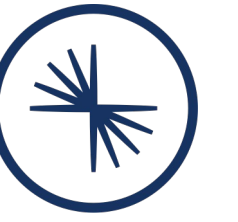


kafka

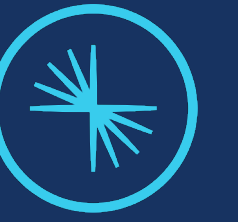
as Data Hub for Crypto, NFT, Metaverse (Beyond the Buzz!)



bitcoin ethereum
 CARDANO HYPERLEDGER
 SOLANA r3.c.rda
 IOTA Meta
 OpenSea

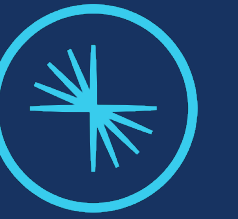


Agenda



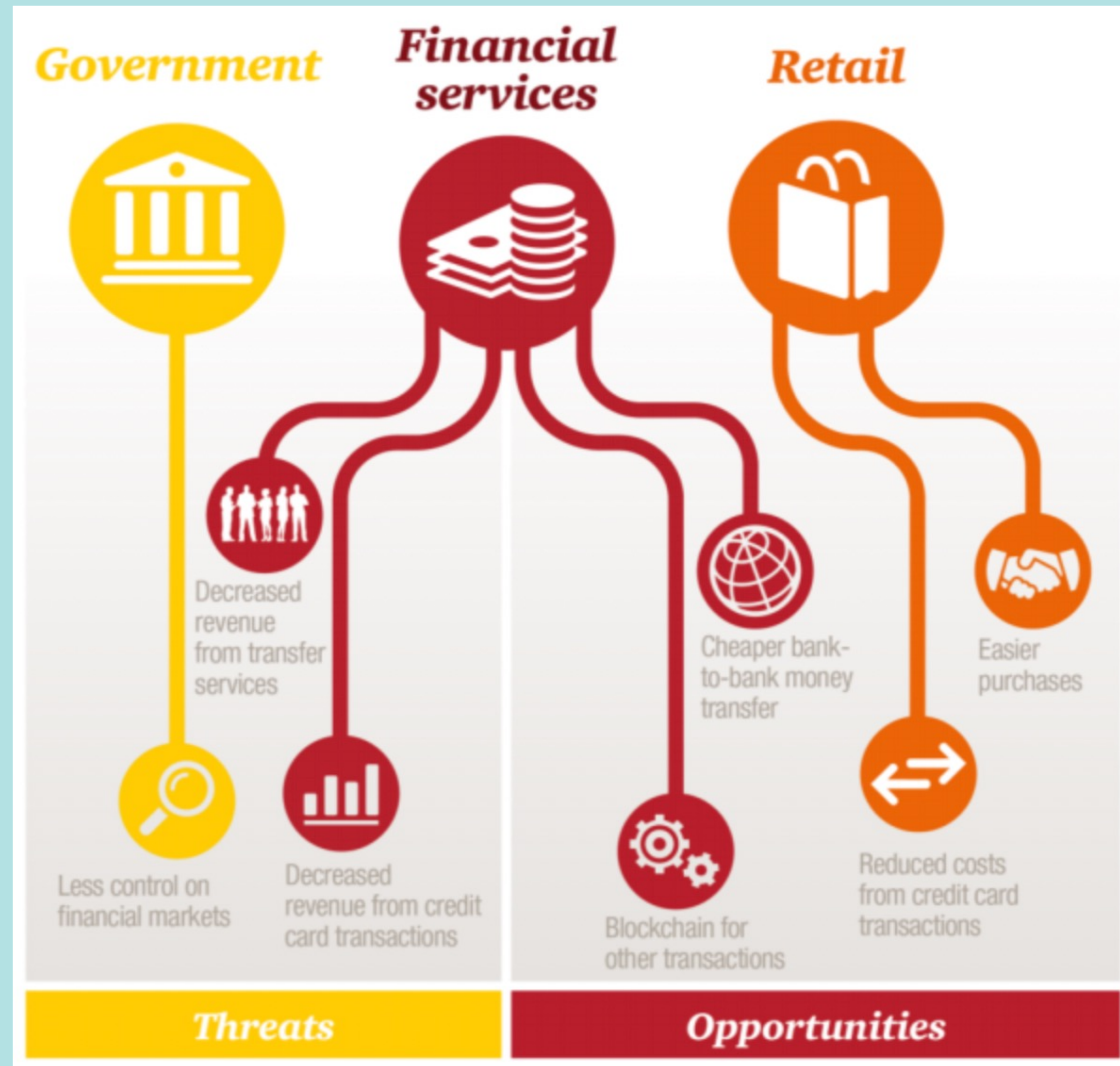
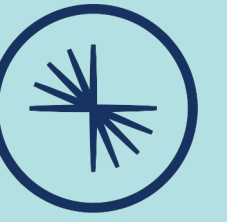
1. When (not) to use Blockchain and Crypto
2. Data Streaming for Real-time Analytics at Scale
3. Examples for Data Streaming with Crypto Data
4. Real-World Projects
5. Choosing the Right Tool for the Job

Agenda

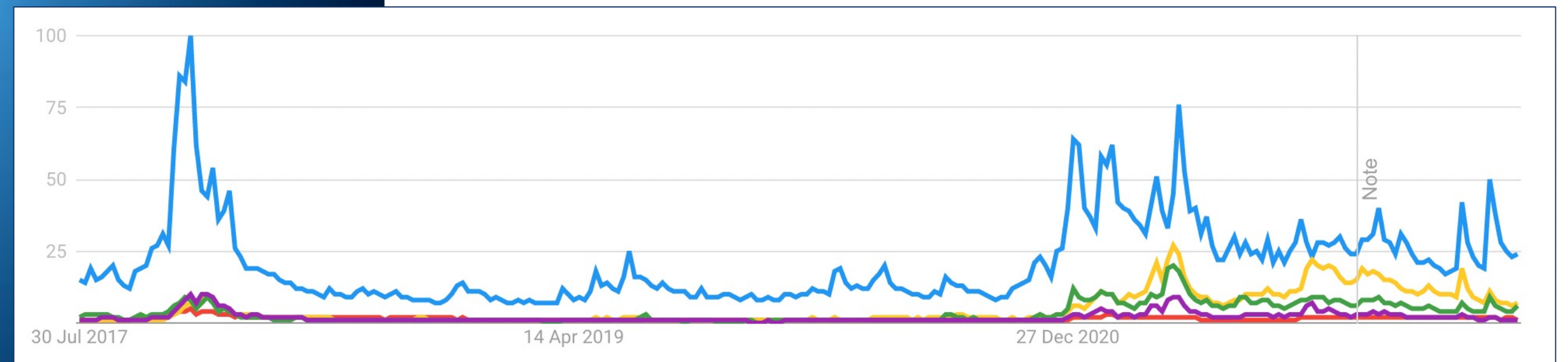
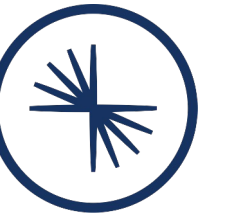


1. **When (not) to use Blockchain and Crypto**
2. Data Streaming for Real-time Analytics at Scale
3. Examples for Data Streaming with Crypto Data
4. Real-World Projects
5. Choosing the Right Tool for the Job

Potential impact of cryptocurrency on financial services

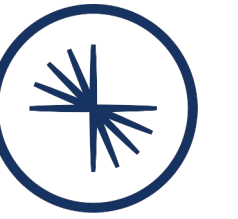


Bitcoin – The Tip of the Iceberg!



'**Bitcoin**' compared to
'Blockchain', 'Crypto', 'Ethereum'

Web3 beyond the hype: 5 types of digital assets



Native tokens, which are the monetary incentives used to compensate nodes for maintaining and updating the respective blockchain

Stablecoins, which are supposed to represent cash on the blockchain and are pegged to fiat currencies like the US dollar, or central bank digital currencies (CBDCs), which are regulated by a central bank ²

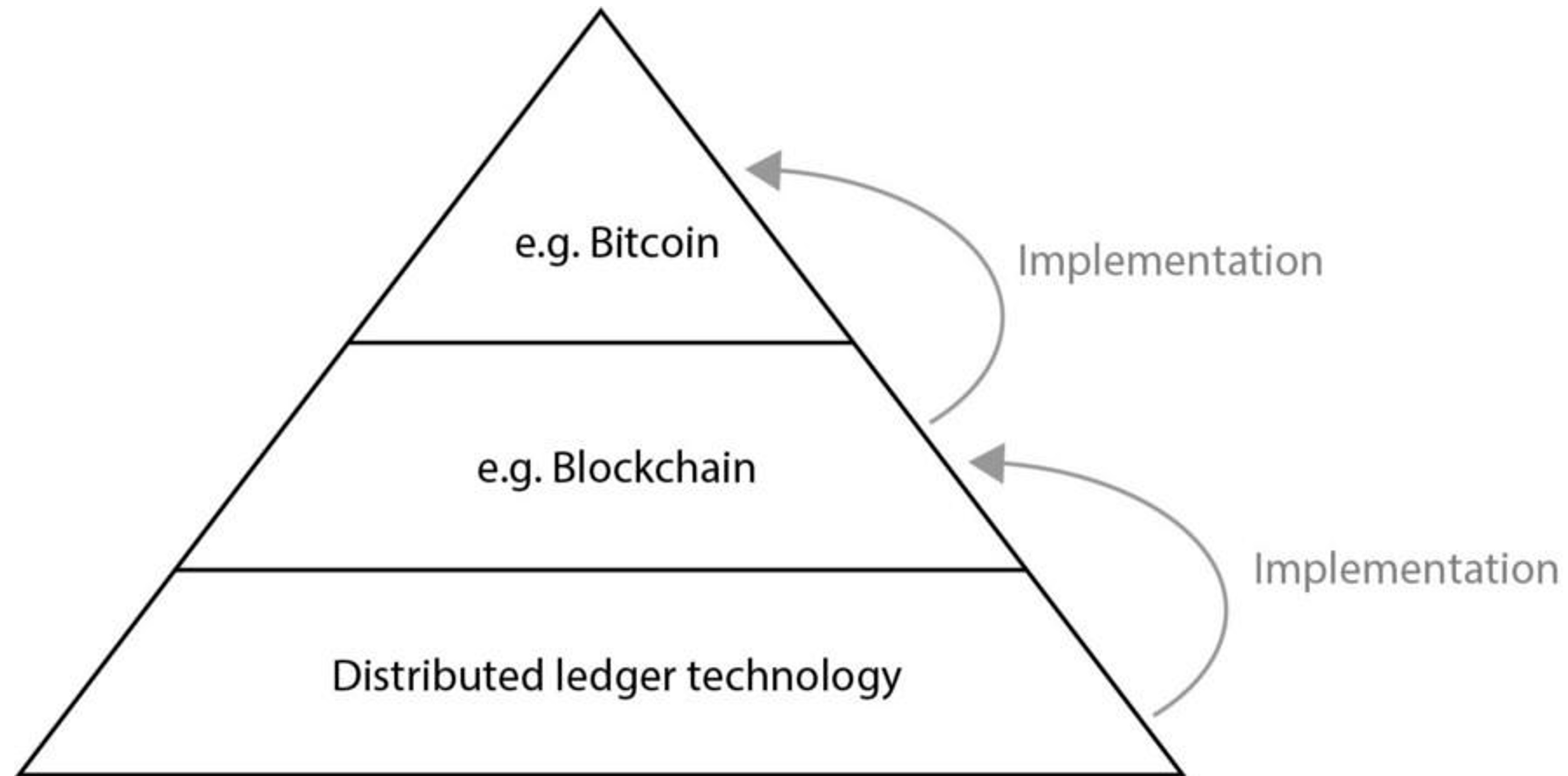
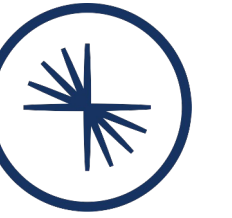
Governance tokens, which are tokens that represent voting rights on the functional parameters of smart contracts

Non-fungible tokens (NFTs), which are a unique, indivisible digital asset with provable ownership

Digital assets that represent claims on real-world assets such as commodities, real estate, or intellectual property, and are "tokenized" into divisible digital assets on the blockchain

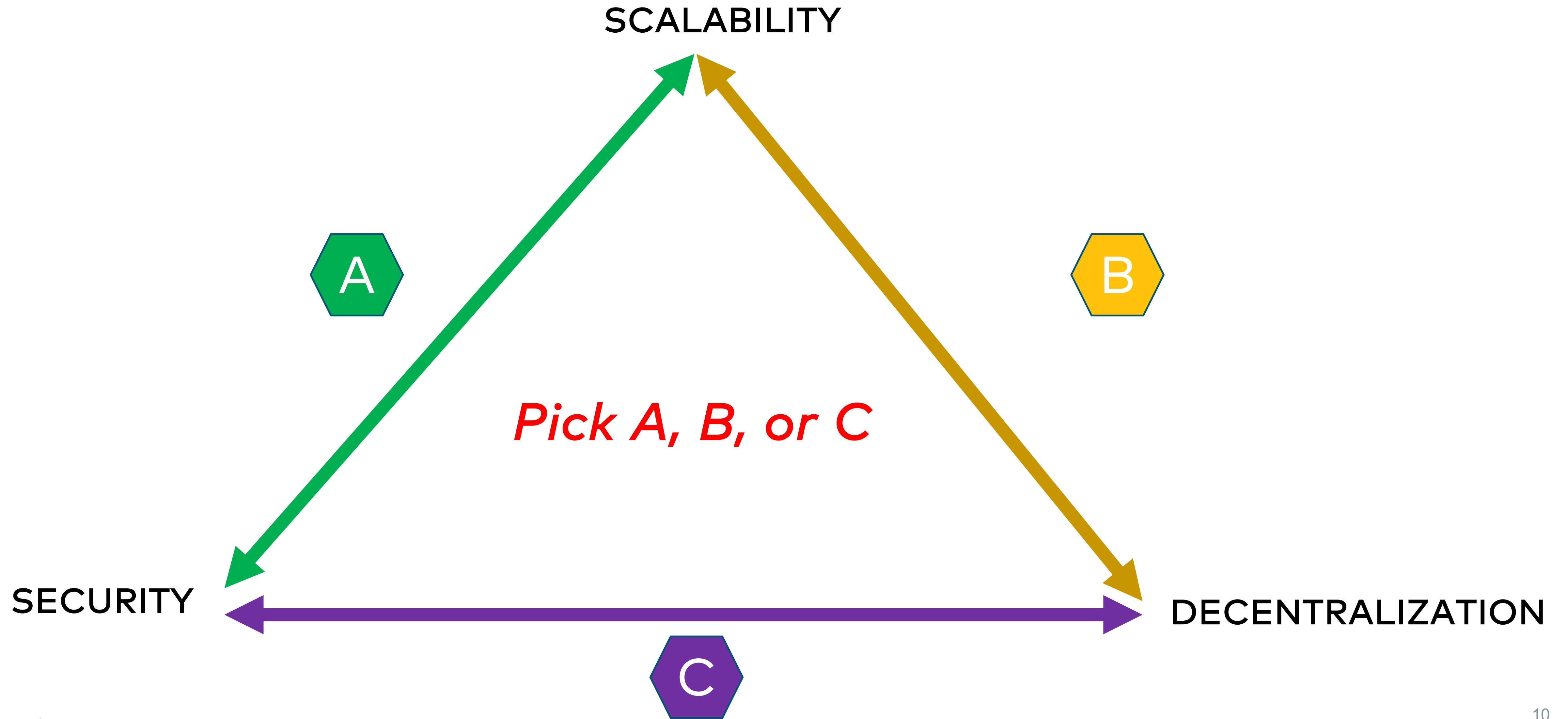
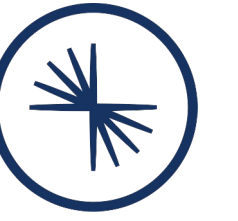
<https://www.mckinsey.com/industries/financial-services/our-insights/web3-beyond-the-hype>

Blockchain vs. Distributed Ledger Technology (DLT)

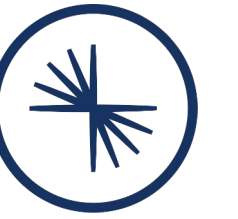


<https://www.da.ventures/post/gaining-clarity-on-key-terminology-bitcoin-versus-blockchain-versus-distributed-ledger-technology>

The Cryptocurrency Trilemma



Challenges and Concerns of Blockchains



- Technical and organizational complexity
- Transaction speed and scalability
- Energy consumption
- Security
- Data tenancy
- Lifecycle costs
- Ease of integration



