



CONFLUENT



Event Driven 101 & Demo

Rémi Forest, Solutions Engineers Manager SEMEA

Olivier Laplace, Solutions Engineer

Florent Ramière, Solutions Engineer

Eric Ortiz, BD and Partners Manager SEMEA

2009



ExxonMobil

Microsoft®

ICBC 

The logo for Industrial and Commercial Bank of China (ICBC), featuring the letters "ICBC" in a bold, black, sans-serif font next to a red circular emblem containing the number "25".

Walmart 

The logo for Walmart, featuring the word "Walmart" in a blue, bold, sans-serif font next to its signature yellow five-pointed starburst symbol.

2019

Data ... Data ... Data ... Data ... Data ...



 **Microsoft**

Alphabet

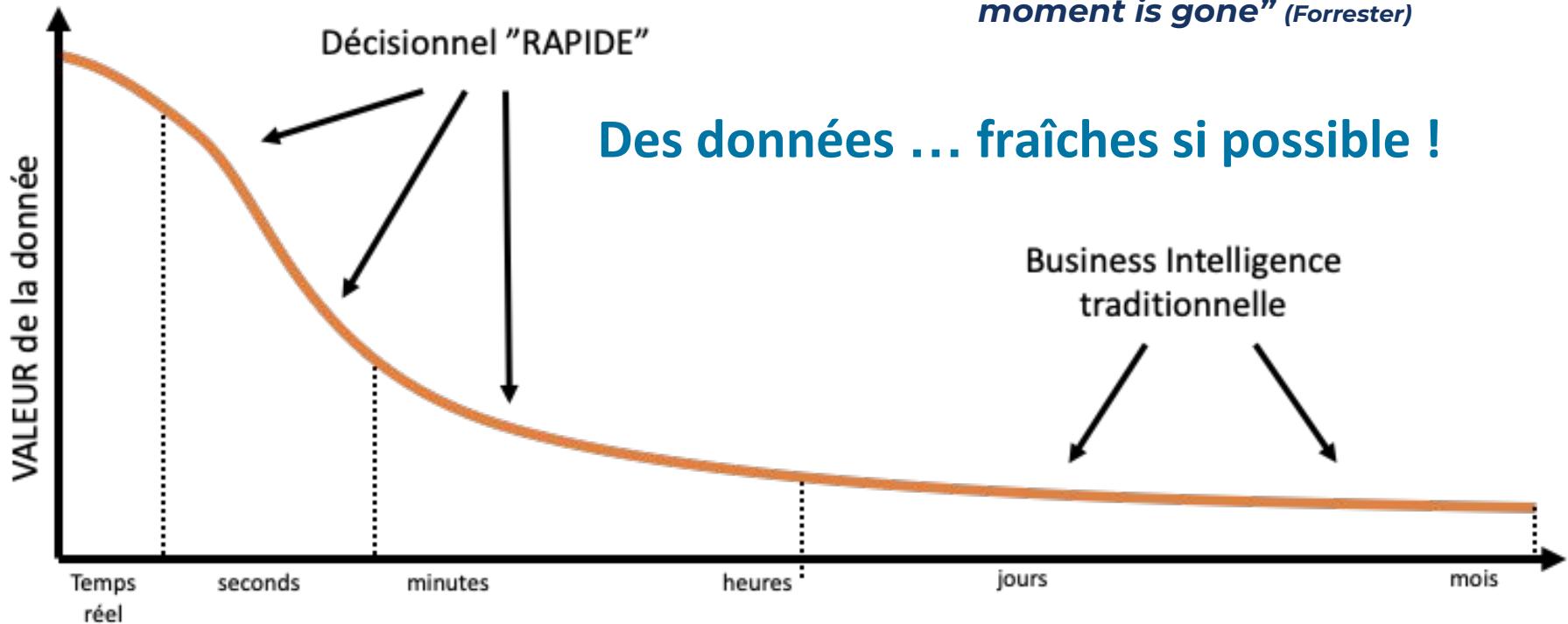
amazon

facebook



“Perishable Insights”

“Insights that can provide incredible value, but the value expires and evaporate once the moment is gone” (Forrester)





Un besoin d'immédiateté : Je veux ...

En tant que personne

- savoir quand mon Taxi/VTC va arriver (Transport: UBER, Lyft)
- être remboursé de mes frais médicaux à la sortie du praticien (Assurance: ALAN)
- avoir des recommandations en temps réel en fonction de mon activité WEB ou Media (Retail: Walmart, Media: Netflix)
- ...

En tant qu'entreprise

- prédire de possibles pannes (maintenance prédictive)
- suivre en temps réel mon stock et mes inventaires
- identifier rapidement des fraudes de paiement
- mettre à jour en temps réel le compte fidélité de mes clients
- appliquer du ML à des modèles dynamiques
- ...

“Quand souffle le vent du changement, certains construisent des murs, d’autres des moulins à vent”

Event Driven 101 :

Poser les fondations pour répondre au monde d’aujourd’hui et de demain





Agenda

01

Event

02

Event Driven

03

Event Driven Architecture

04

Event Streaming Platform

05

Attention Points

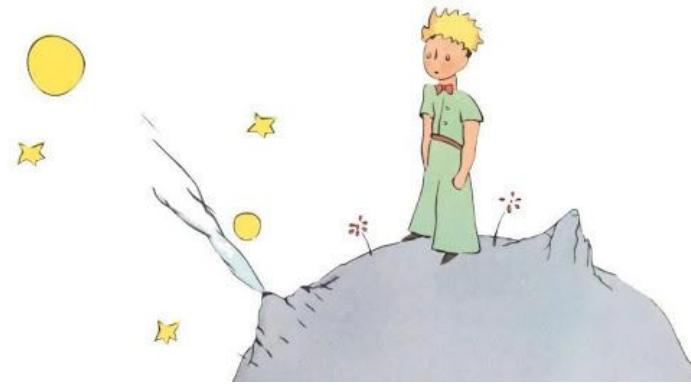
06

Demo CP & CC



01 - Event

What's an Event?



Dessine moi un Event ...

