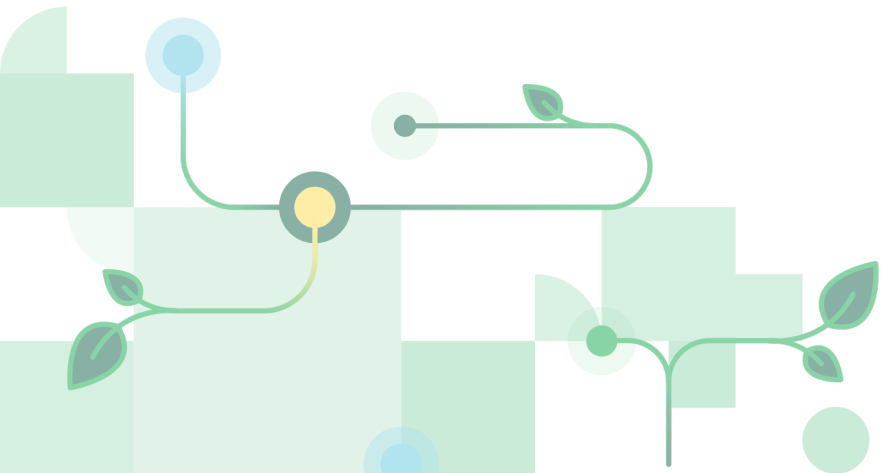
- ✓ developed open-source
- ✓ officially supported by MongoDB
- ✓ Verified Gold certified by Confluent



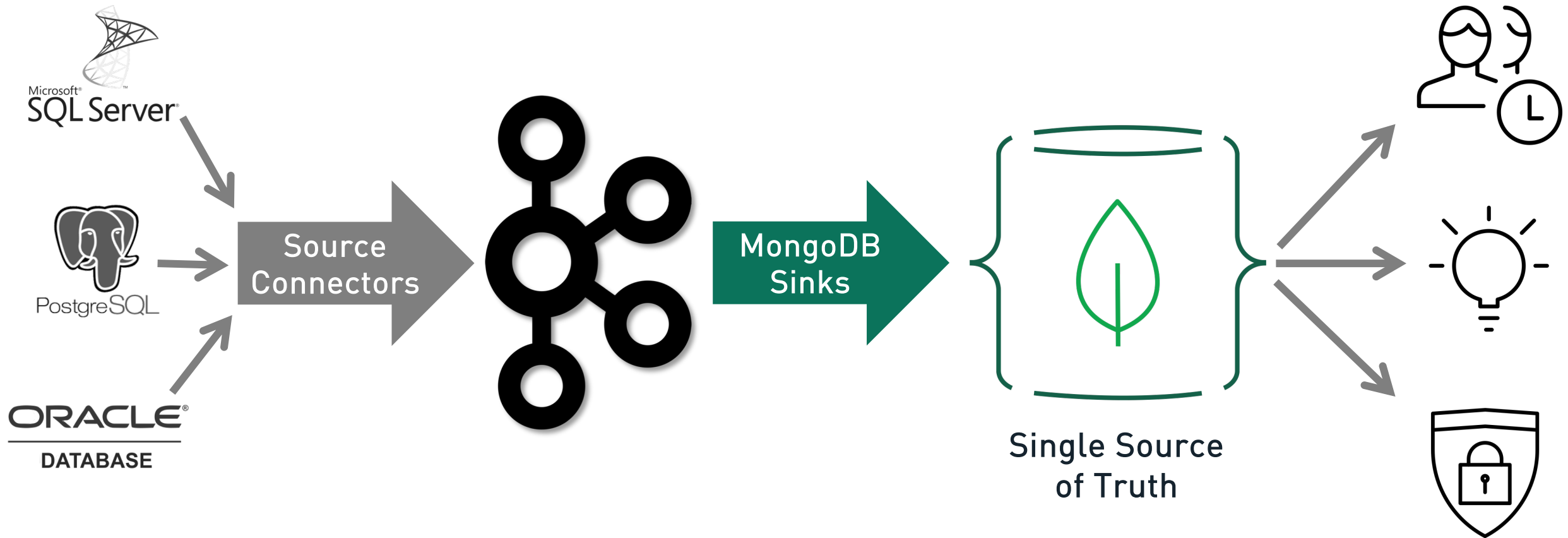
Available for Download on the Confluent Hub
<https://www.confluent.io/hub/mongodb/kafka-connect-mongodb>