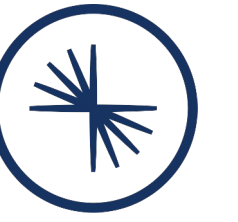
HYPERLEDGER



ethereum

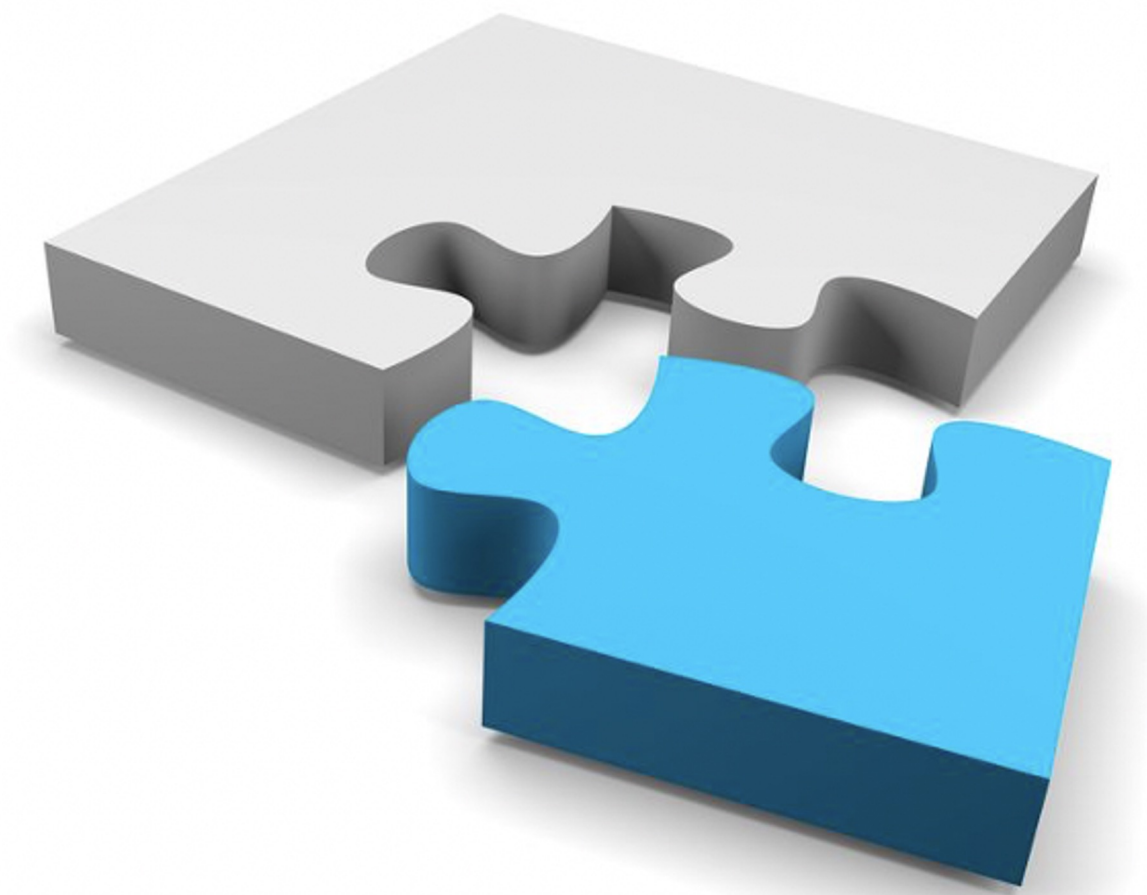
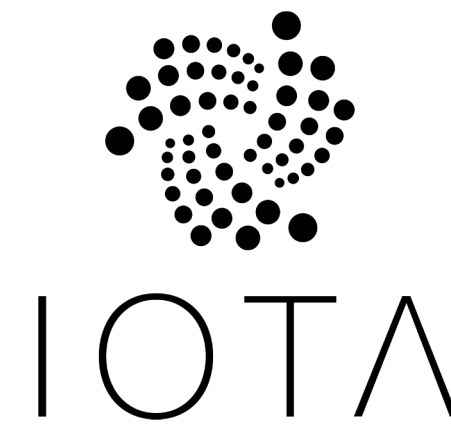
Non-Blockchain DLTs

to solve the challenges to provide better performance and scalability



Purpose-built solutions

- Ripple: Payment settling, currency exchange and remittance system
- Diem (former Libra): Global payment system and financial infrastructure
- IOTA: Internet of Things transactions

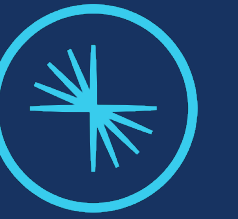


Open and flexible solution for many use cases

- Apache Kafka: Open, scalable, real time Data Streaming

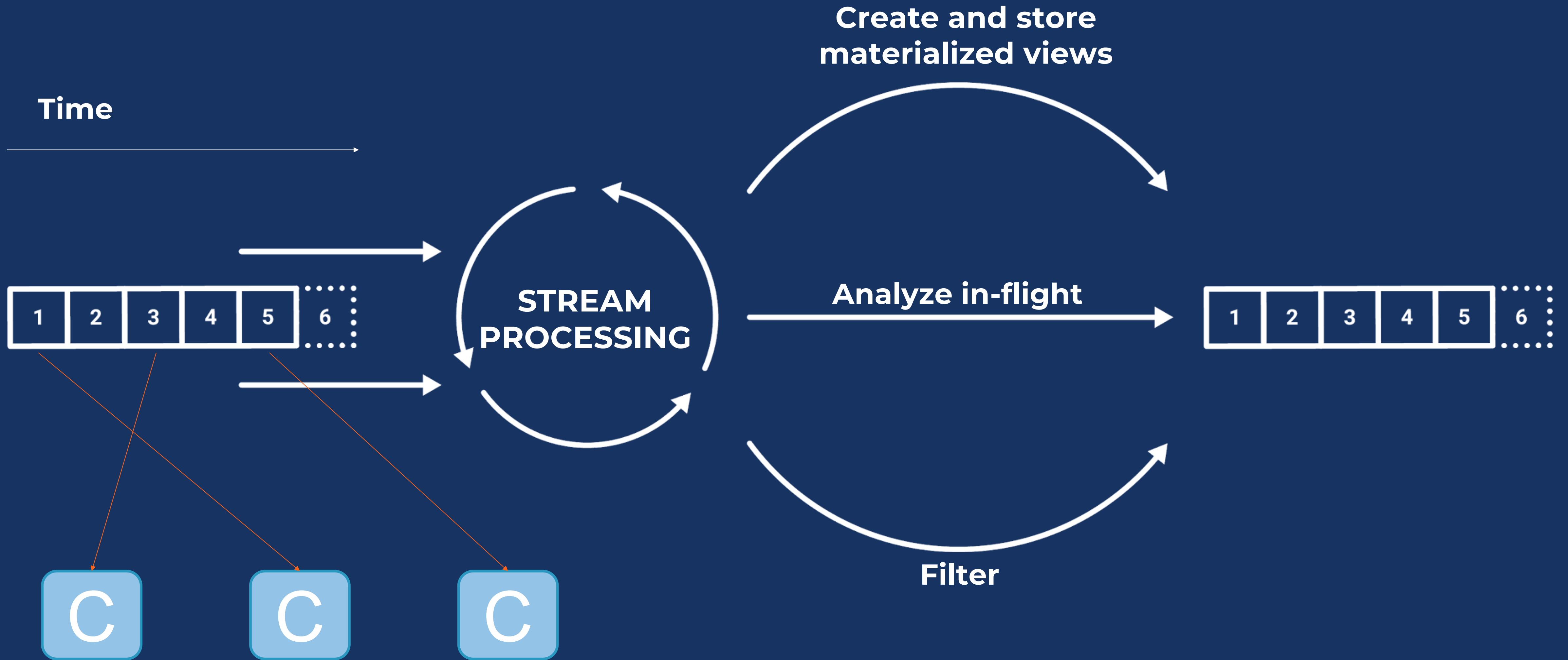
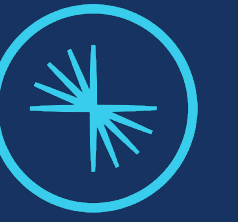


Agenda

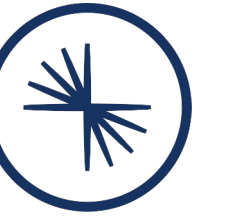


1. When (not) to use Blockchain and Crypto
- 2. Data Streaming for Real-time Analytics at Scale**
3. Examples for Data Streaming with Crypto Data
4. Real-World Projects
5. Choosing the Right Tool for the Job

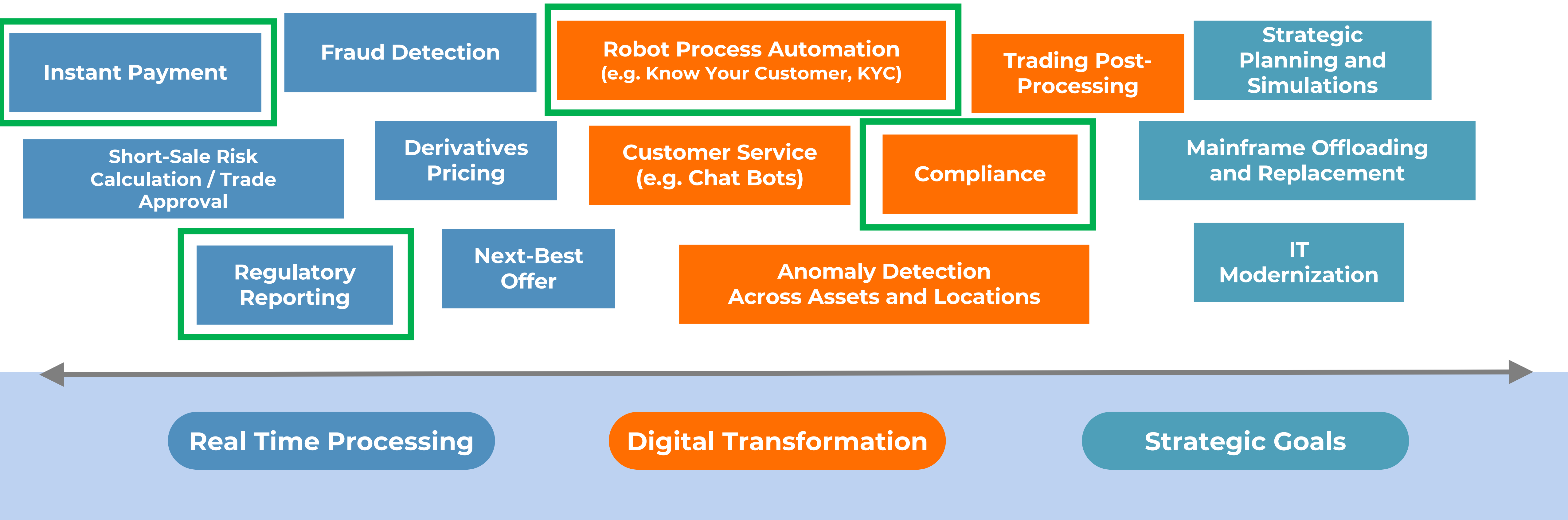
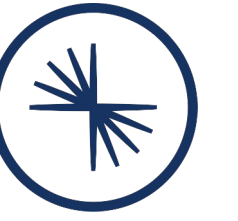
Data Streaming



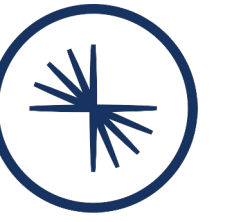
Data Streaming in the Finance Industry



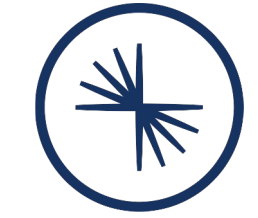
Tamper-Proof Data Streaming for (some of the) Use Cases in the Finance Industry



Apache Kafka – The de facto standard for Data Streaming is the Underpinning of an Event-driven Architecture

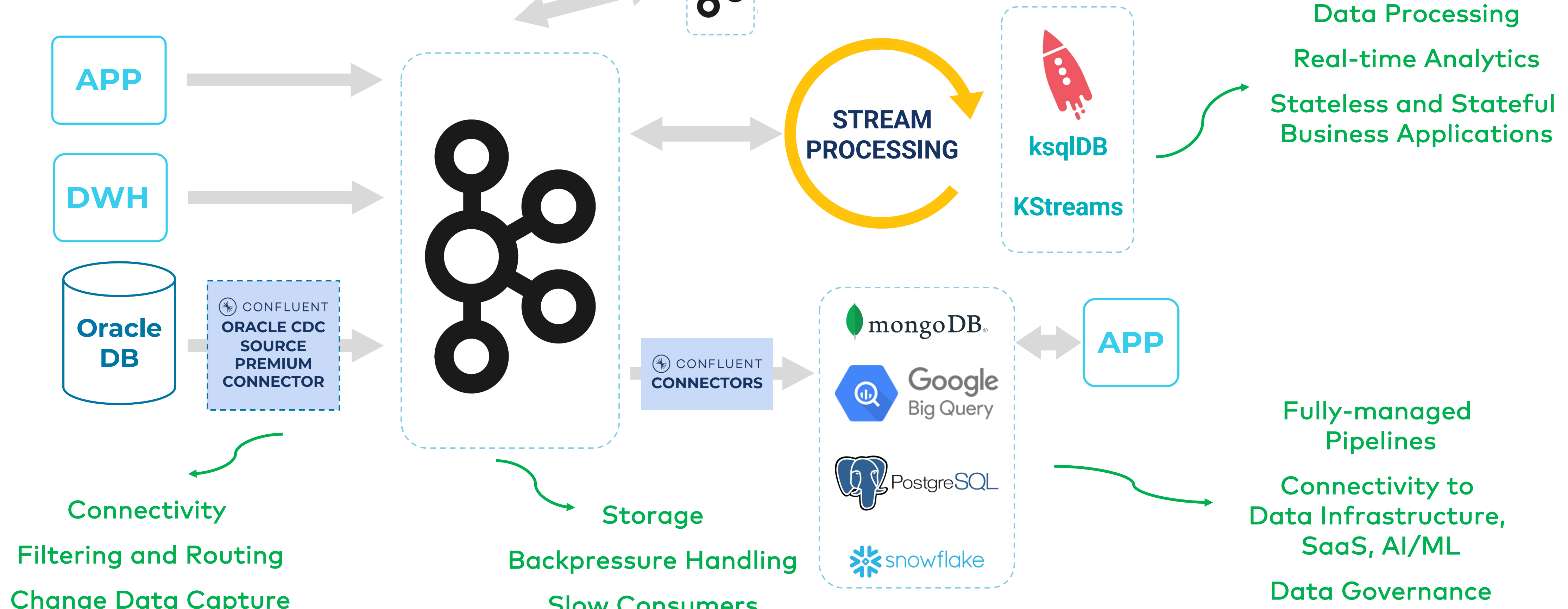


Kafka is a cloud-native data streaming platform!



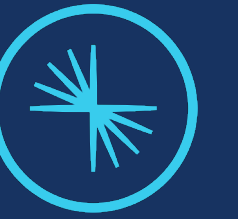
→ More than just data ingestion or message queue

Real-time Data Sharing
across Hybrid and Multi-Cloud



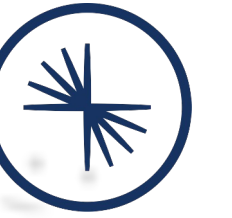
Built-in Scale and Fault Tolerance

Agenda



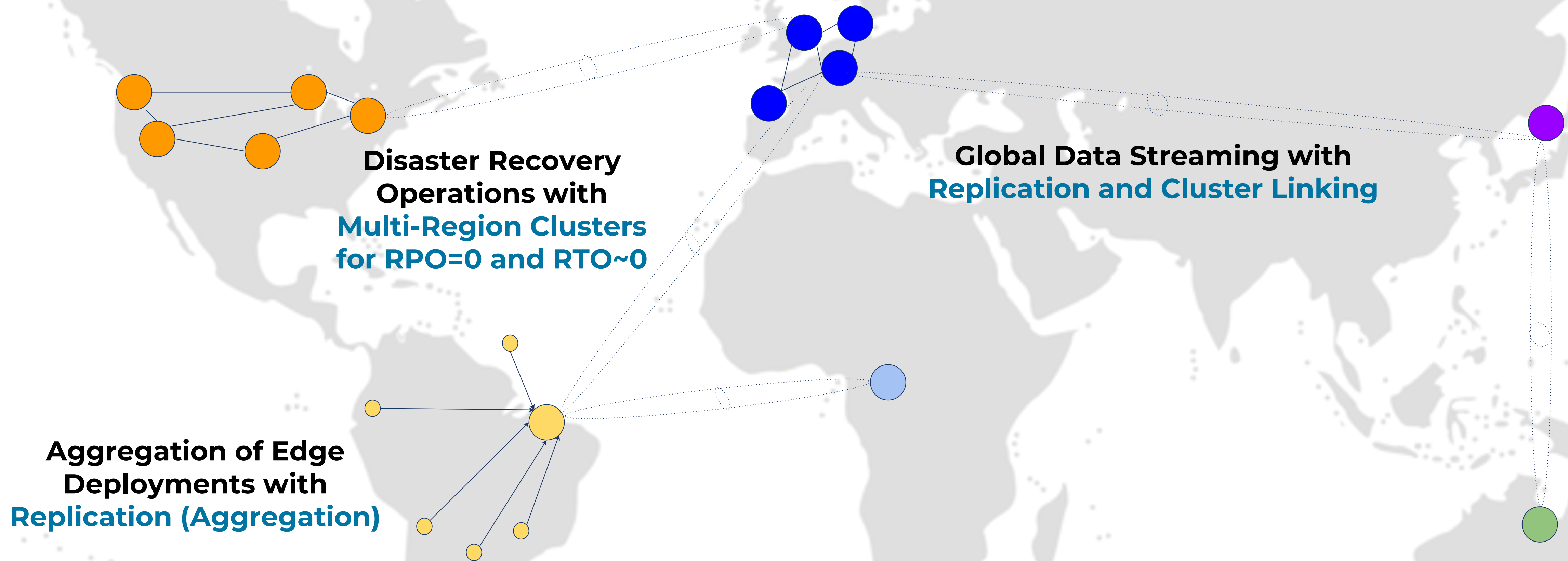
1. When (not) to use Blockchain and Crypto
2. Data Streaming for Real-time Analytics at Scale
- 3. Examples for Data Streaming with Crypto Data**
4. Real-World Projects
5. Choosing the Right Tool for the Job

Data Streaming across the Globe



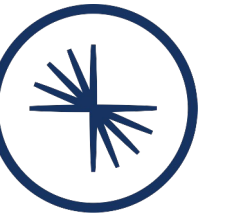
Streaming Replication between Kafka Clusters

Bridge to Databases, Data Warehouses, Data Lakes, Apps, APIs, SaaS





kafka is not a Blockchain!