Dictionnaire (Larousse) :

“Evénement [EVENT] : *Tout ce qui se produit, arrive ou apparaît”*

versus

“Etat [STATE] : *Situation de quelque chose à un moment donné”*



Events vs State

Streams
record history

- 1. e4 e5
- 2. Nf3 Nc6
- 3. Bc4 Bc5
- 4. d3 Nf6
- 5. Nbd2



“The sequence of moves”

EVENTS

Tables
represent state



“The state of the board”

STATE

Events are at the Heart of Every Business

A business is defined through a series of events and its ability to respond to them



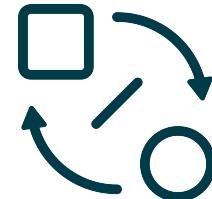
A **digital business** represents **events in data**



A Customer
Experience



A Sale



A Trade



An Invoice

Events are everywhere in every business...



- A good was sold
- A row was updated in a database table
- A wind turbine sensor measured 14 revolutions per minute
- An action occurred in a video game
- Temperature of an engine
- A click in a Web Application
- A chess move such as “White moved the e2 pawn to e4”
- A payment of \$200 was made by Frank to Sally on Nov. 24, 2019, at 17:11
- ...



02 - Event Driven

Becoming Software-Defined is a Competitive Requirement

“By 2020 **eventsourced, real-time** situational awareness will be a required characteristic for 80% of digital business solutions. And **80% of new business** ecosystems **will require support for event processing.**”

Gartner



Every Industry is Moving from Batch/Manual to Software-Defined/Event Driven

	Software-using	Software-defined
Auto / Transport	Spreadsheet-driven driver schedule	Real-time ETA
Banking	Nightly credit-card fraud checks	Real-time credit card fraud prevention
Retail	Batch inventory updates	Real-time inventory management
Healthcare	Batch claims processing	Real-time claims processing
Oil and Gas	Batch analytics	Real-time analytics
Manufacturing	Scheduled equipment maintenance	Automated, predictive maintenance
Defense	Reactive cyber-security forensics	Automated SIEM and Anomaly Detection



JPMORGAN CHASE & Co.

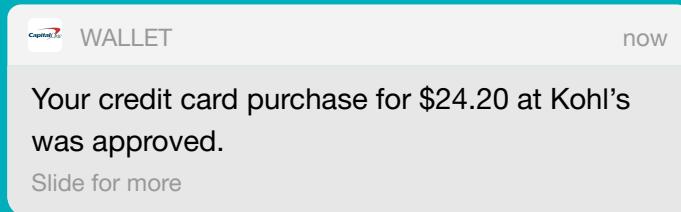


Morgan Stanley



U.S. Defense Agencies ¹⁴

Real time with historical context: **Contextual**

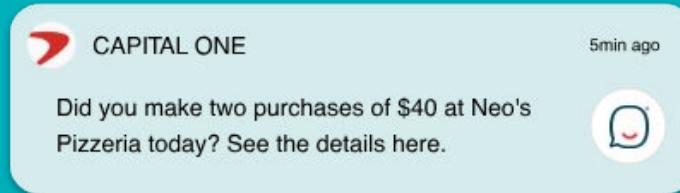


Event-Driven App (Purchase Notification)

Only Real-Time Events

Messaging Queues and Event Streaming Platforms can do this

Noisy, low value communication



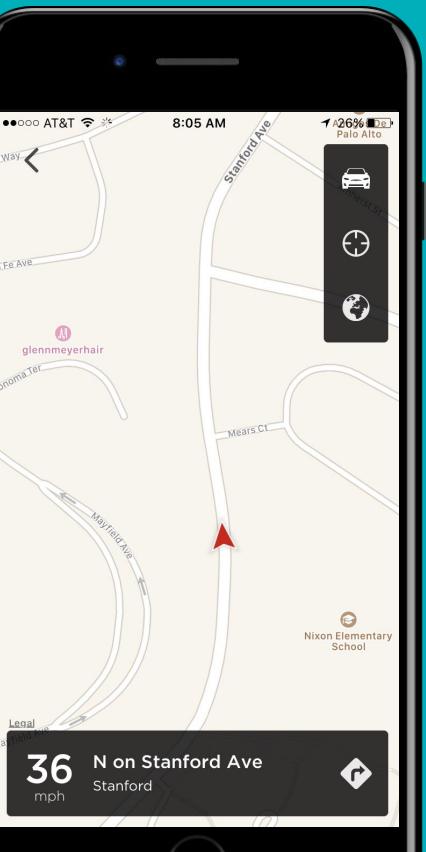
Contextual Event-Driven App (Unusual Activity Alert)

Real-Time combined with stored data

Only Event Streaming Platforms can do this

Intelligent, high-value communication

Why add the word **Contextual**?



Event-Driven App (Location Tracking)

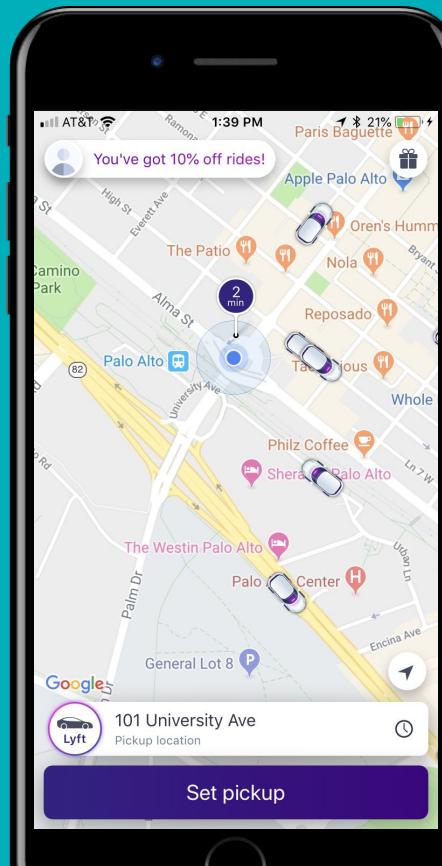
Only Real-Time Events
Messaging Queues and
Event Streaming
Platforms can do this

Where is my
driver?

Contextual Event-Driven App (Estimated Time Arrival)

Real-Time combined with
stored data
Only Event Streaming
Platforms can do this

When will my
driver get here?