Use Cases

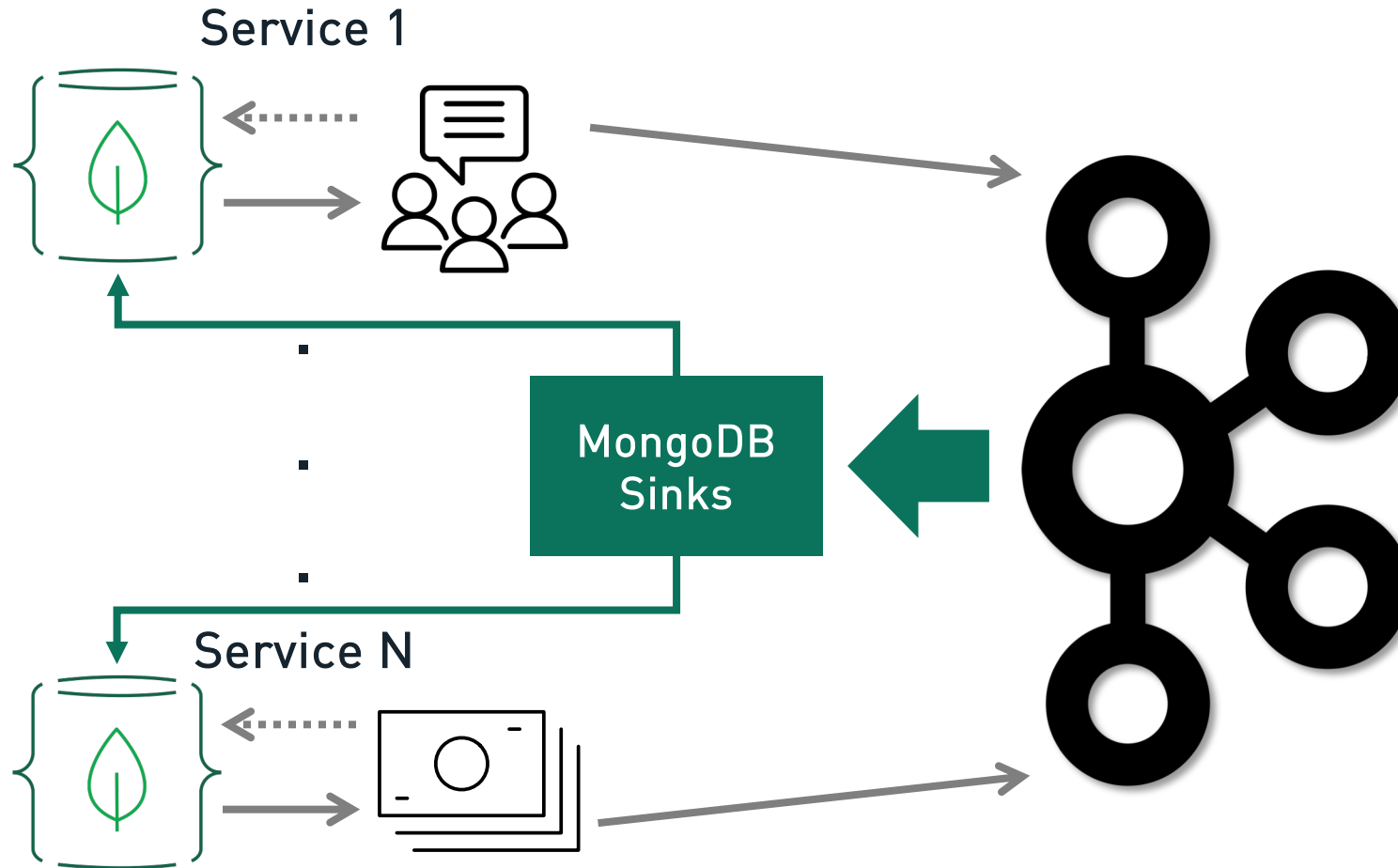
MongoDB Connector for Apache Kafka



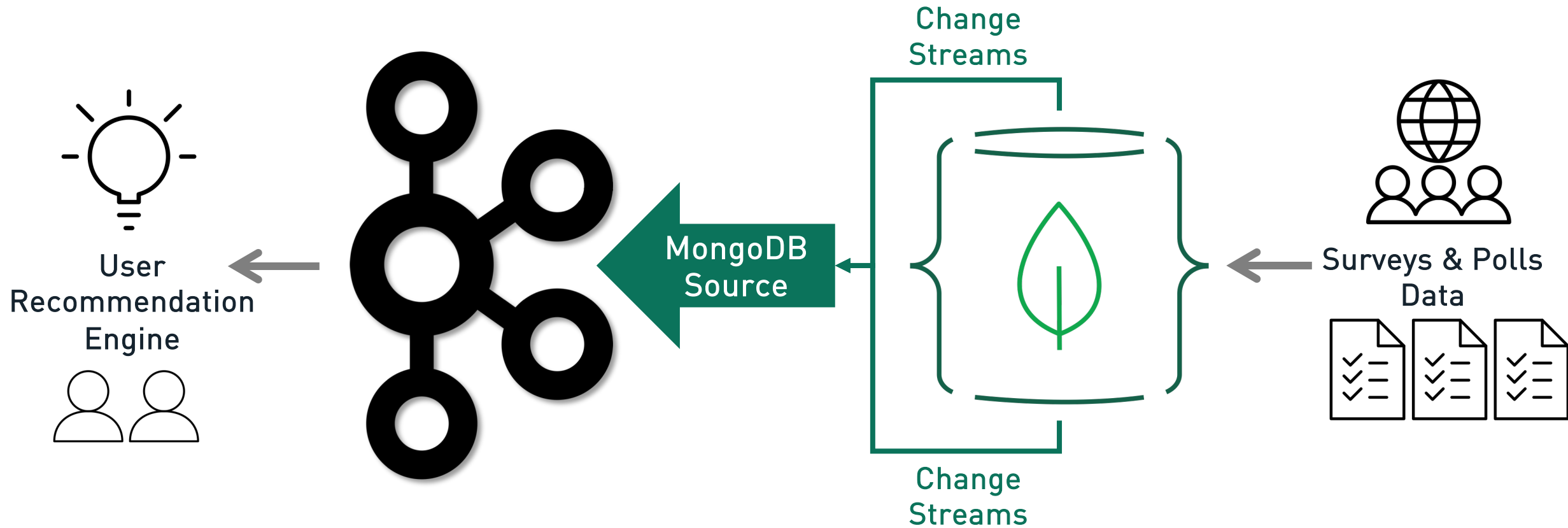
Single Customer View for eCommerce