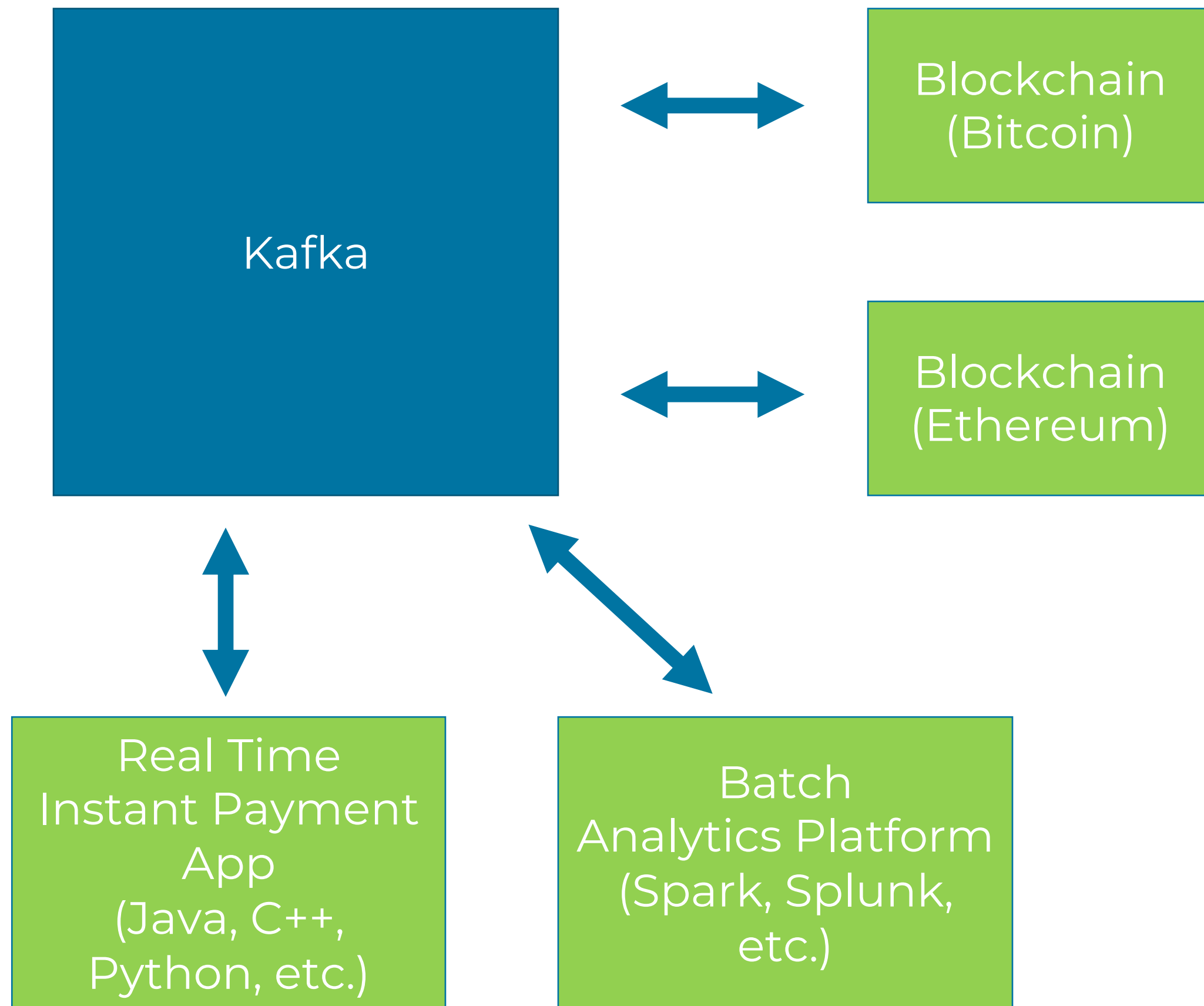


Real-time
High Throughput
Decentralized database
Distributed log of records
Immutable log
Replication
High availability
Decoupling of applications / clients
Role-based access control to data

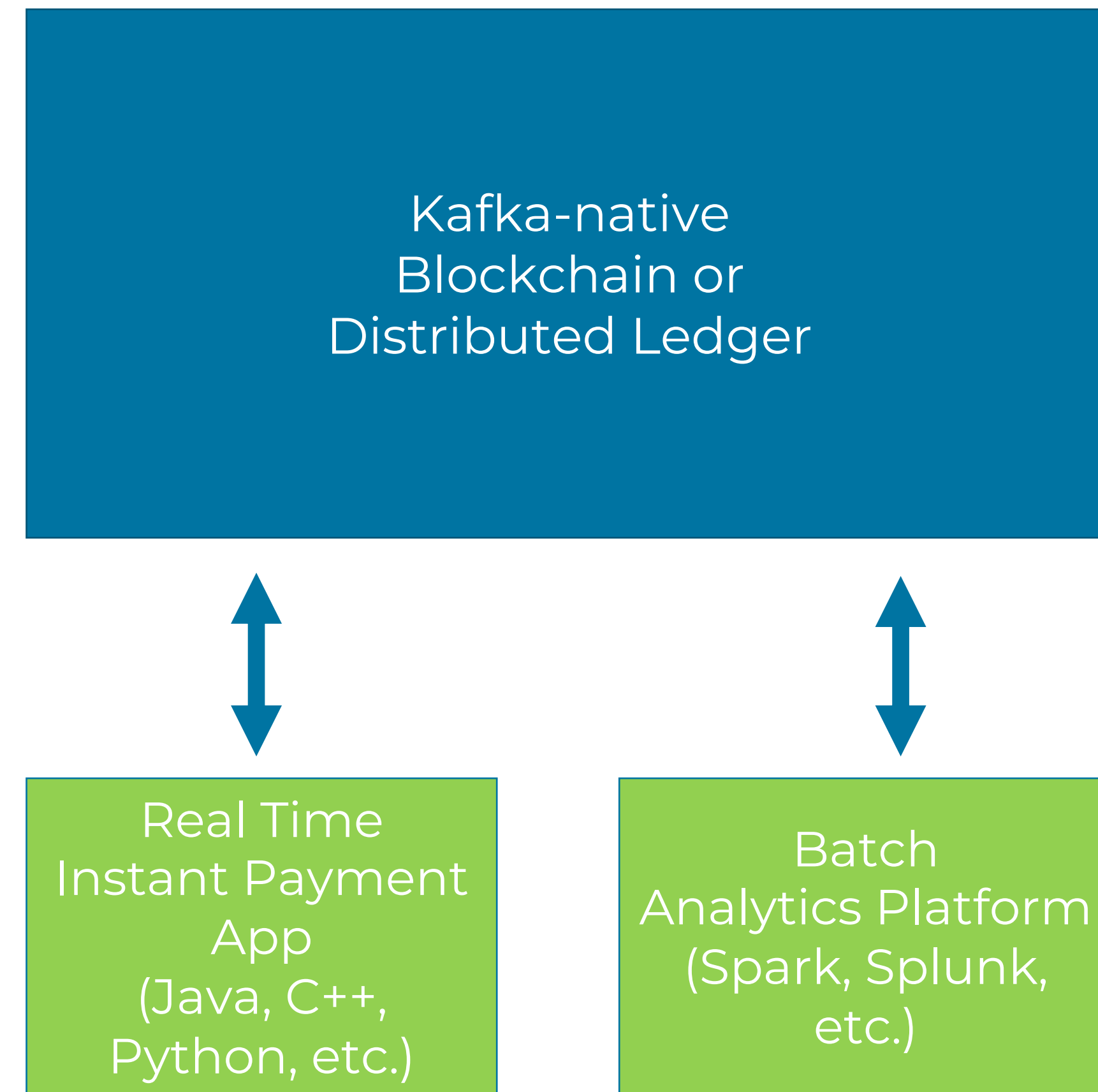
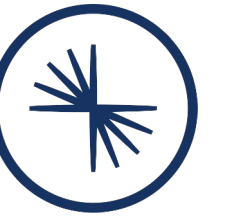
Tamper-Proof
Encrypted payloads
Deployment across independent organizations



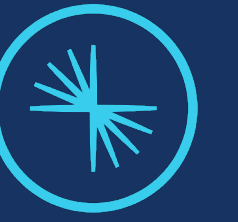
Kafka AND Blockchain



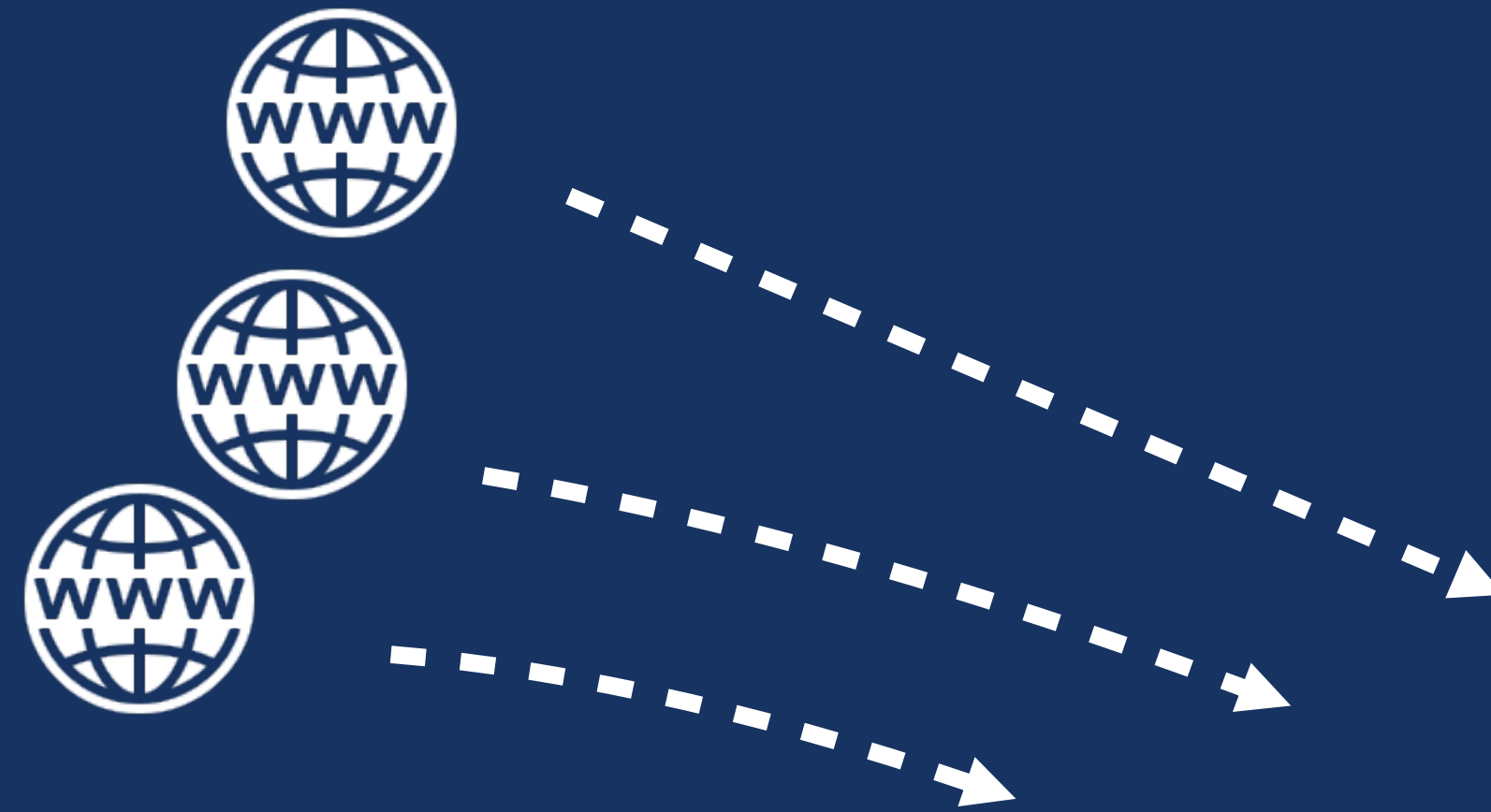
Kafka AS Blockchain



Data Producers for Crypto Cybersecurity



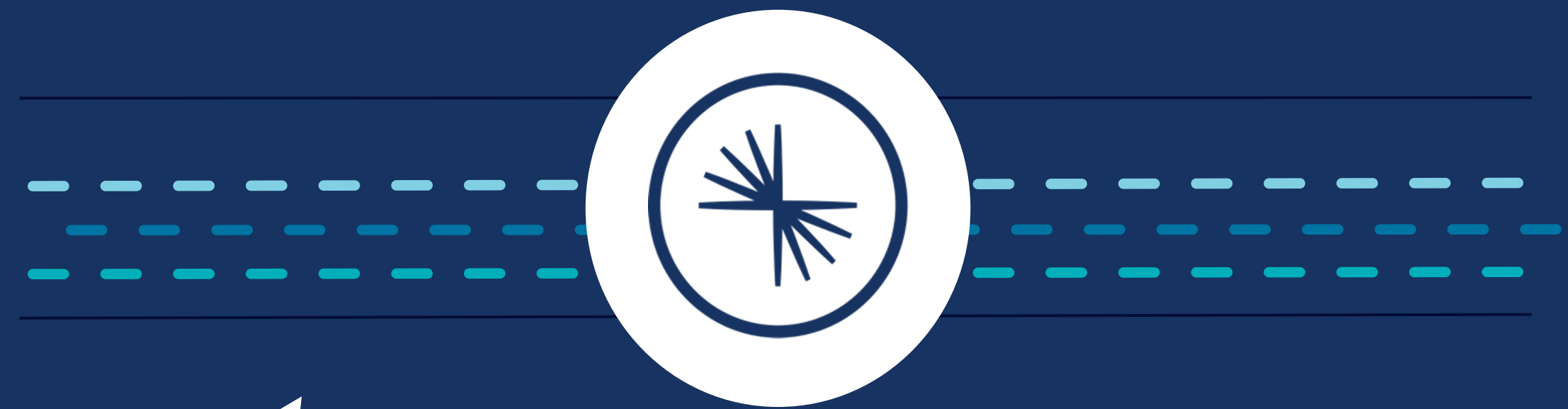
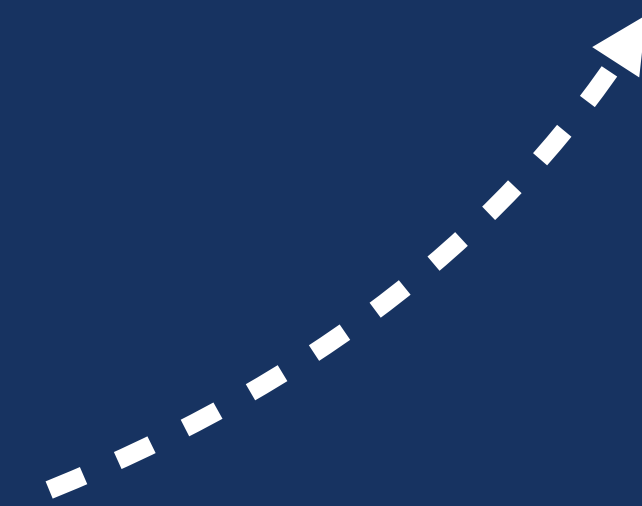
Application Logs
for Analytical Workloads



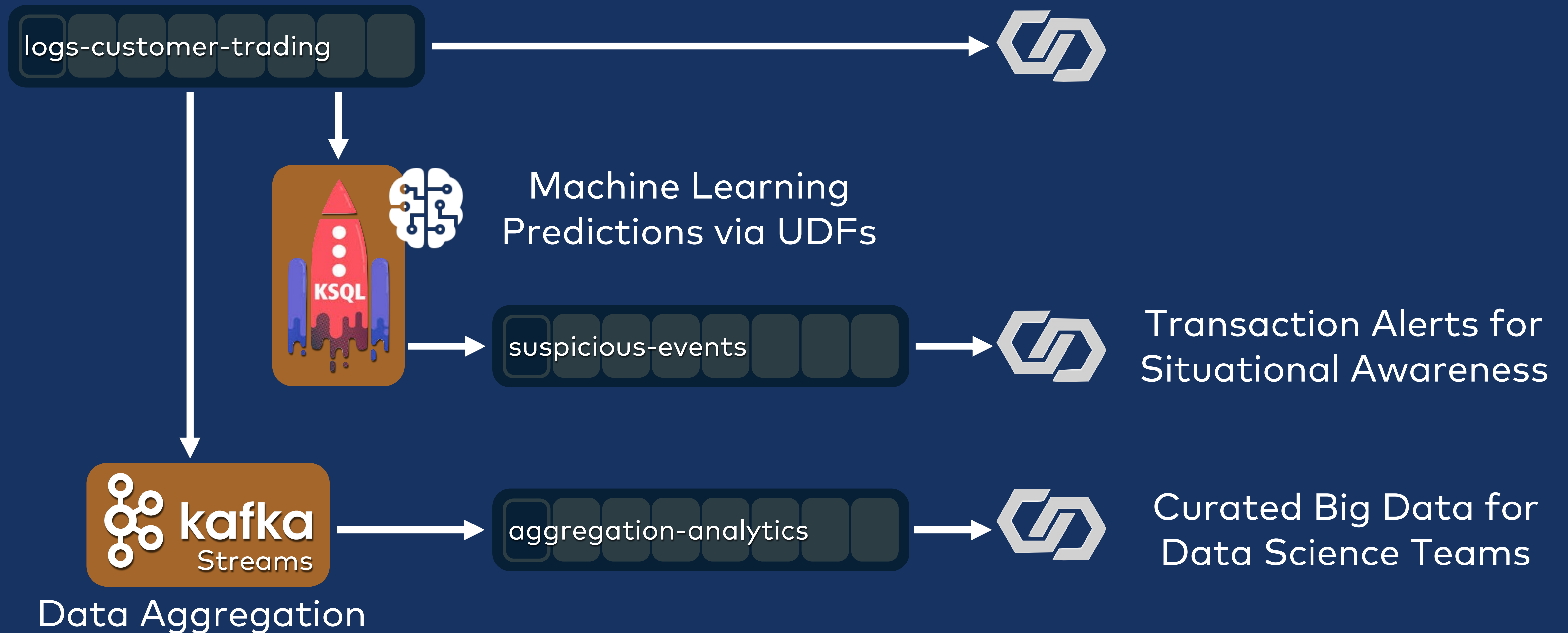
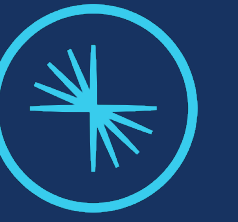
Trading /
Payments
for Transactional
Banking Workloads



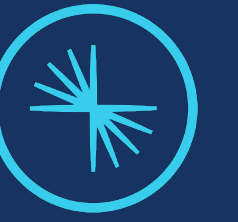
3rd Party Crypto Data
as Foundation of the Crypto Application