***“Streams capture events,
series of events capture behaviors,
they tell a story”***



03 - Event Driven Architecture

ETL/Data Integration

-

Batch

Expensive

Time Consuming

+

High Throughput

Durable

Persistent

Maintains Order

Messaging

+

Fast (Low Latency)

-

Difficult to Scale

No Persistence

Data Loss

No Replay

ETL/Data Integration

-

Batch

Expensive

Time Consuming

+

High Throughput

Durable

Persistent

Maintains Order

Stored records

Messaging

+

Fast (Low Latency)

-

Difficult to Scale

No Persistence

Data Loss

No Replay

Transient Messages

ETL/Data Integration

Messaging

Batch Processing
Expensive
Time Consuming

+
Durable
Maintains Order

**Both of these are a complete mismatch
to how your business works.**

Stored records

Transient Messages

ETL/Data Integration

-

Batch

Expensive

Time Consuming

Event Streaming Paradigm



High Throughput



Fast (Low Latency)

Durable

Persistent

Maintains Order

Messaging

-

Difficult to Scale

No Persistence

Data Loss

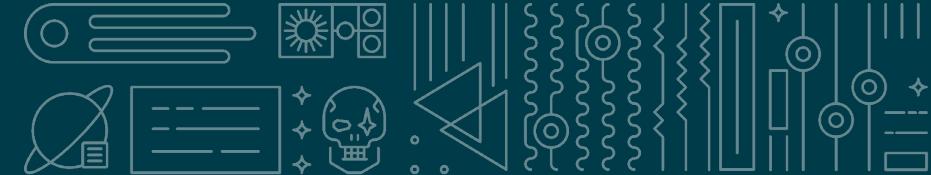
No Replay

Stored records

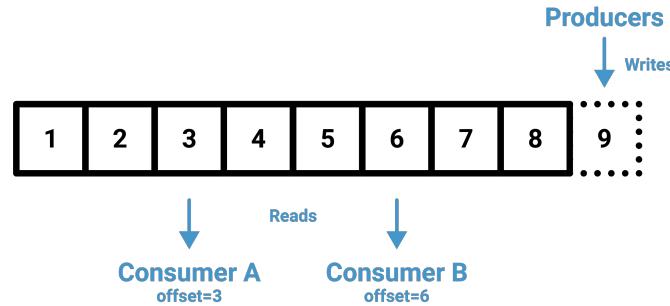
Transient Messages

Event Streaming Paradigm

To rethink data as not stored records or transient messages, but instead as a continually updating stream of events

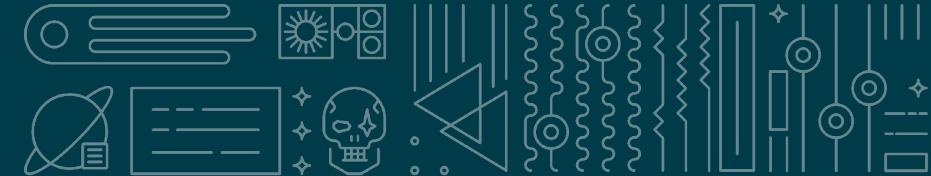


Event Streaming Paradigm



Event Streaming allows

**Contextual Event Driven
Applications**



Software-using: How we managed Data



BUILT FOR REAL-TIME EVENTS	SCALABLE FOR ALL DATA	PERSISTENT & DURABLE	CAPABLE OF ENRICHMENT
----------------------------	-----------------------	----------------------	-----------------------

Databases	✗	Good for transactional applications	✓
-----------	---	-------------------------------------	---

Messaging	✓	Good for ultra low-latency, fire-and-forget use cases	✗
-----------	---	---	---

ETL	✗	Good for batch data integration	✓
-----	---	---------------------------------	---

Data Warehouse	✗	Good for historical analytics and reporting	✓
----------------	---	---	---

Event Streaming	✓	The Essential Data Platform for Becoming Software-Defined	
-----------------	---	---	--

(Scalable Messaging + Real-Time Data Integration + Stream Processing)

Data Platform Requirements for Becoming Software-Defined



Software
-using

Built for
Historical Data

Scalable for
Transactional Data

Transient

Raw data

Software
-Defined

① Built for Real-
Time Events

- State vs. Change
- Historical analysis vs. real-time operations

② Scalable for
ALL data

- Non-transactional data is 10x transactional data
- IoT, Logs, Security events...

③ Persistent +
Durable

- Mission critical apps require zero data loss
- Mission critical systems require replay

④ Enriched
data

- Stream Processing (SQL on RT events)
- Context & situational awareness (ex. ETA)

Key principles of an EDA



The event-driven architecture is made up of **highly decoupled**, single-purpose event processing components that **asynchronously receive** and **process** events.



04 - Event Streaming Platform

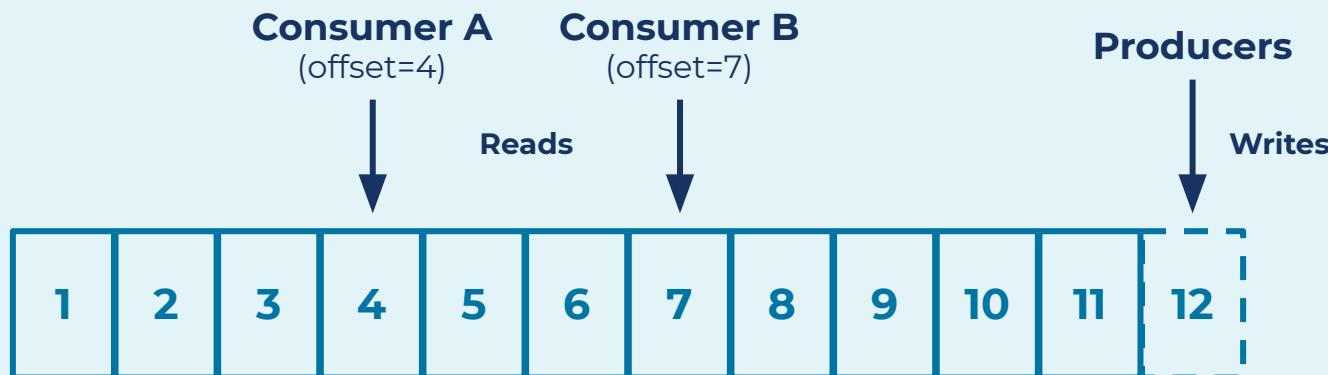
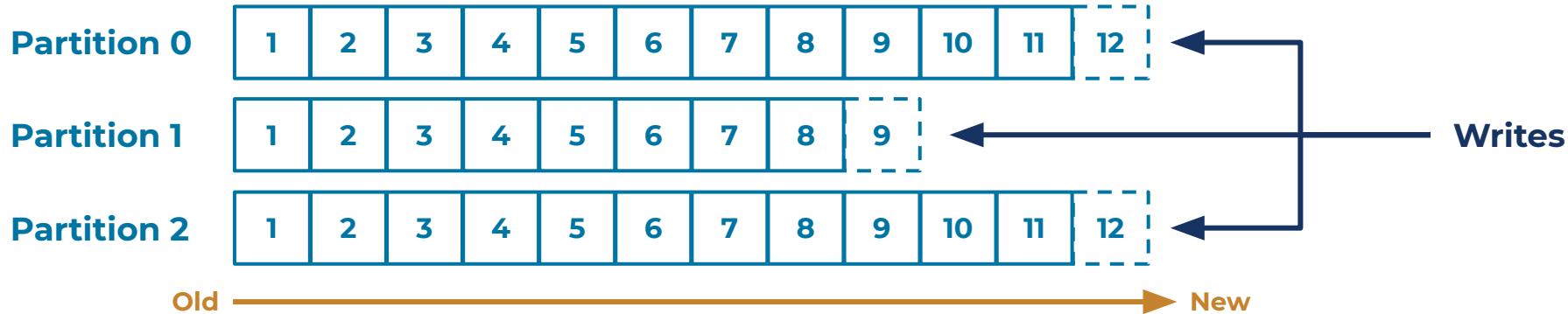
What Confluent Event Streaming Platform brings