Data Synchronization between Microservices

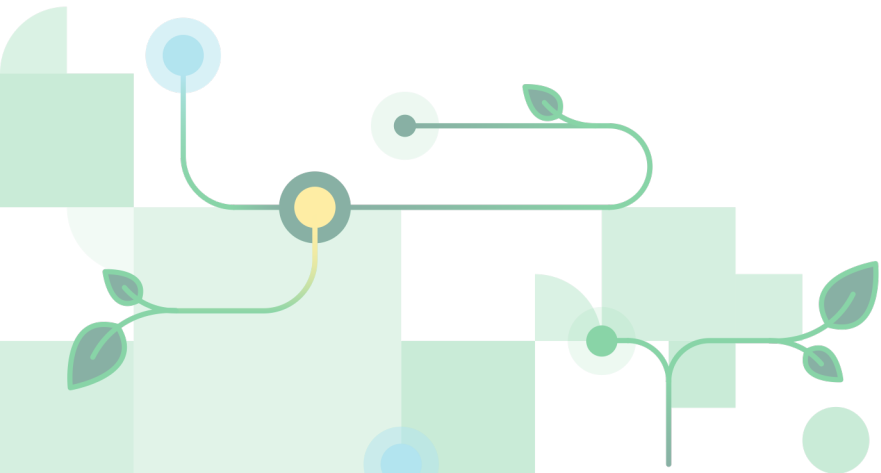


Recommendation Engine for Opinion Mining