Real-Time Data Processing for Crypto Threat Intelligence



Data Consumers for Alerting and Regulatory Reporting



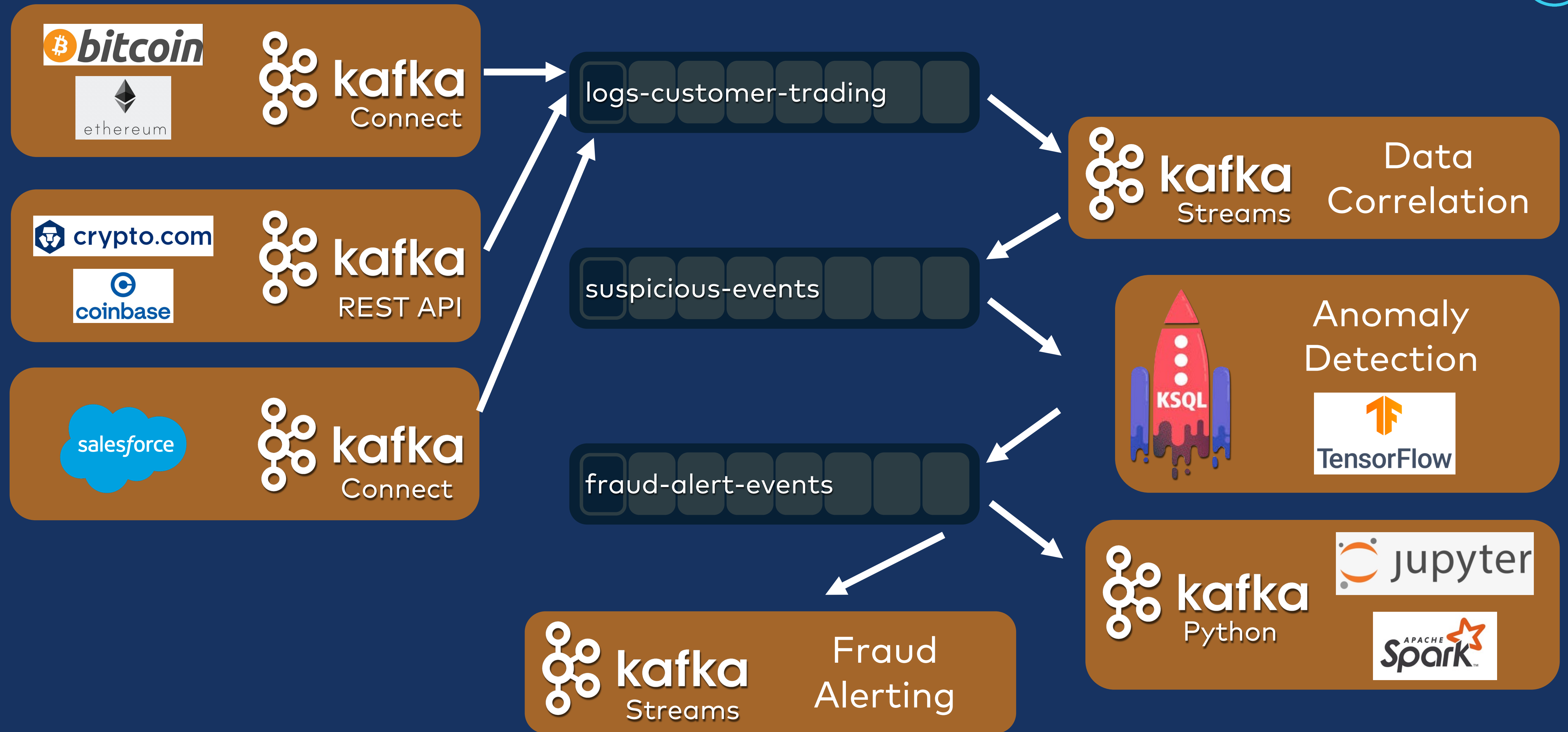
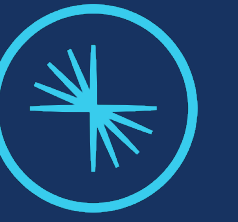
No constraints on integration flows

Data curation on the fly

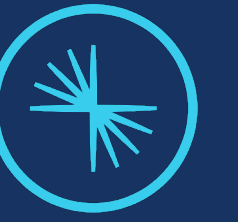
Flexible choice of (multiple) consumers



Kafka as Data Hub for Real-Time Cyber Analytics in the Crypto World



Metaverse: Live commerce with real-time data correlation



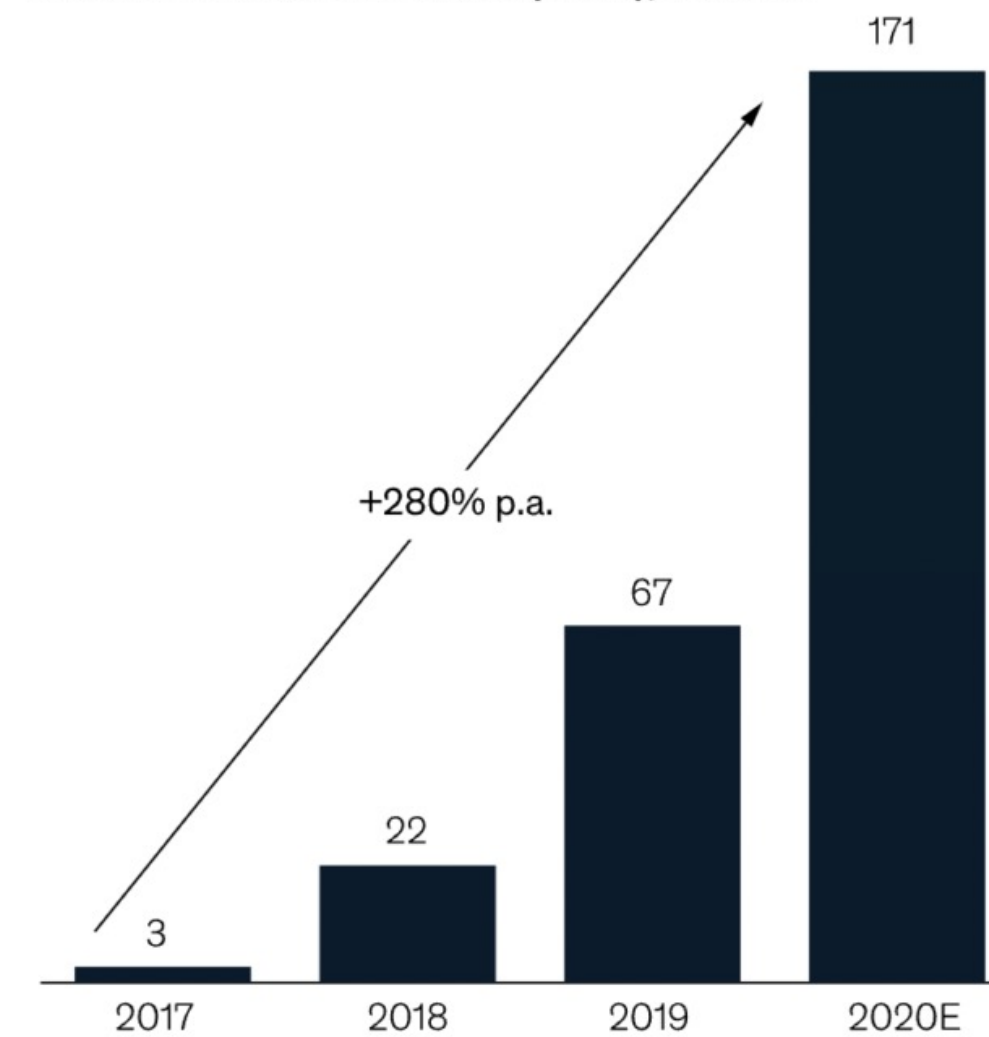
Online and offline shopping via social commerce

Including integration of CRM, loyalty, inventory, chatbots, location-based services, augmented reality, etc.



China's live commerce reached an estimated \$171 billion in value in 2020.

Gross merchandise value (GMV), \$ billion¹



¹Total GMV generated by livestreaming in B2C; includes mainstream brands, influencer brands, and refunded items.
Source: Everbright Securities; iResearch; McKinsey analysis

McKinsey
& Company



Data Streaming as Data Hub for Crypto, Blockchain, and Metaverse

Interact with other people in the metaverse.
Upsell rare items for your games.
Provide context-specific pricing.
Integrate with crypto wallets and NFT markets.
All automated. In real-time. At scale.

**Mobile App
+ AR Glasses**
(Swift on iOS)

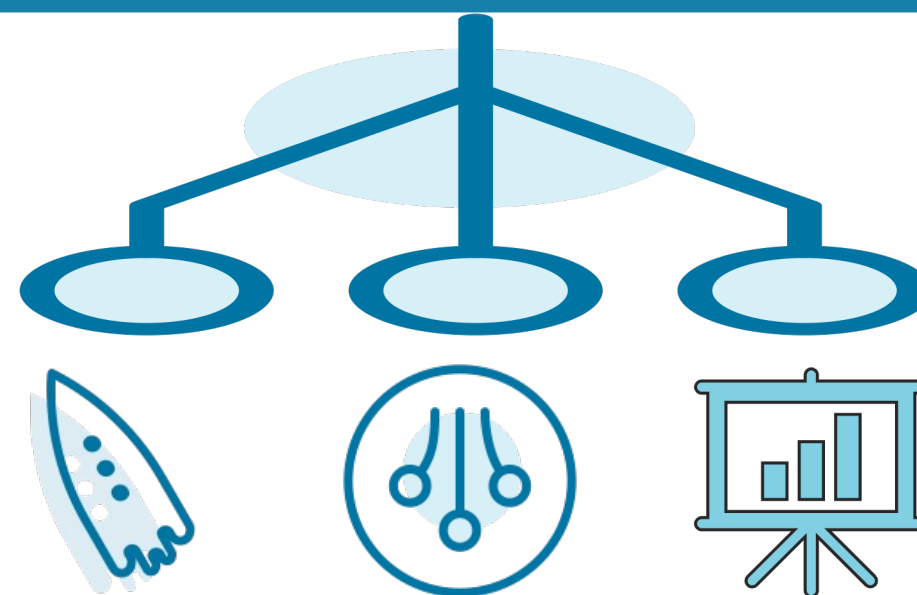


(1) Interest in rare game item

(4) Receive NFT recommendation and context-specific price



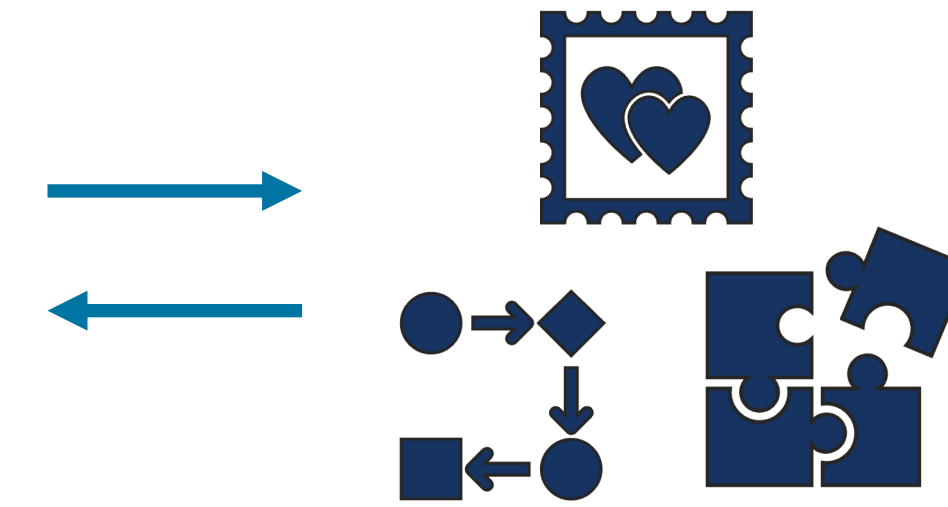
(5) Order and pay NFT for game item



Data Consolidation

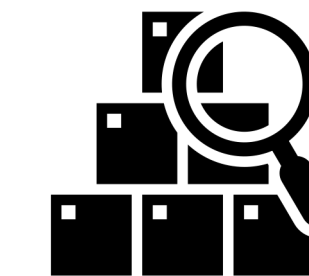
Streaming ETL

(Kafka Connect + Kafka Streams)



(2) Live Q&A with chatbot

(3) Context-specific pricing engine



Chatbot

External Speech NLP
(REST API)

CRM / Loyalty

Salesforce

(Kafka Connect)

NFT Marketplace

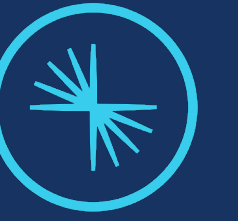
Live monitoring and item bidding process
(ksqlDB + OpenSea API)

Crypto Payment Service
Buy Now Pay Later (BNPL)

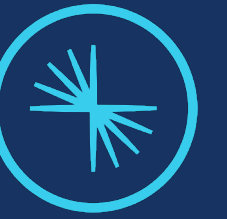
(Ethereum API)

(6) Execute Crypto transaction and confirm NFT

Agenda

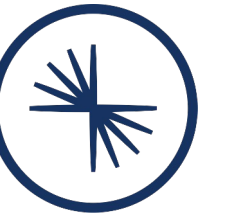


1. When (not) to use Blockchain and Crypto
2. Data Streaming for Real-time Analytics at Scale
3. Examples for Data Streaming with Crypto Data
- 4. Real-World Projects**
5. Choosing the Right Tool for the Job



Kafka AS Blockchain

R3 Corda: Kafka AS Blockchain / DLT



Rearchitected: V5 migrates to Kafka

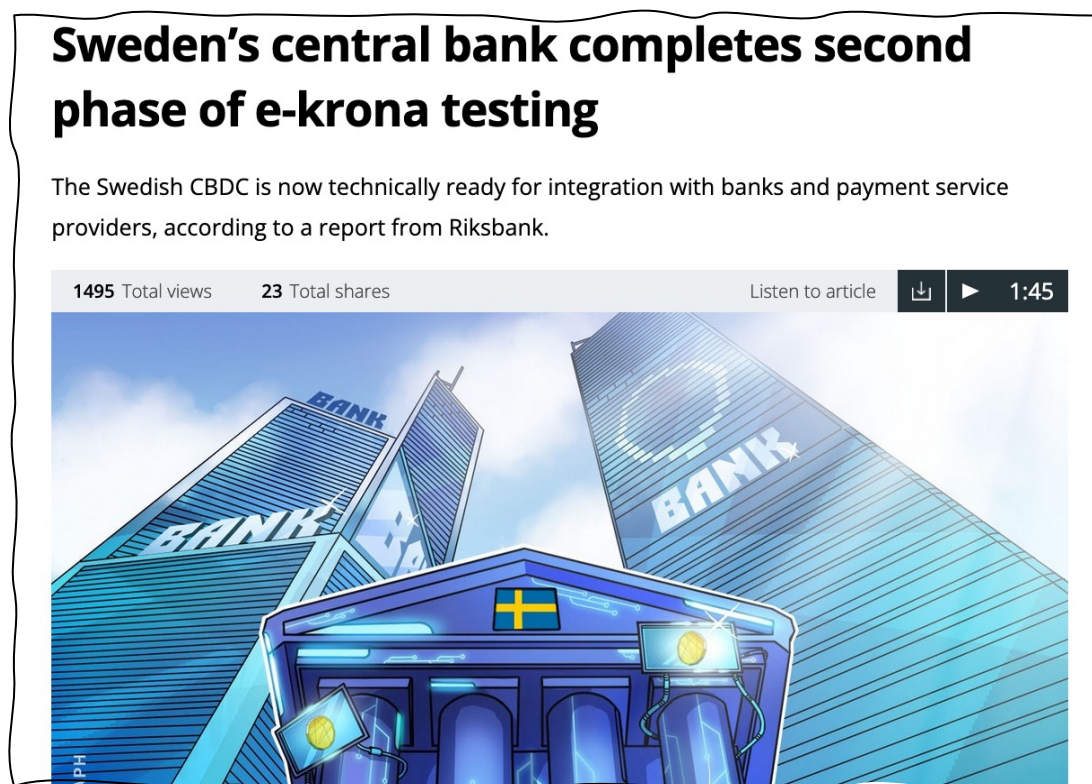
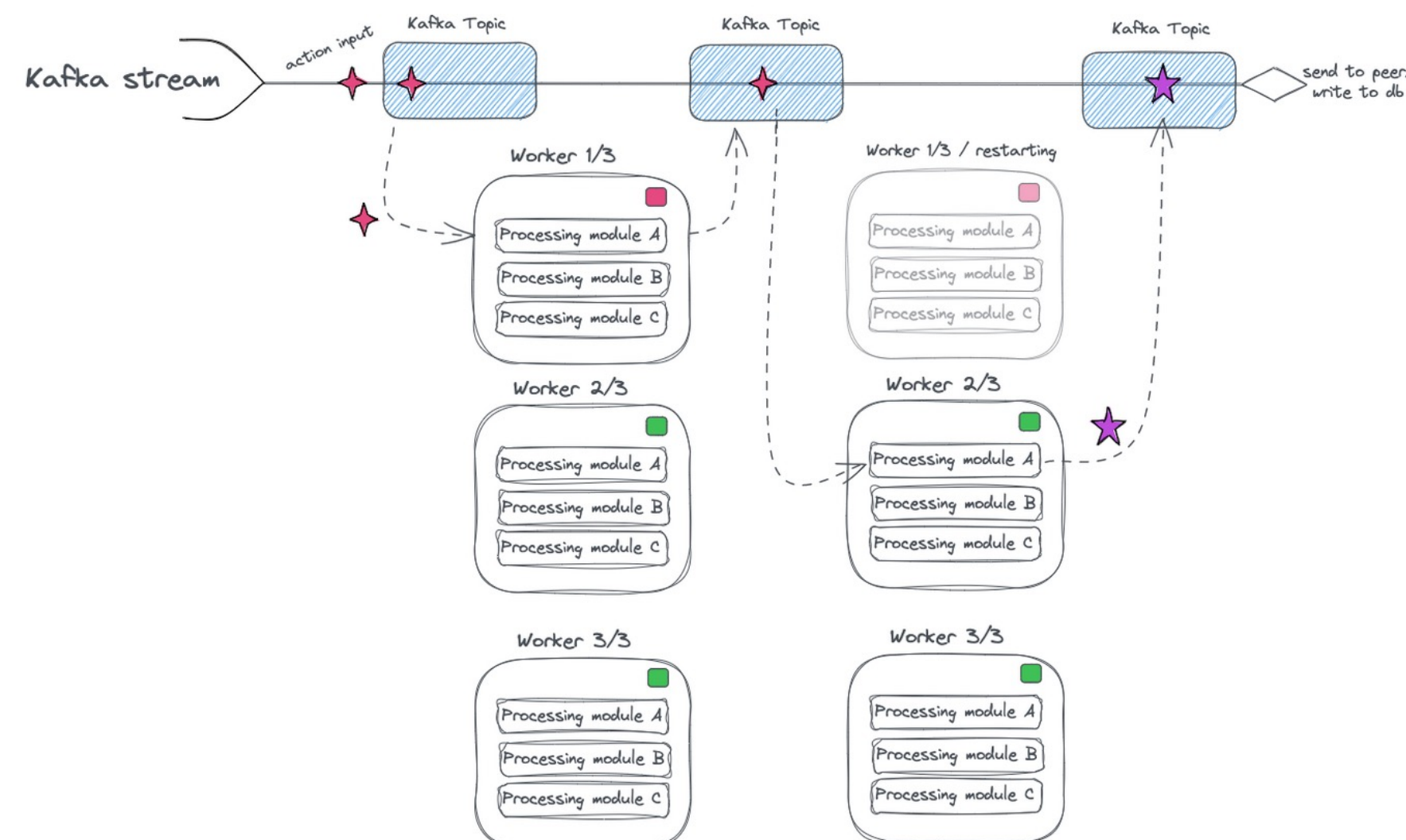
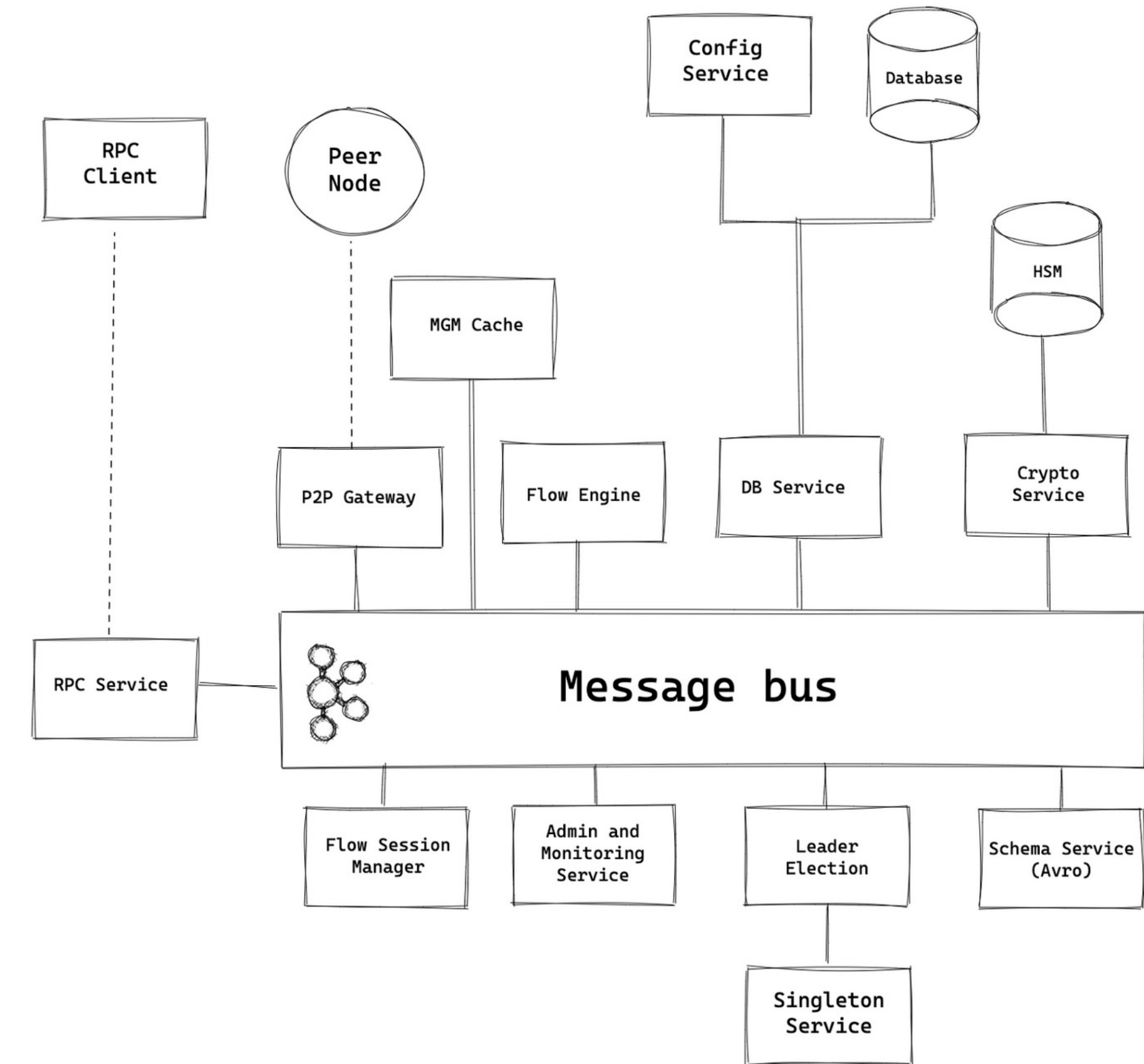
High availability for critical services + cost-effective way to scale (horizontally) to deal with 'burst-y' and high-volume throughputs