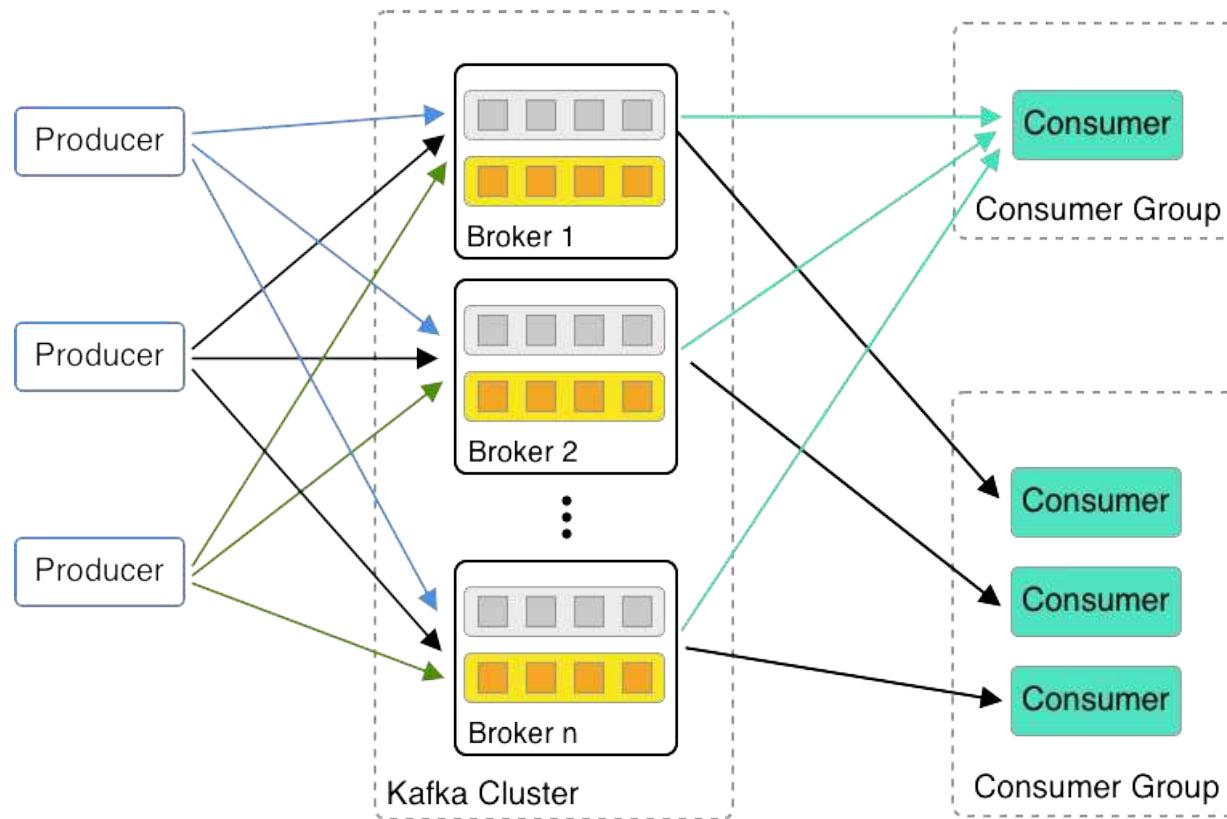
- It lets you publish and subscribe to events
- It lets you store events for as long as you want
- It lets you process and analyze events

in a scalable, fault-tolerant, and reliable manner!

Anatomy of a Kafka Topic



Scalable Data Pipeline: Producer & Consumer



Stream Processing: Push vs Pull queries



Processing Layer
(ksqlDB, Kafka Streams)

Point-in-time lookups of static information

Pull queries

Table

plus aggregation

alice	Rome
bob	Sydney

Subscribe to changing query results in real-time

Push queries

Stream

plus schema
(SerDes)



Storage Layer
(Brokers)