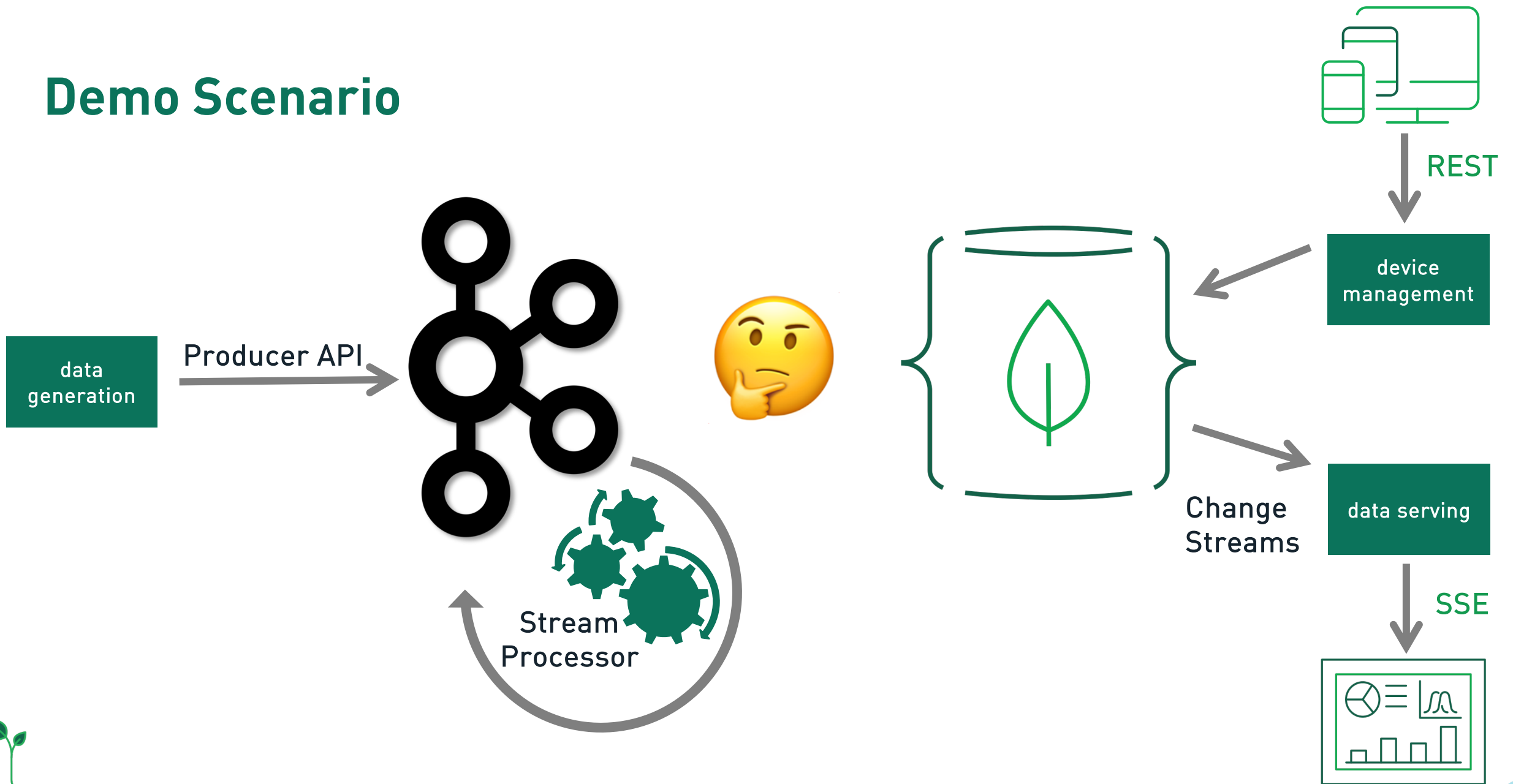


Demo Scenario

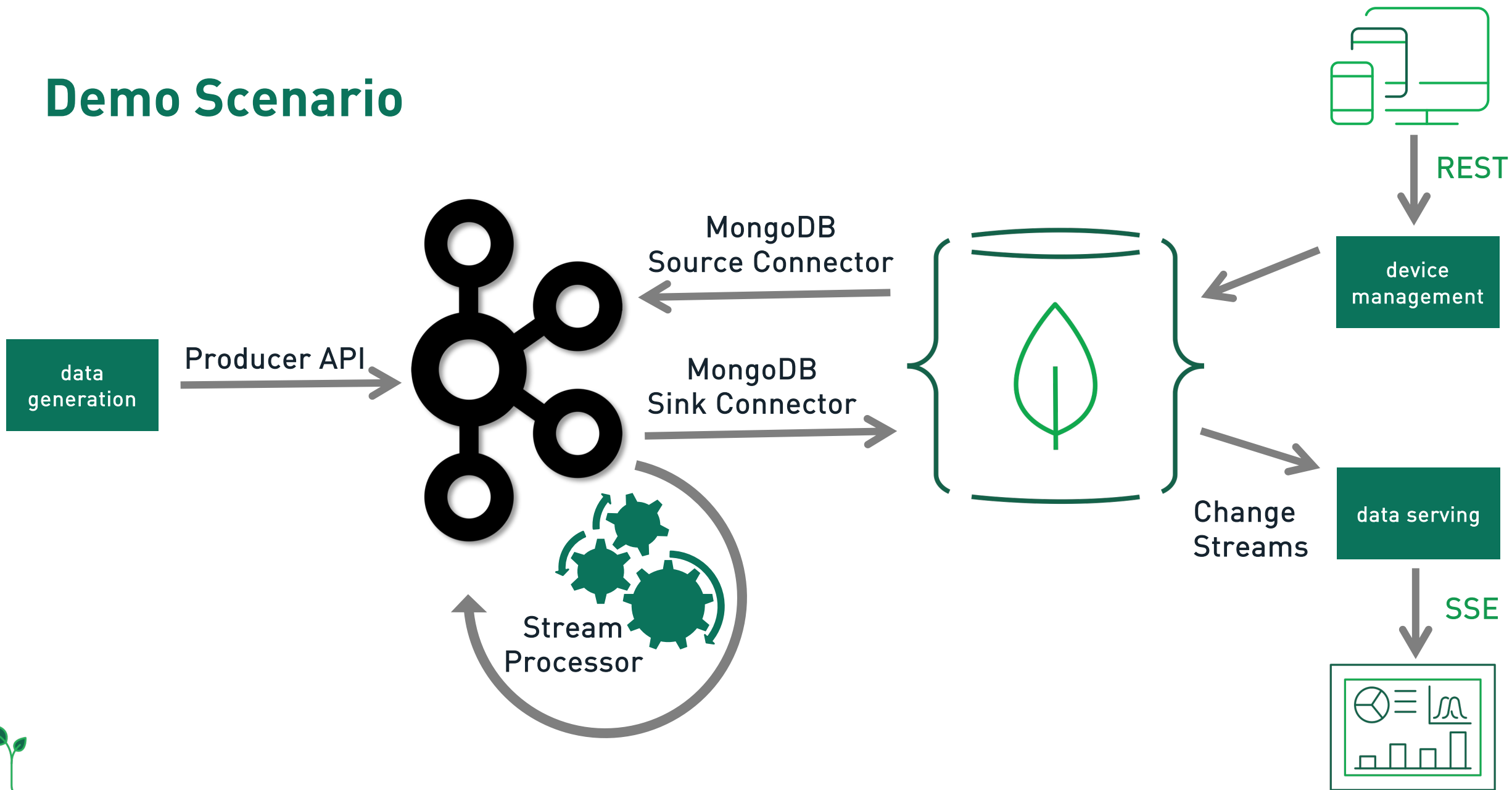
Let's see it in action!



Demo Scenario



Demo Scenario





MONGODB
.LOCAL

THANK YOU!





Streaming Data on the Shoulders of Giants [NETCONOMY]

Hans-Peter Grahsl

<https://www.surveymonkey.com/r/KXC6DB6>



MONGODB LOCAL