Fully redundant, worker-based architecture

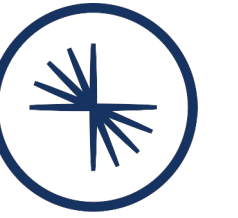
Kafka cluster as the message broker to facilitate communication between node services

Compacted topics (logs) as the mechanism to store and retrieve most recent states

Key components: Process engine, database integration, Crypto integration, RPC service (HTTP), Monitoring



<https://medium.com/corda/corda-5-the-road-ahead-part-3-introduction-to-the-architecture-8c43bb19ac81>



Industry standard oracle network for connecting smart contracts to the real world

Transitioning from traditional time series-based monitoring toward an event-driven architecture and alerting approach

Data Engineer, Observability

APPLY FOR THIS JOB

NEW YORK CITY / REMOTE / ENGINEERING / REMOTE - FULL-TIME

Be the data engineer powering the solutions to a unique Observability challenge -- monitoring uptime and reliability of independent/3rd party oracle providers. Observability at Chainlink Labs is going through a transition from traditional time series-based monitoring toward an event-driven architecture and alerting approach. You will have a significant impact as we grow the Chainlink ecosystem and ensure the best experience for our customers by ensuring reliable uptime.

You'll develop and build highly scalable, secure, and reliable software that will change the way smart contracts function at a fundamental level. You'll have the opportunity to learn and master the latest research concerning distributed systems, cryptography, blockchains, game theory, consensus algorithms, and decentralized applications.

You will be given a high level of autonomy/ownership over your projects, the opportunity to expand your scope of knowledge, and the chance to help build the decentralized infrastructure of the future.

<https://jobs.lever.co/chainlink/b0094754-6d1e-4c55-8dbc-622bc4d8a5ab>

Your Impact

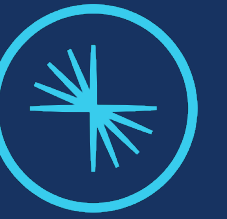
- Lead the design and deployment of data pipelines that power our real time monitoring/observability services to detect and alert the team of needed action.
- Make recommendations to ensure sufficient metrics are collected to create alerts with every new feature release.
- Thinking creatively about attack vectors, possible failures, and disaster scenarios, modeling them in reproducible test environments, and developing fixes
- Implementing resilient distributed systems to achieve extremely high reliability in a variety of blockchain environments

Responsibilities

- 3+ years of professional experience as a software developer / DevOps engineer or equivalent
- Experience with Kafka required
- Deep knowledge of go or Kafka Streams apps (including Java/the JVM) a plus
- Experience administering Kafka Connect, Confluent Platform, and/or Kubernetes is a plus
- Experience with test-driven development and the use of testing frameworks
- Strong communication skills, specifically giving/receiving constructive feedback in a collaborative setting

Our Stack

Golang, Kafka, Postgres, Kubernetes, AWS

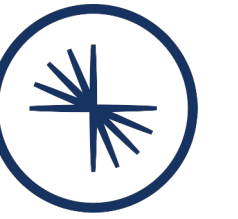


Kafka IN Blockchain Tools

TokenAnalyst



TokenAnalyst



Integrate blockchain data with its analytics tools.

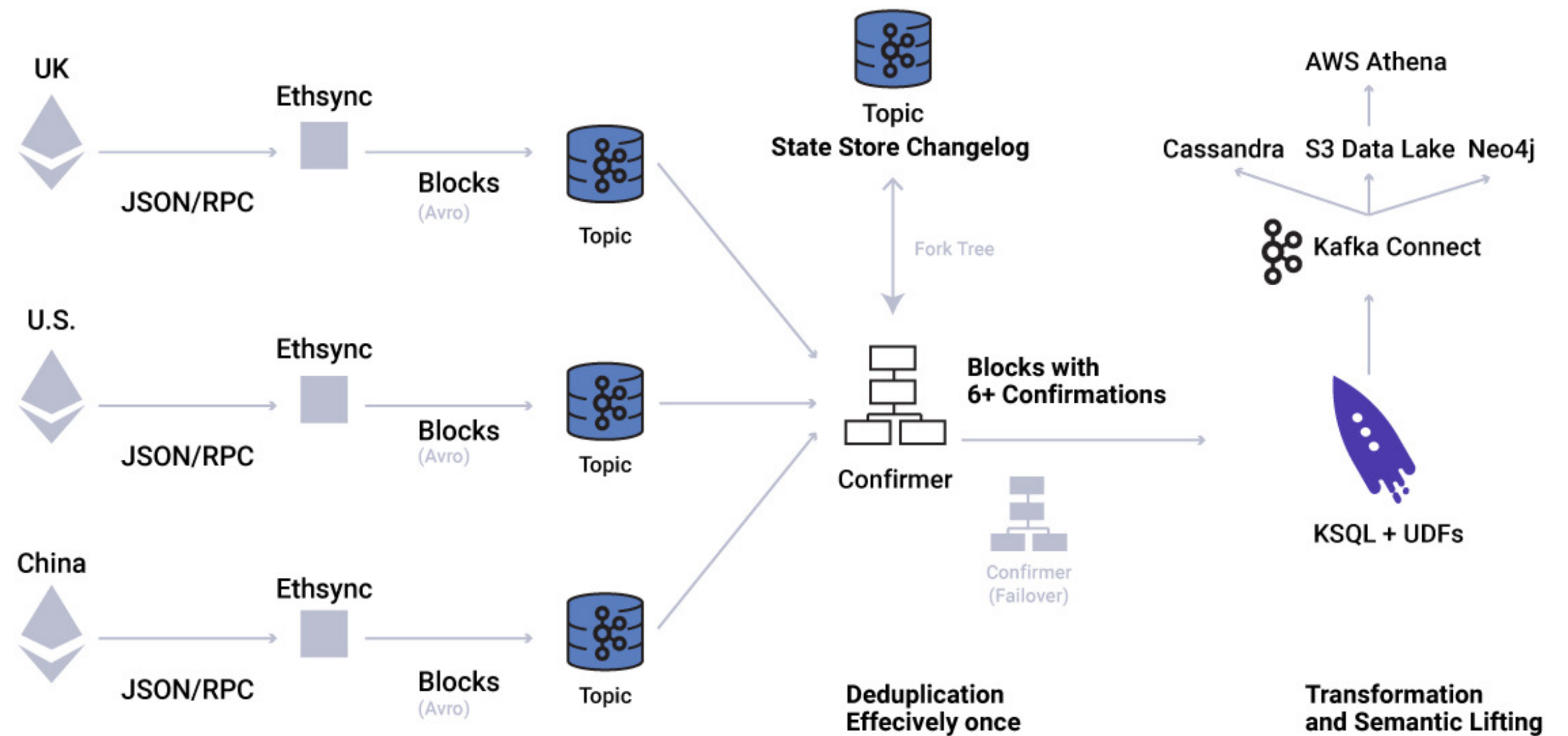
Kafka Streams provides a stateful streaming application to prevent using invalid blocks in downstream aggregate calculations.

Kafka Connect for the integration with databases and data lake.

Blockchain Market Intelligence

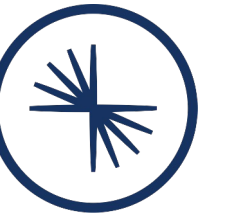
Enterprise-grade data and tools to understand and access blockchains.

Exchange	Token	Inflow	Change	Outflow	Change
Binance	USDT_ERC20	\$94,646,850	-0.06%	\$54,597,677	0.10%
Binance	BTC	\$56,288,631	-0.30%	\$73,647,325	-0.23%
Huobi	BTC	\$47,311,406	0.04%	\$46,265,302	-0.19%
Bitstamp	BTC	\$25,256,040	0.01%	\$19,590,864	-0.06%
Kraken	BTC	\$23,466,015	0.05%	\$26,371,287	-0.21%



<https://www.confluent.io/en-gb/blog/reliable-fast-access-to-on-chain-data-insights/>

EthVM – Blockchain Explorer

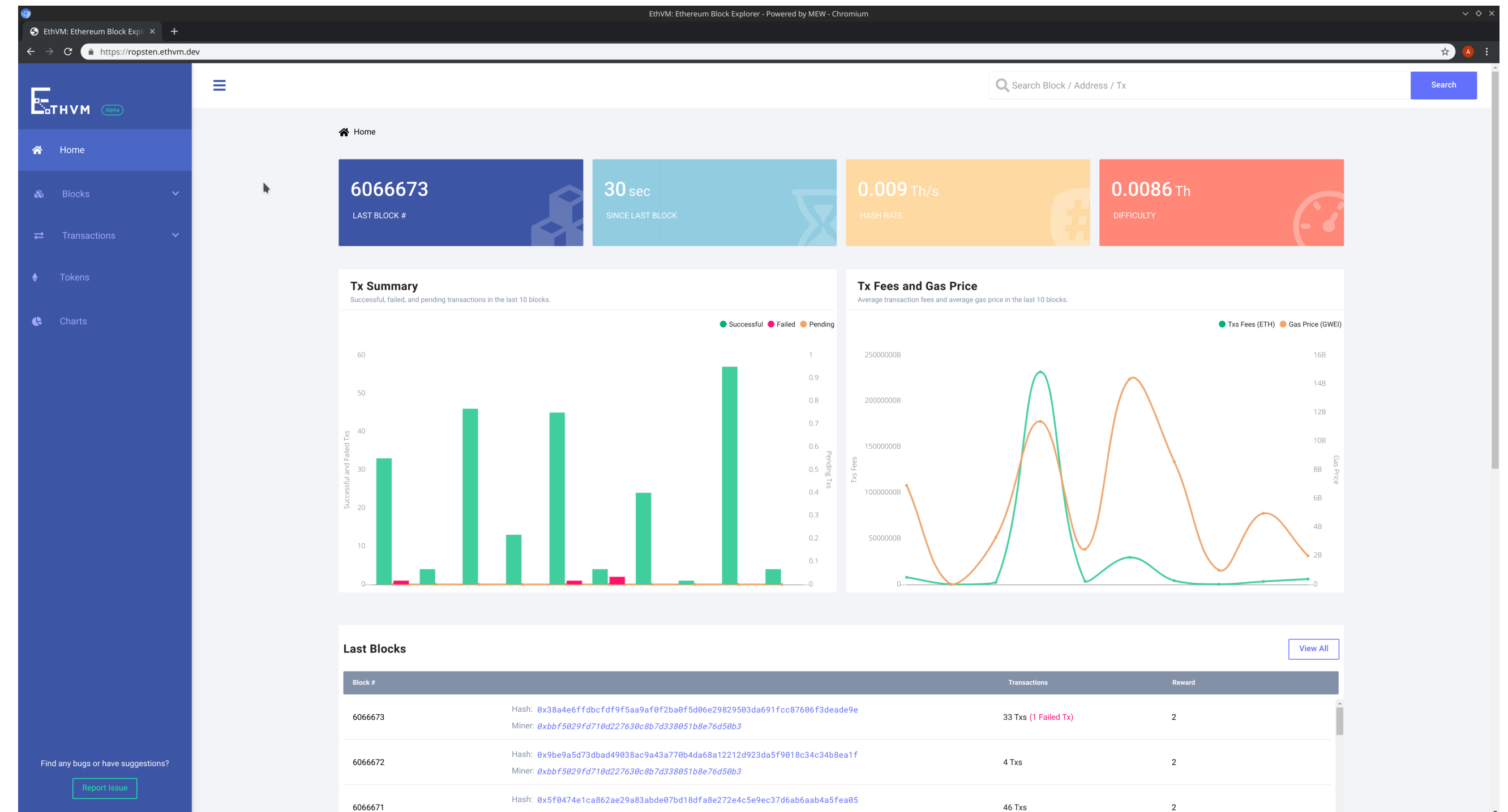
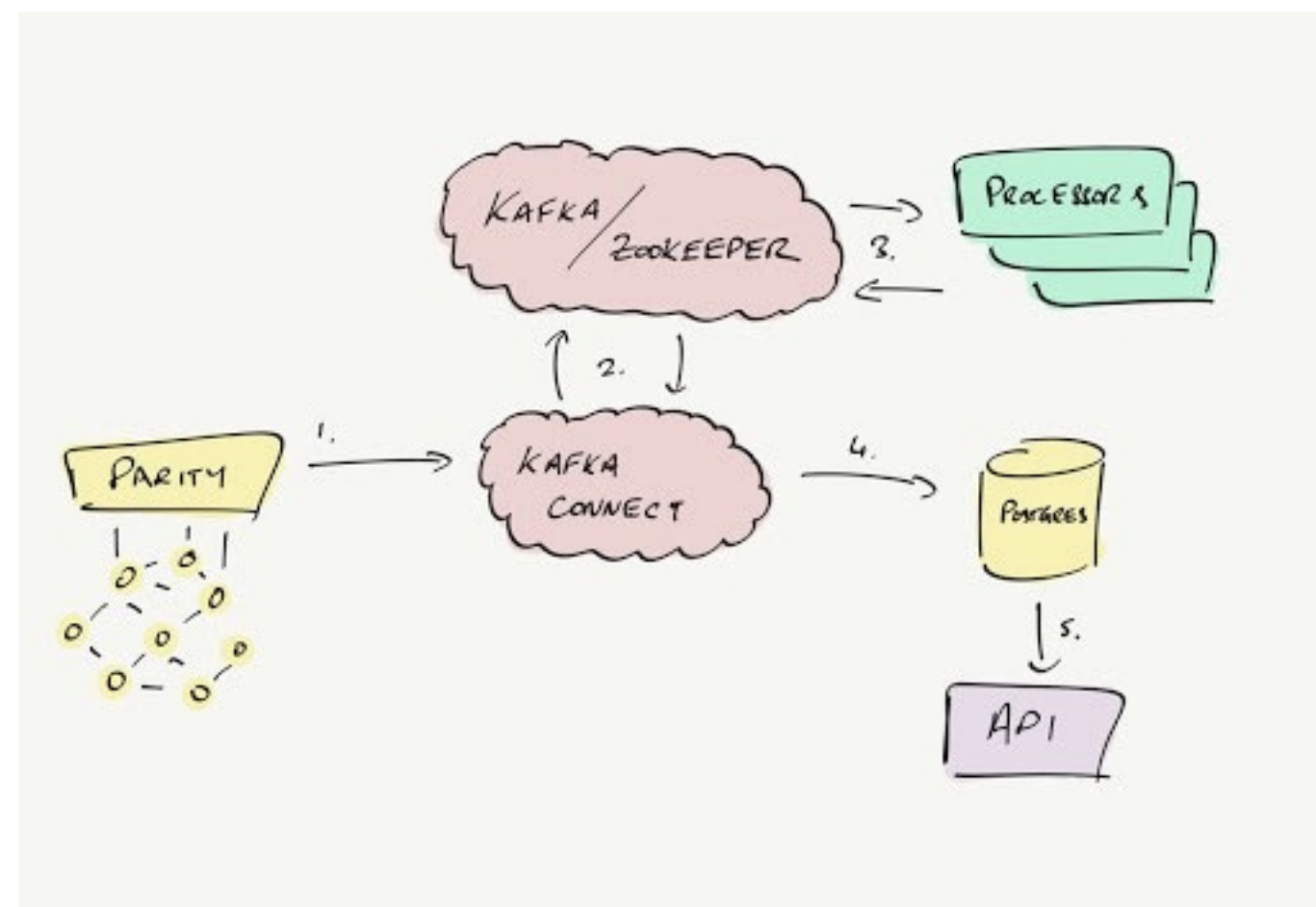