Topic



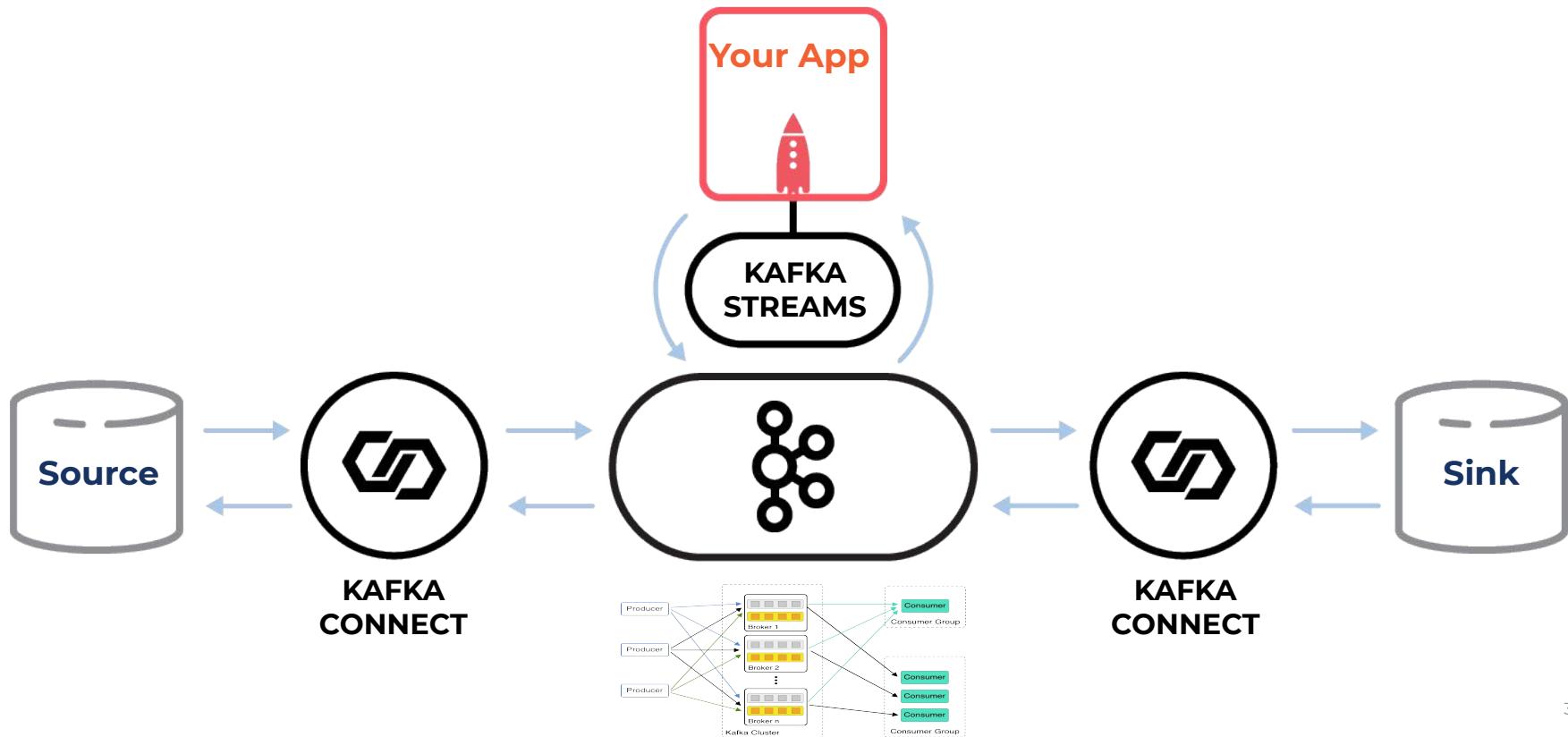
Instantly Connect Popular Data Sources & Sinks