Open-source Ethereum Blockchain Data Processing and Analytics Engine with a client-side Block Explorer

Tool for blockchain auditing and decision

Verify the execution of transactions and smart contracts, check balances, and monitor gas prices

Built with Kafka Connect, Kafka Streams, Schema Registry



<https://github.com/EthVM/EthVM>

<https://www.bitcoininsider.org/article/66671/ethvm-first-open-source-block-explorer-powered-kafka>

kaleido – Rest API for Crypto

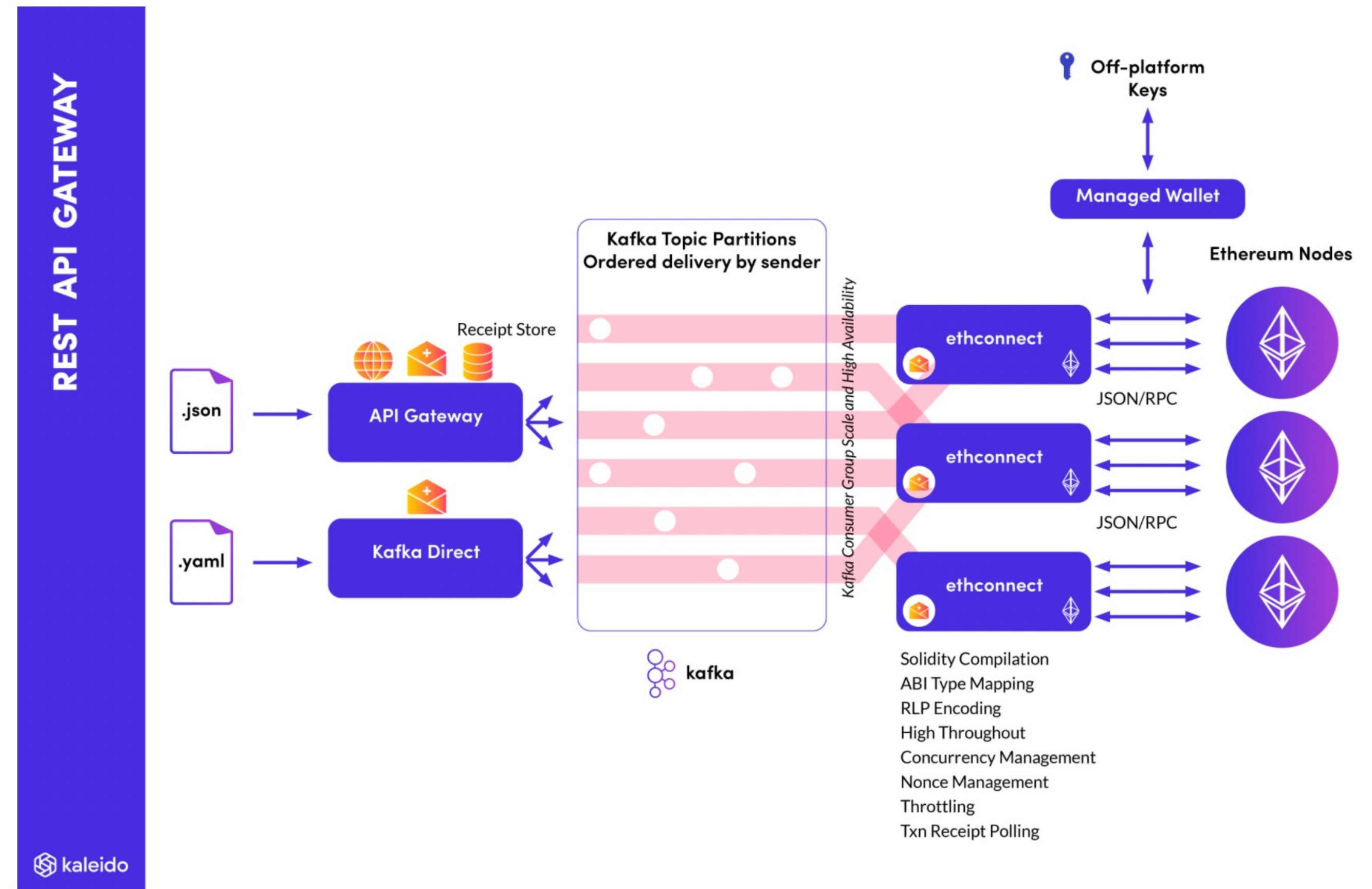


Enterprise grade blockchain APIs to deploy and manage Smart Contracts, send Ethereum transactions, and query blockchain data

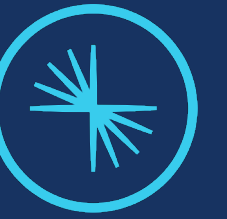
Hides the complexities of Ethereum transaction submission, thick Web3 client libraries, nonce management, RLP encoding, transaction signing, and smart contract management.

REST APIs for your on-chain logic & data

Backed by a fully-managed high throughput Apache Kafka infrastructure



<https://www.kaleido.io/blockchain-platform/rest-api-gateway>

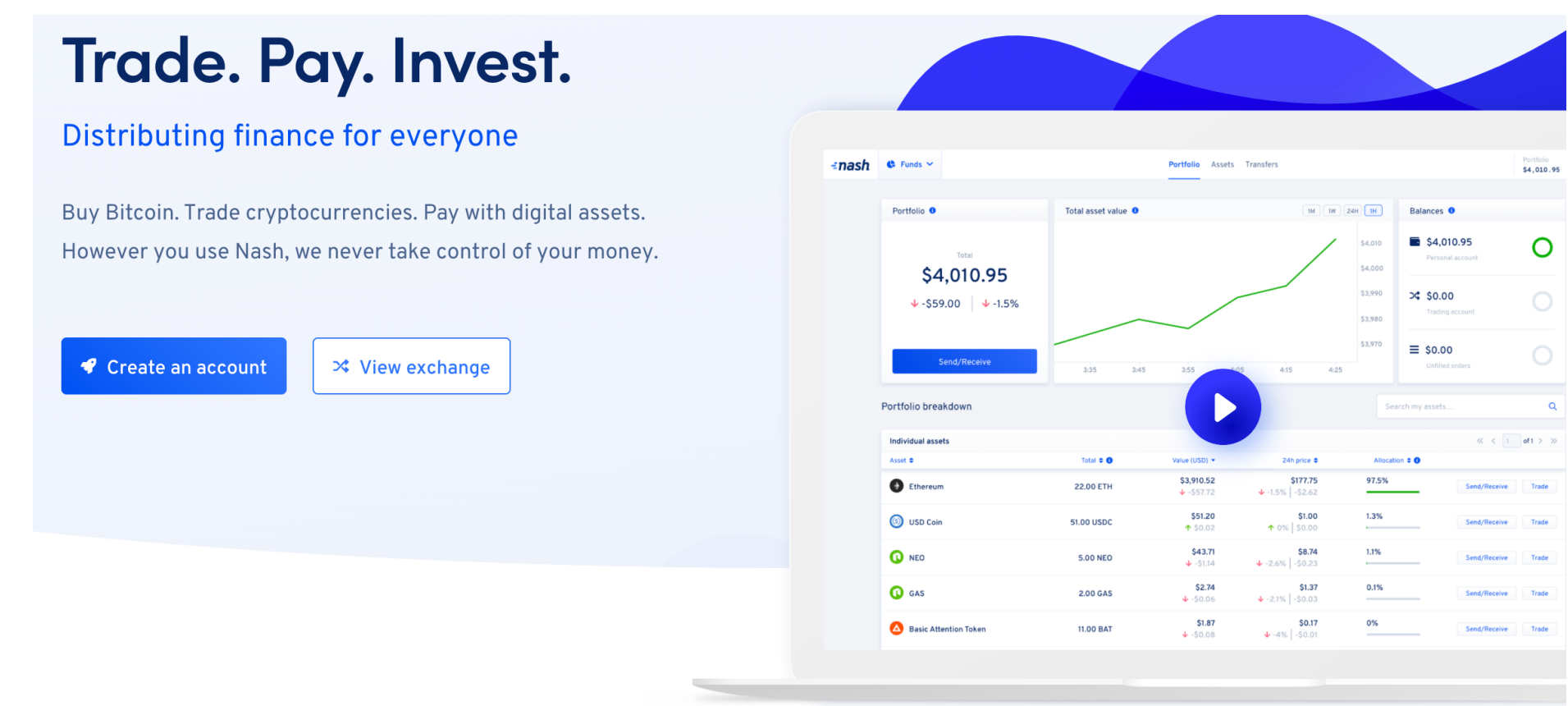


Kafka AND Blockchain

Kafka AND Blockchain: Financial services platform combining Kafka and Blockchains



- Provides the speed and convenience of traditional exchanges and the security of non-custodial approaches.
- Invest in, make payments with and trade Bitcoin, Ethereum, NEO, and other digital assets.
- The exchange is the first of its kind, offering non-custodial cross-chain trading with the full power of a real order book
- Deterministic replayability in its exact order at any time

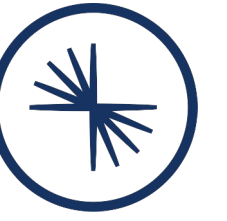


Revolutionary technology built
with the user in mind



<https://community.nash.io/t/first-of-its-kind-non-custodial-exchange-using-confluent-cloud/8254>

CUSTODIGIT – A Platform for Digital Assets



- For cryptocurrencies (bitcoins, etc.) and digital assets
- Secure storage of wallets
- Sending and receiving on the blockchain
- Trading via brokers and exchanges
- Regulated environment

- Kafka as central nervous system
- Workflow orchestration with distributed saga design pattern

CUSTODIGIT

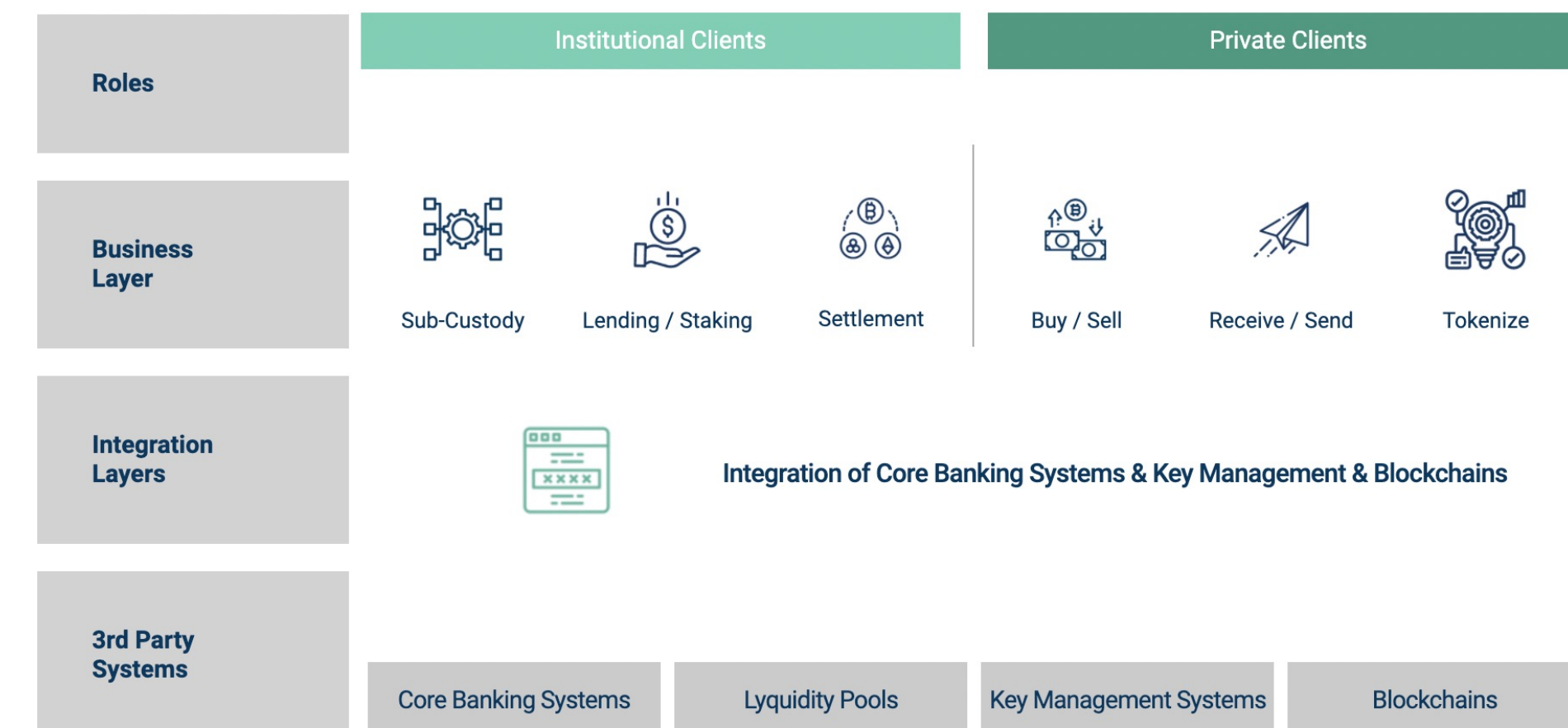
Your gateway to a digital asset offering



Custodigit is the easiest and most secure way for regulated financial service providers to offer digital assets to their clients.

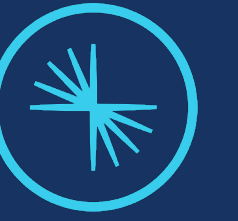
Seamless digital asset ecosystem

Our integration layer provides a reliable, seamless connection between your core banking system, every major liquidity provider in the industry, and your client interface.



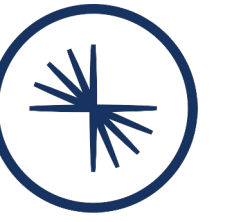
<https://events.confluent.io/datainmotiontour20211>

Agenda



1. When (not) to use Blockchain and Crypto
2. Data Streaming for Real-time Analytics at Scale
3. Examples for Data Streaming with Crypto Data
4. Real-World Projects
5. **Choosing the Right Tool for the Job**

Kafka vs. Blockchain?



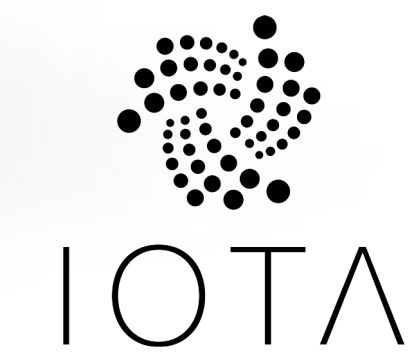
HYPERLEDGER



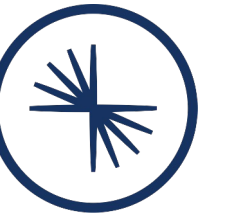
ethereum



r3.c.rda



Kafka vs. Blockchain?

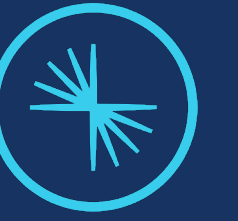


Use Apache Kafka for

- Enterprise infrastructure
- Open, scalable, real-time requirements
- Flexible architectures for many use cases

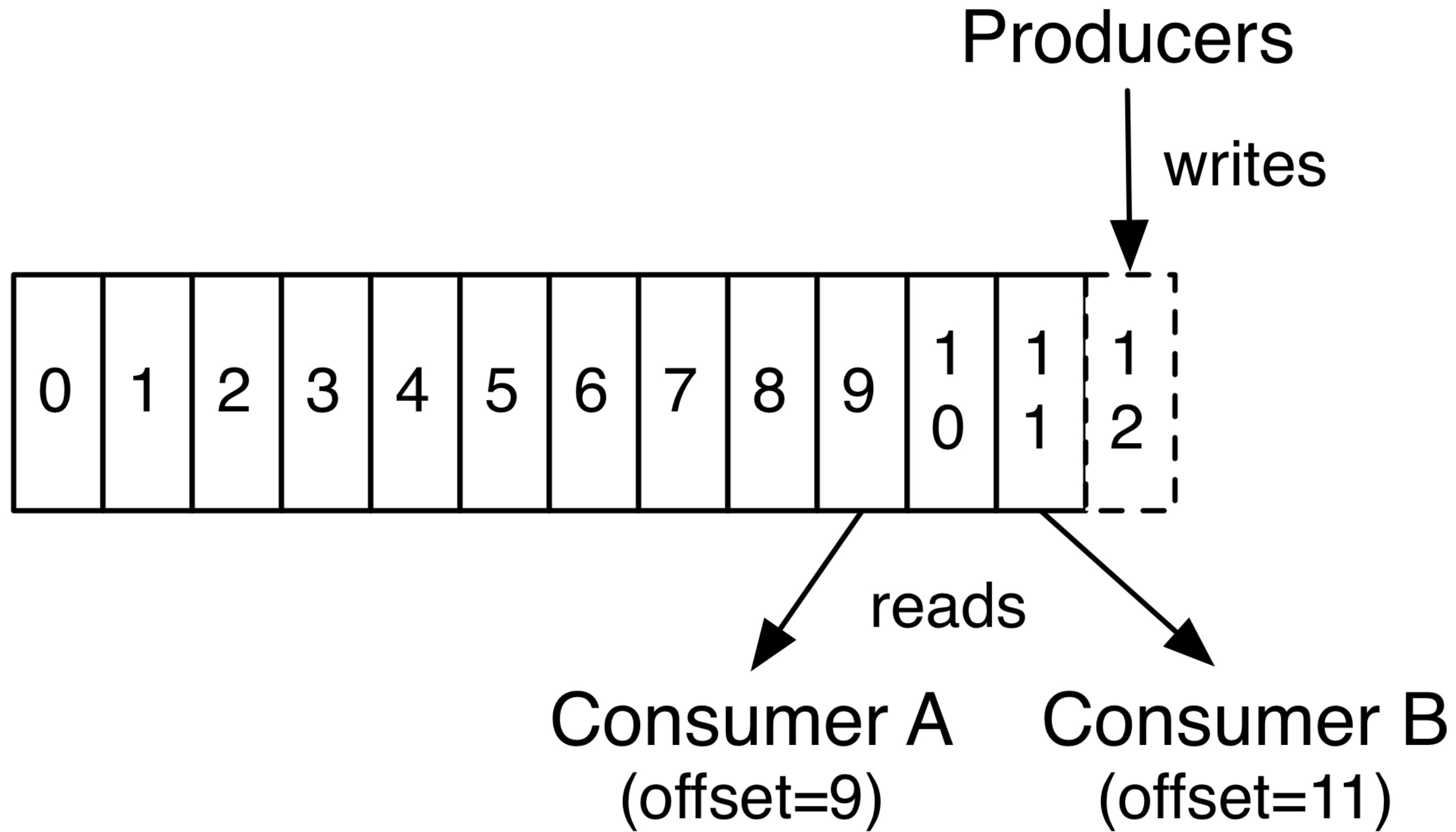
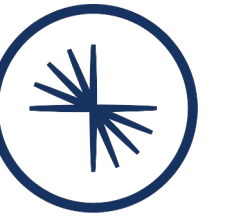
Use Blockchain for

- Deployment over various independent organizations
 - Participants verify the distributed ledger contents themselves.
- Specific use cases
- Server-side managed and controlled by multiple organizations
- Scenarios where the business value overturns the added complexity and project risk

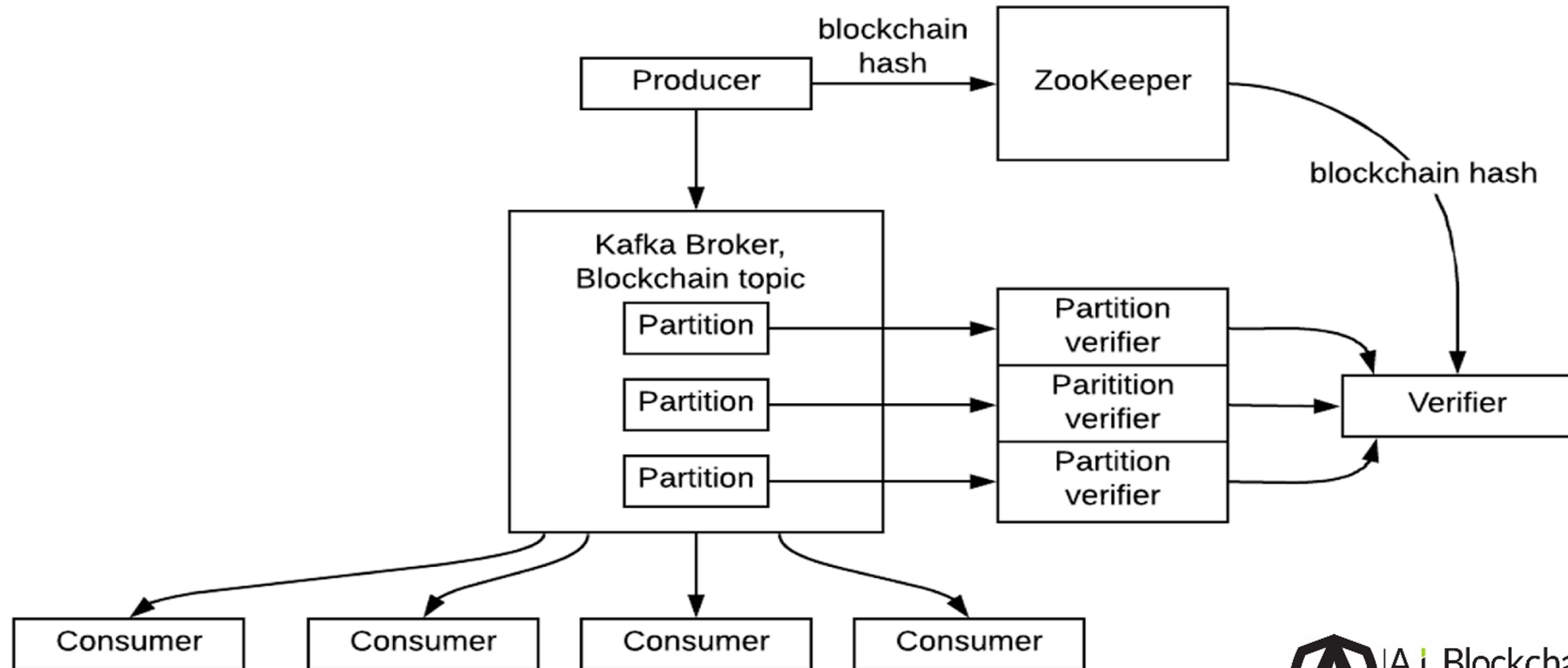
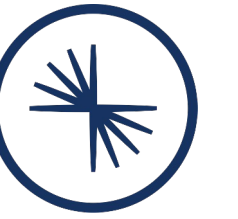


Is 'tamper-proof'
all you need?

Apache Kafka is an immutable append-only commit log

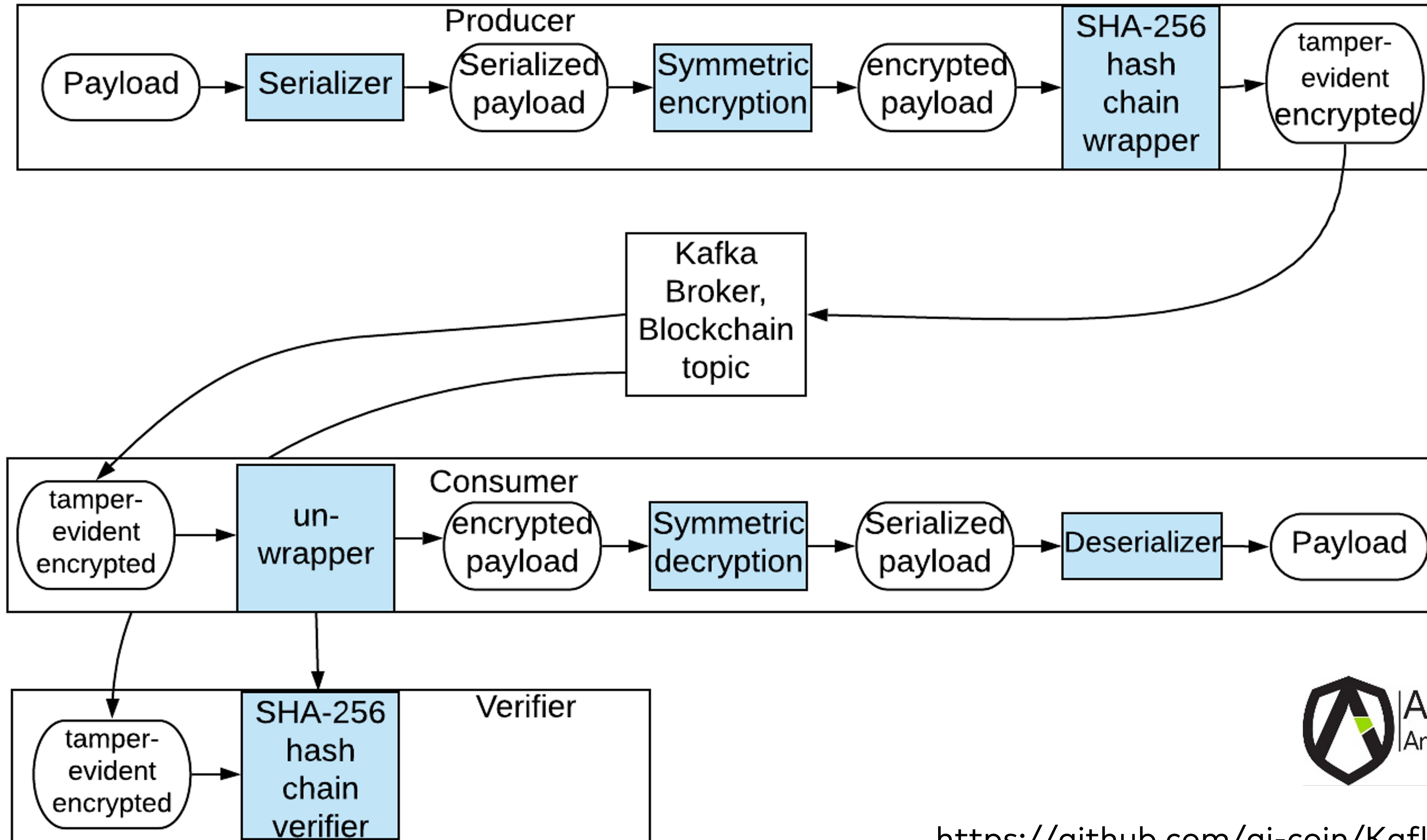
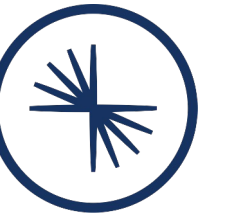


KafkaBlockchain – A library for tamper-proof Kafka streaming



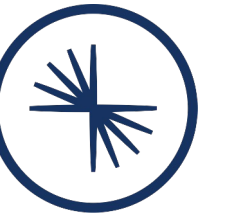
<https://github.com/ai-coin/KafkaBlockchain>

KafkaBlockchain – A library for tamper-proof Kafka streaming



<https://github.com/ai-coin/KafkaBlockchain>

What about long-term storage in Kafka?

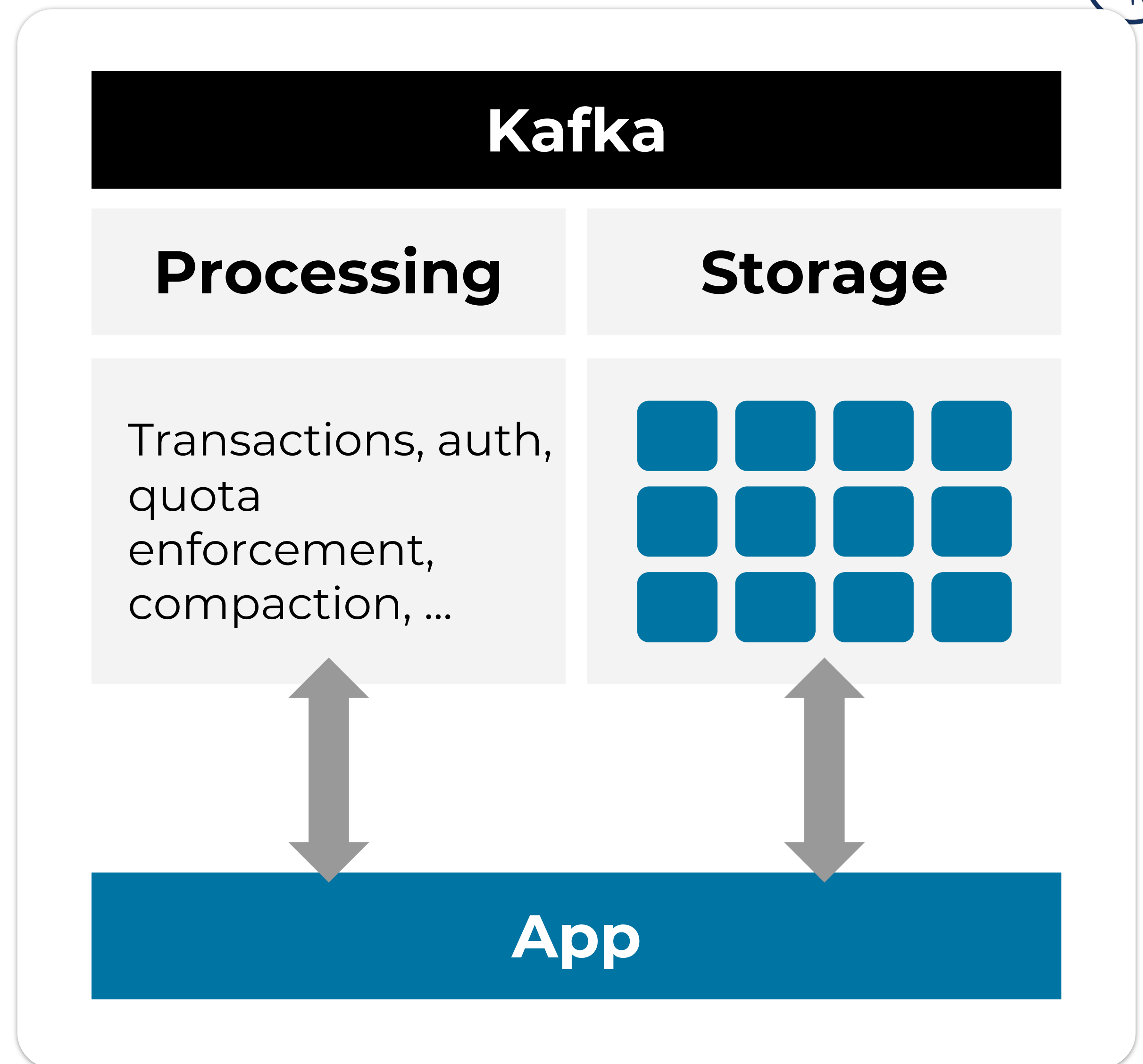


Today, Kafka works well for recent events, short horizon storage, and manual data balancing

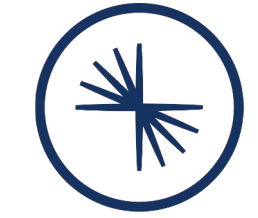
Kafka's present-day design offers **extraordinarily low messaging** latency by storing topic data on fast disks that are collocated with brokers. This is usually good.

But sometimes, you need to store a **huge amount of data for a long time.**

Blockchain is such a use case!



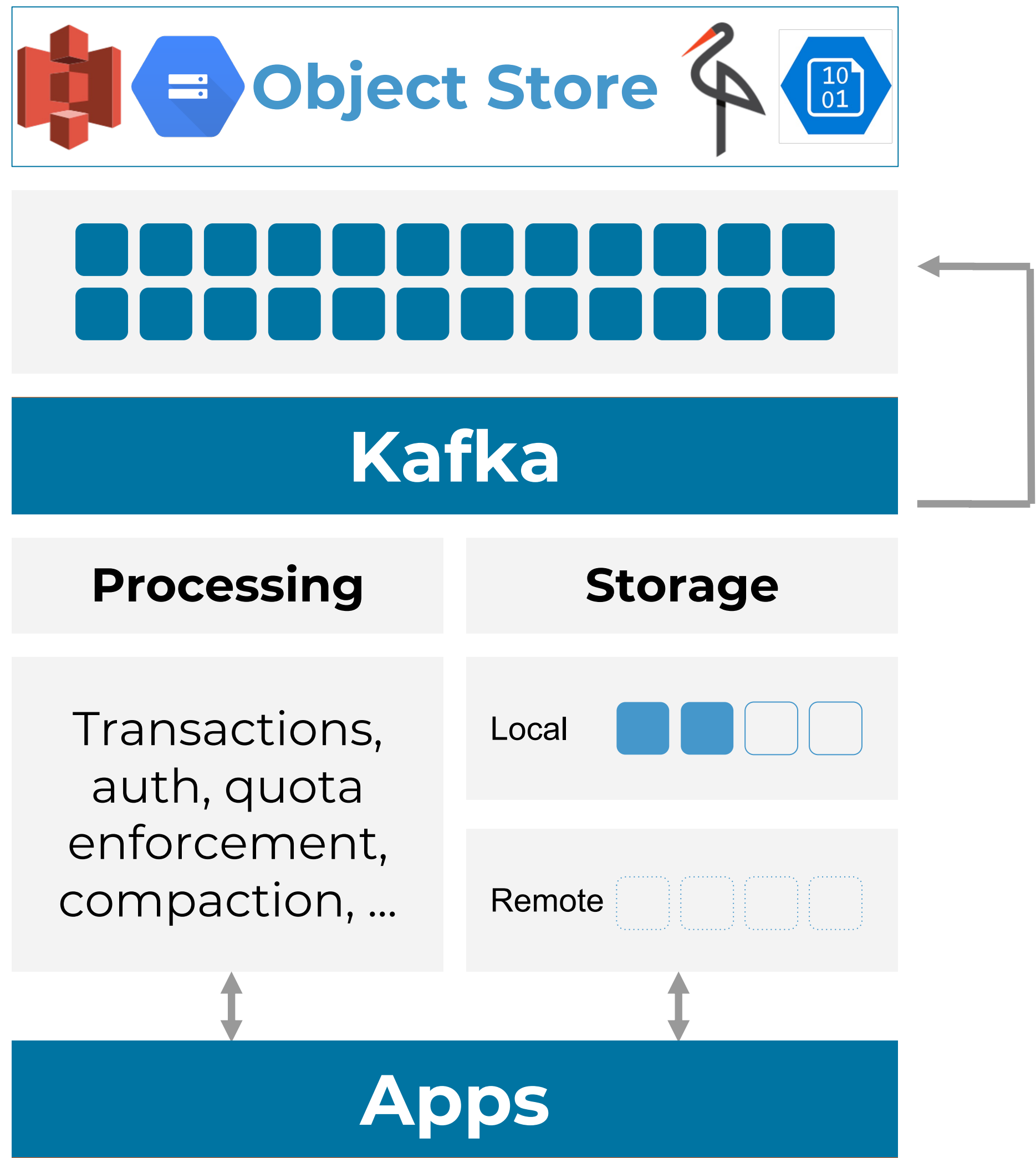
Confluent Tiered Storage for Kafka



Store Forever (Tamper-Proof)
Older data is offloaded to inexpensive object storage, permitting it to be consumed at any time. Using KafkaBlockchain, storage can be made tamper-proof and immutable