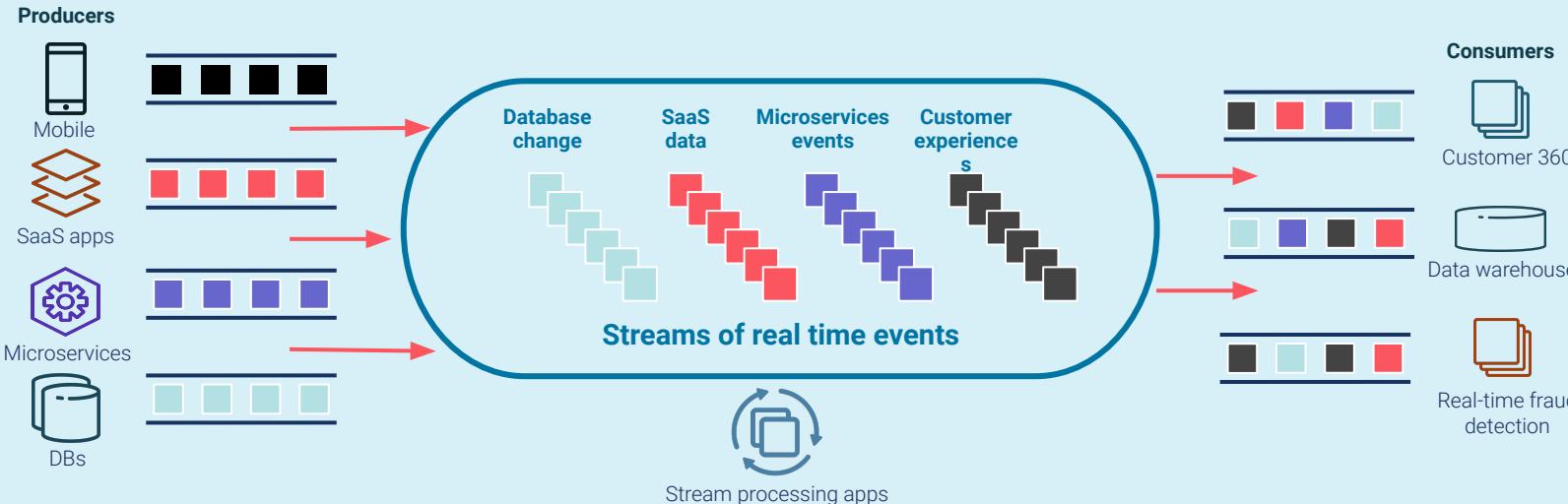
80+ Confluent Supported

20+ Partner Supported, Confluent Verified

Kafka + Kafka Connect + Kafka Streams



A Streaming Platform is the Underpinning of an Event Driven Architecture



Ubiquitous connectivity

Globally scalable platform for all event producers and consumers

Immediate data access

Data accessible to all consumers in real time

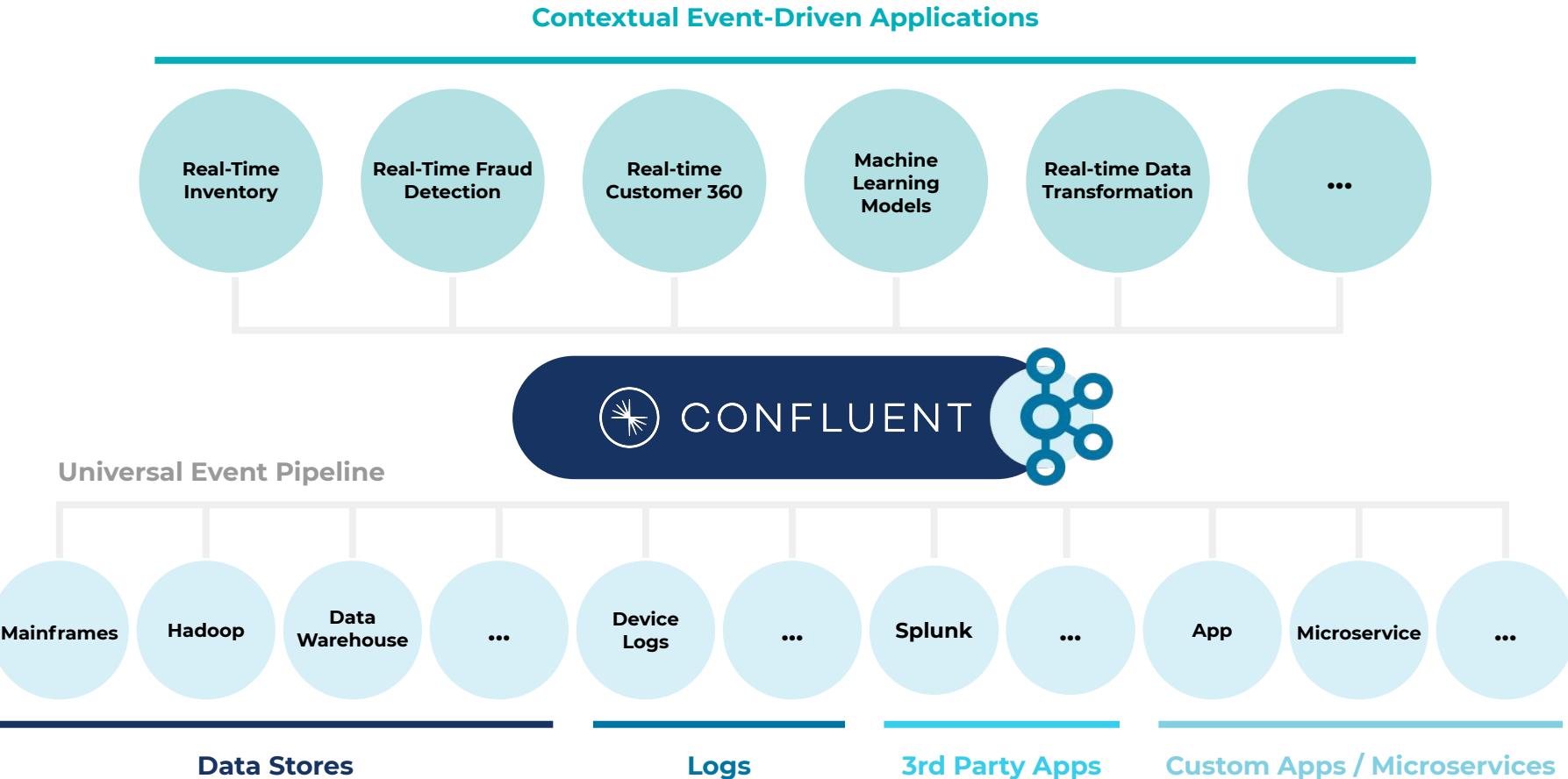
Single system of record

Persistent storage to enable reprocessing of past events

Continuous queries

Stream processing capabilities for in-line data transformation

Confluent is this platform

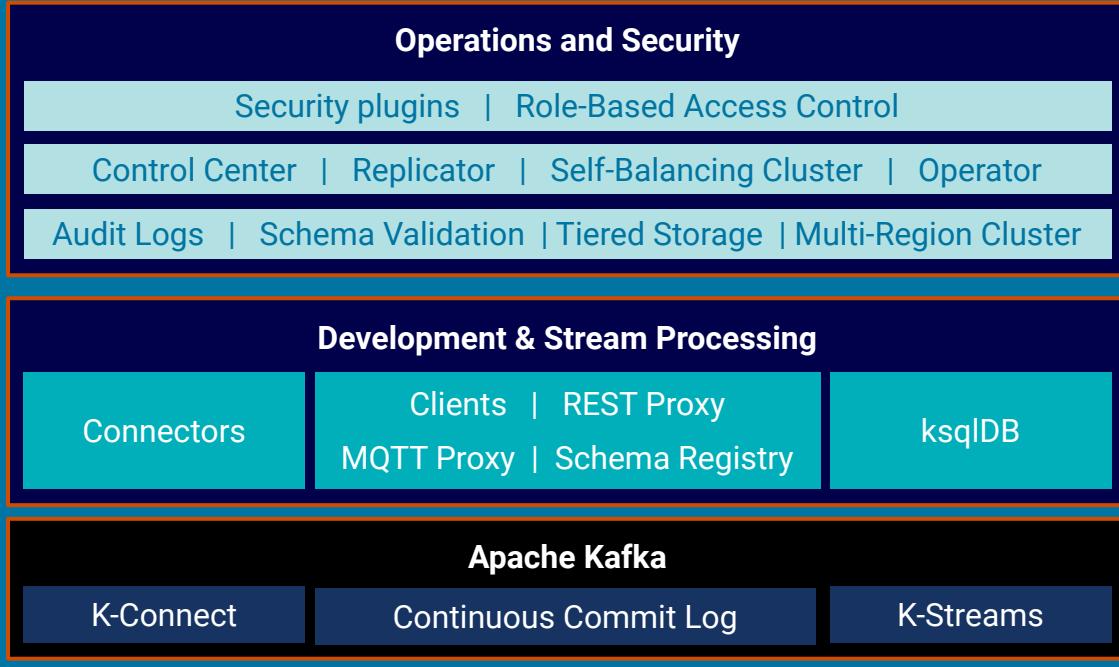




Confluent combines **Event Streaming** (ingestion is a component), **Stream Processing** (ability for microservices and apps to use and manipulate the streaming data), and an **Integration Layer** (by enabling multiple systems to connect input and output). **No other single platform does this.**

It **fundamentally** changes how organizations work with data.

Confluent Platform 6.0



**Mission-critical
Reliability**

Complete

**Event Streaming
Platform**

Freedom of Choice

Self-Managed Software

Fully-Managed Service

Datacenter

Public Cloud

Confluent Cloud



05 - Attention Points

What should you ask?

Beyond the need for Real Time, questions related to use cases when considering an **Event Driven Architecture**

- Is it needed to **persist the data? How long?** Is it needed to **replay**?
- Is it trusted? With which semantic? meaning transactional, **exactly once, at most or at least once**? Will it scale?
- Is **stateless processing**, such as filtering, projection, cleaning or enrichment required? Is **stateful processing**, such as aggregations or stateful sequence processing required?
- Is a **materialized view** against the stream **required?** How many **Tx per second** are required for a windowed view?
- What about **volume, latency, scalability?**
- Does it **support error handling**, such as error flows and dead letter queue?
- Does it need to **integrate** with other solutions or applications?
- Does it **react** and **drive intelligence** from the state collected from a stream (**stream processing**)?