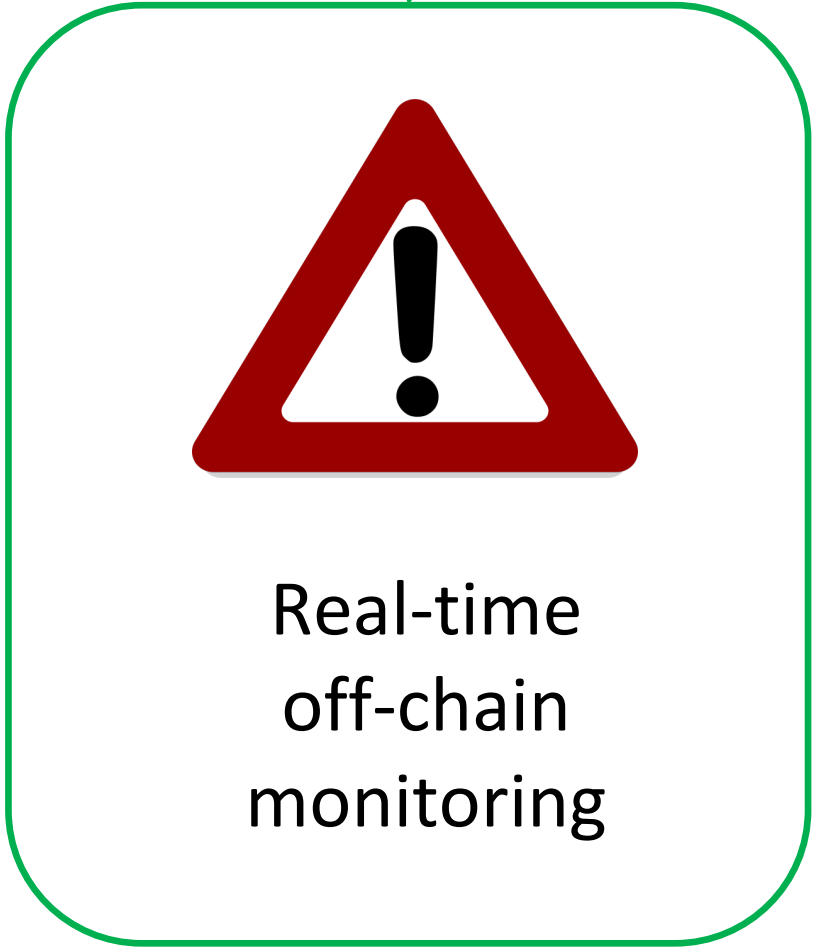
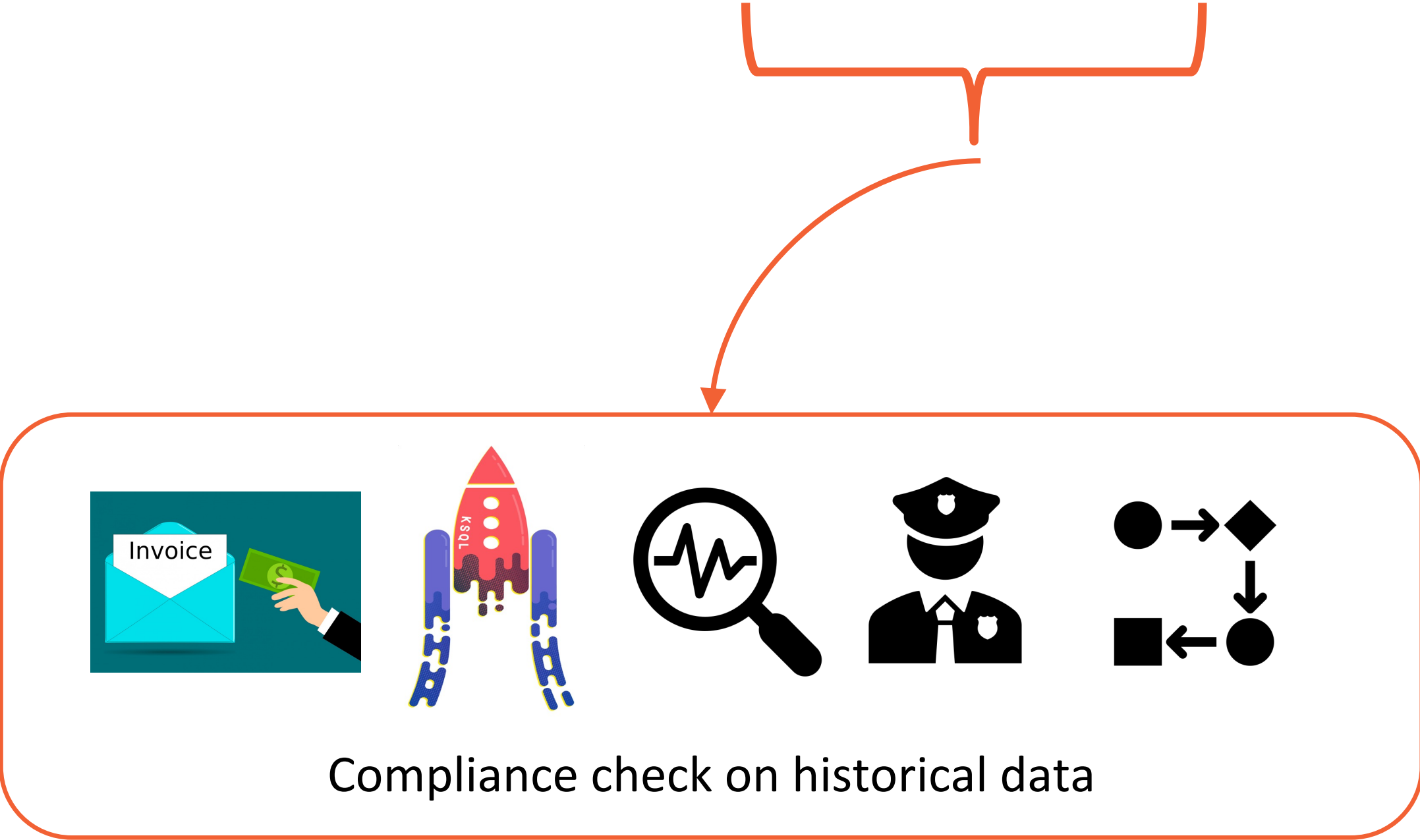
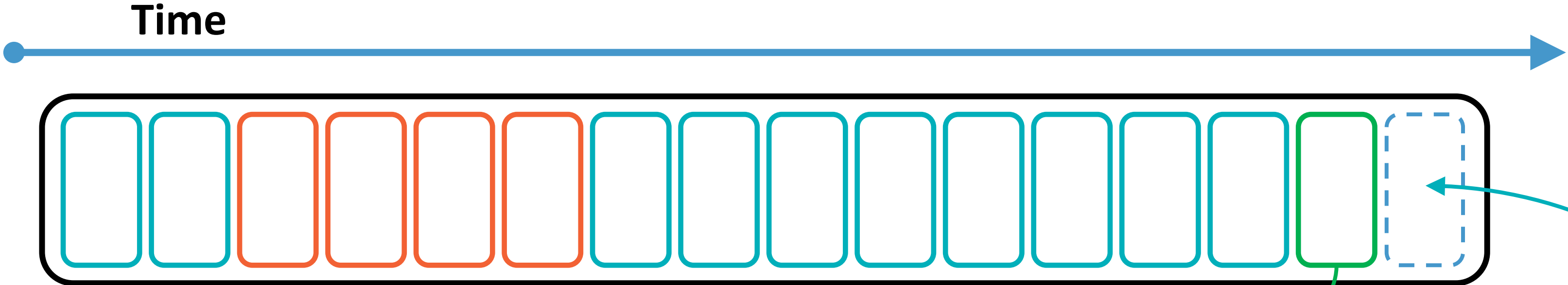
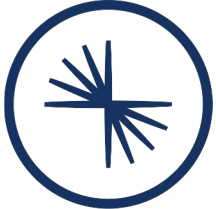
Save \$\$\$
Storage limitations, like capacity and duration, are effectively uncapped.

Instantaneously scale up and down
Your Kafka clusters will be able to automatically self-balance load and hence elastically scale



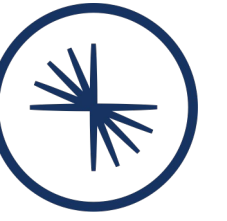
Secure, tamper-proof, encrypted off-chain data streaming

Real-time data processing and analytics of historical events with one scalable infrastructure

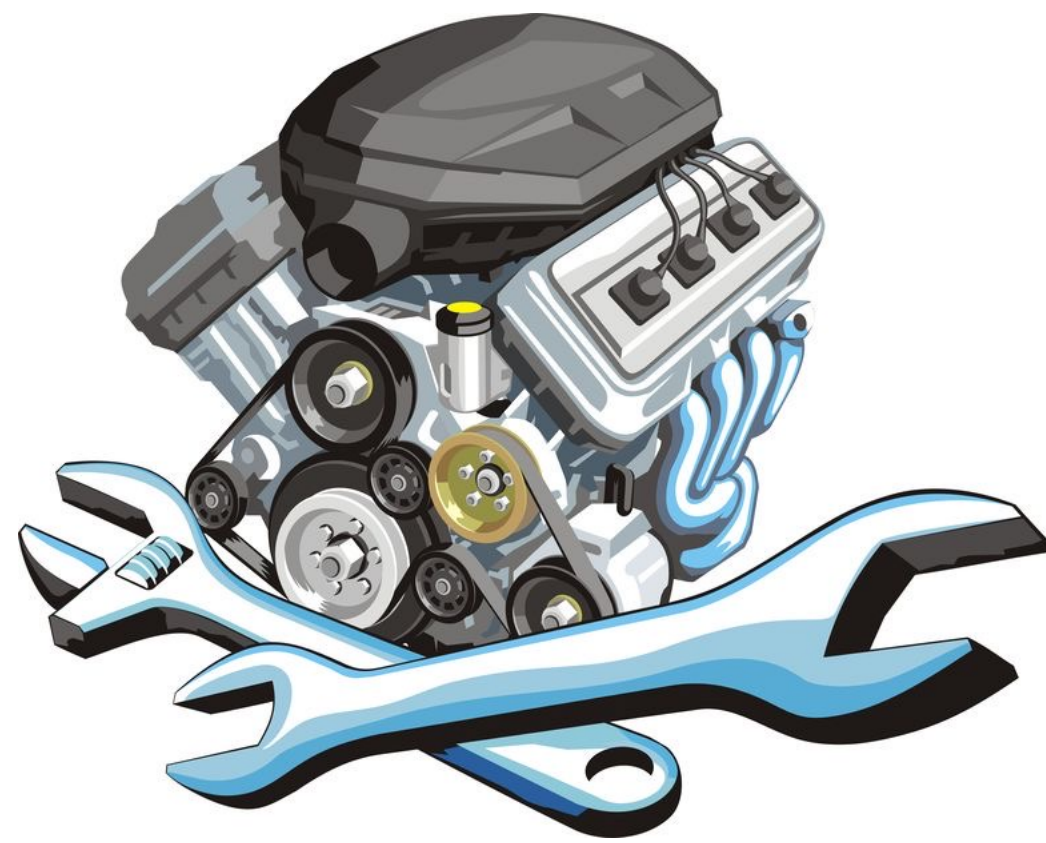


Why Confluent

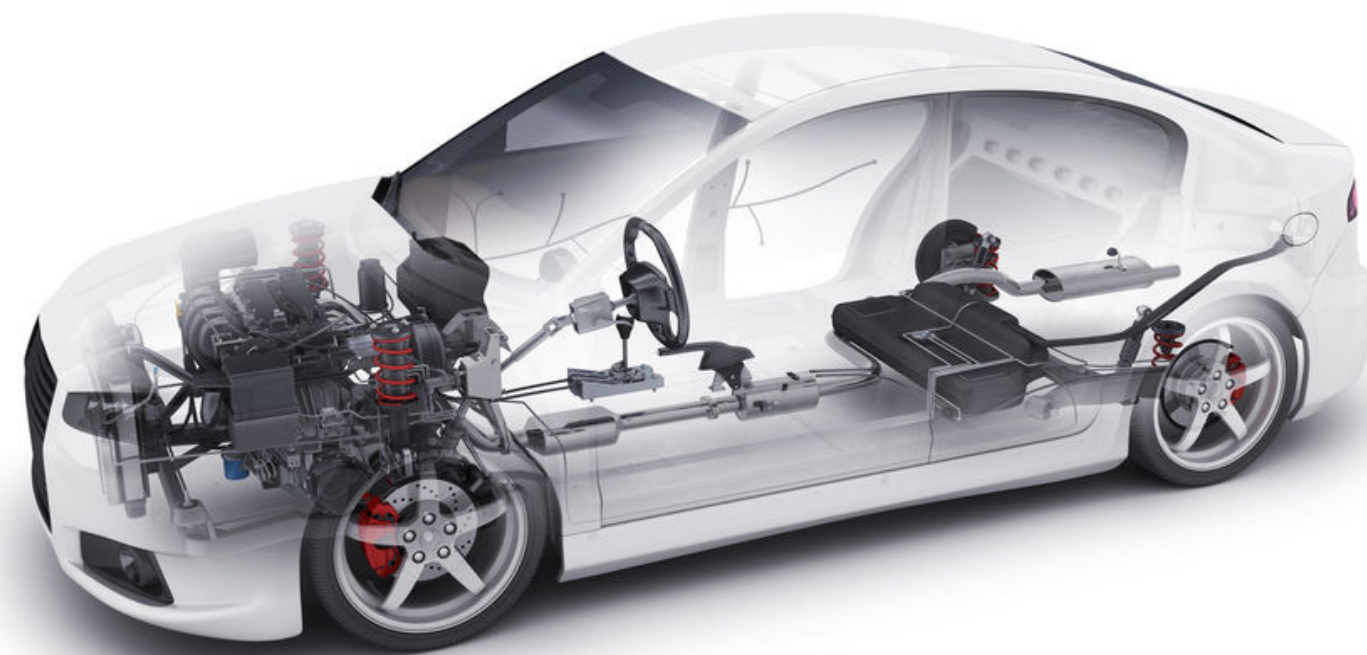
Confluent completes Apache Kafka. Cloud-native. Everywhere.



Car Engine

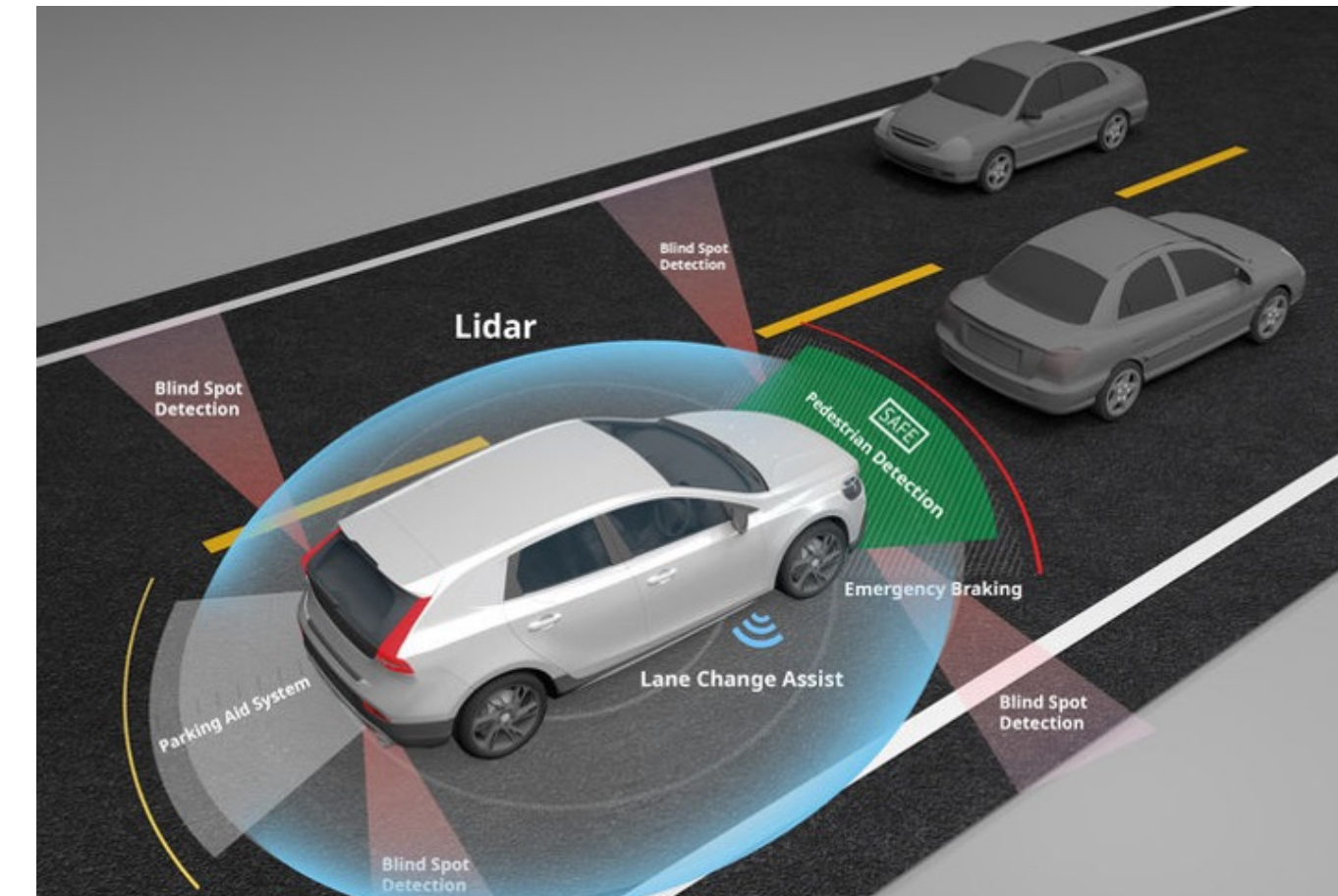


Car



CONFLUENT

Self-driving Car



Confluent Cloud





Questions? Feedback?
Let's connect!

Kai Waehner

Field CTO

kai.waehner@confluent.io
confluent.io
kai-waehner.de
@KaiWaehner
linkedin.com/in/kaiwaehner