Is it (really) easy?

What are the potential technical or organizational difficulties when building an Event Driven Architecture?

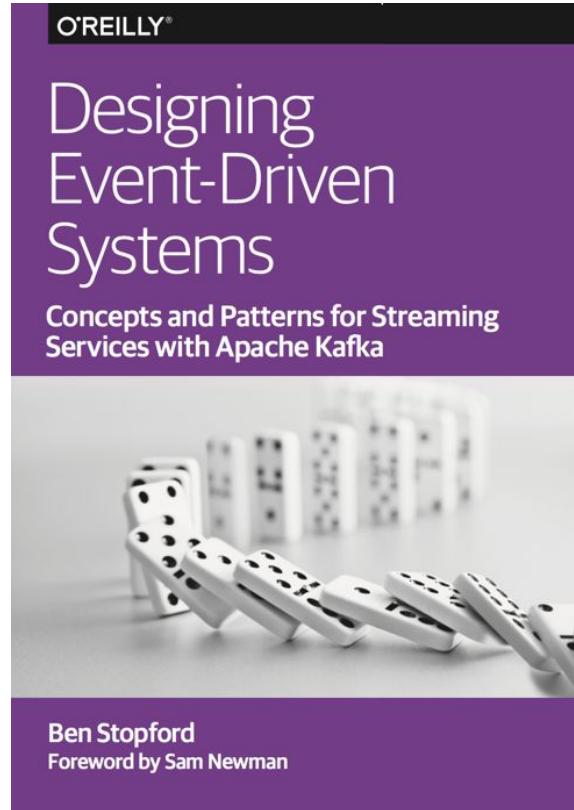
- Asynchronous and Distributed ⇒ various distributed architecture issues
- Granularity ⇒ Which Event can / can't run independently?
- Lack of Atomic transactions ⇒ How to maintain a Tx unit of work for a global business process?
- Data Governance ⇒ Each Event must have a data format and a (versioned) contract right from the start
- Event Stream vs Stream of States ⇒ Stream Processing is important
- Tests ⇒ Independently, globally, when all is fine... and when it's not
- Decoupled but... ⇒ teams must communicate
- It's a multi domain area ⇒ Data Integration, Security, Monitoring, Topic & Schema Design, Data Governance, Error Handling, RPO/RTO, Multi DC, Throughput, Latency, Durability, Availability...



Need one book?



<http://bit.ly/designing-event-driven-systems>





06 - Demo CP & CC

Two Ways to Deploy Confluent



SELF-MANAGED SOFTWARE



Confluent Platform

The Enterprise Distribution of Apache Kafka

In the datacenter



FULLY-MANAGED SOFTWARE



Confluent Cloud

Apache Kafka Re-Engineered for the Cloud

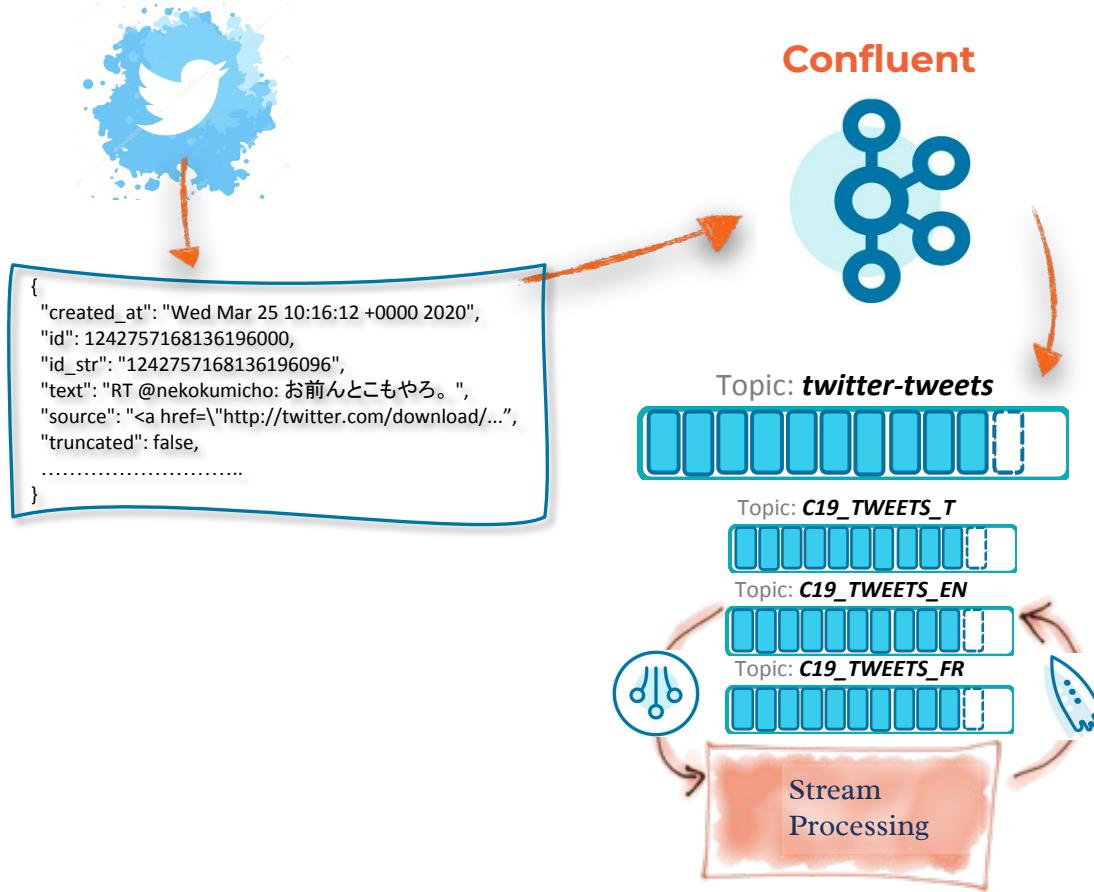
In the cloud





Confluent Platform

Démo : Topologie

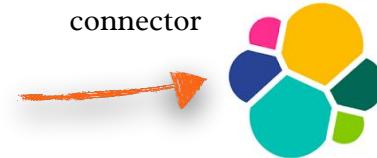


English
messages



```
{  
  "USER_NAME": "JamesBond@spy.com",  
  "USER_LOCATION": "Londres",  
  "TEXT": "RT @Spy: James Bond fight Convid19",  
  "SOURCE": "<a href=\"http://twitter.com/download/...\"",  
  "LANG": "en",  
  "CREATED_AT": "Wed Mar 25 10:17:42 +0000 2020"  
}
```

Elastic (sink)
connector



File (sink)
connector



French
messages

```
{  
  "USER_NAME": "Claire Saphore@CAAquitaine",  
  "USER_LOCATION": "Bordeaux",  
  "TEXT": "RT @Amalmezat: Le @CAAquitaine se mobilise ...",  
  "SOURCE": "<a href=\"http://twitter.com/download/...\"",  
  "LANG": "fr",  
  "CREATED_AT": "Wed Mar 25 10:17:42 +0000 2020"  
}
```



Démo Confluent Platform : étapes

1. Confluent Platform 6.0 (installation basique)
2. Naviguer dans la console Control Center
 - a. Monitoring, Alerting
 - b. Tiered Storage, Self Data Balancer
 - c. Proactive SupportAdministration / Supervision / Monitoring
3. Publier des tweets dans 1 Topic Confluent (twitter-tweets) Producer API / Ingestion
4. "Jouer" un peu avec les données (Stream Processing - ksqlDB)
 - a. Créer une Streams C19_TWEETS ⇒ Push Queries (1)
 - b. Créer une Table pour compter les messages FR (4-2)
 - c. Requêter la Tables - Pull Query ⇒ Pull Queries (5)
 - d. Créer le Topic C19_TWEETS_T (2)
 - e. Créer le Topic C19_TWEETS_FR (10)
 - f. Créer le Topic C19_TWEETS_EN (11)Stream Processing : Filtre / Agrégation / Transformation
5. Publier les données traitées dans des fichiers - Connect / Connector File
 - a. Créer le Connecteur fichier (12)
 - b. Visualiser le contenu du fichier (13)Intégration : Connecteur Fichier (sortie)
6. Publier les données traitées dans ELK - Connect / Connector ELK
 - a. Créer le Connecteur fichier (14)
 - b. Valider le contenu avec KibanaIntégration : Connecteur Elastic (sortie)



Démo Confluent Platform : étapes

1. Confluent Platform 6.0 (installation basique)
2. Naviguer dans la console **Control Center**
 - a. Monitoring, Alerting
 - b. Tiered Storage, Self Data Balancer
 - c. Proactive Support
3. Publier des tweets dans 1 Topic Confluent (twitter-tweets) ⇒ Producer API
4. "Jouer" un peu avec les données (Stream Processing - **ksqldb**)
 - a. Créer une Streams C19_TWEETS ⇒ Push Queries (1)
 - b. Créer une Table pour compter les messages FR (4-2)
 - c. Requêter la Tables - Pull Query ⇒ Pull Queries (5)
 - d. Créer le Topic C19_TWEETS_T (2)
 - e. Créer le Topic C19_TWEETS_FR (10)
 - f. Créer le Topic C19_TWEETS_EN (11)
5. Publier les données traitées dans des fichiers - **Connect / Connector File**
 - a. Créer le Connecteur fichier (12)
 - b. Visualiser le contenu du fichier (13)
6. Publier les données traitées dans ELK - **Connect / Connector ELK**
 - a. Créer le Connecteur fichier (14)
 - b. Valider le contenu avec Kibana



Confluent Cloud

Confluent Cloud : 3 Options



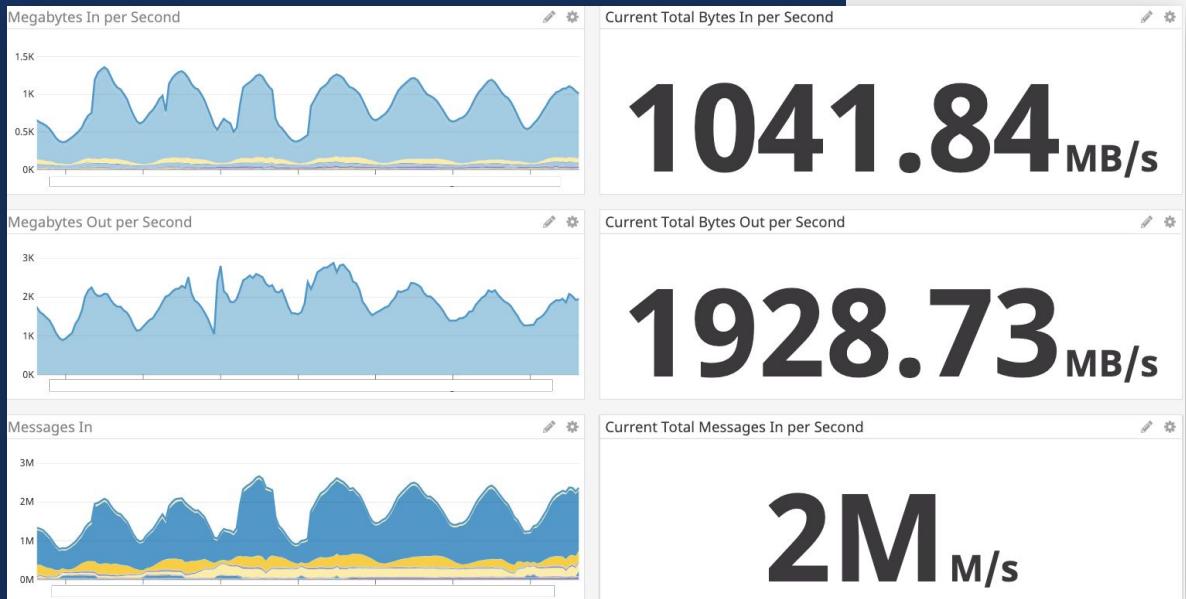
	Multi-tenant	Multi-tenant	Dedicated
	Basic Get started with no minimums	Standard Production-ready for most applications	Dedicated Customizable for any application
Sizing	No sizing required Stream up to 100MBps Store up to 5TB	No sizing required Stream up to 100MBps Store up to 5TB	Limits based on provisioned capacity (CKUs)
Replication options	Single AZ	Single & Multi AZ	Single & Multi AZ
SLAs	-	99.95%	99.95%
Private networking options	-	-	VPC/VNet Peering AWS Transit Gateway
Ideal for	Prototyping & early development	Most production use cases	Mission-critical applications at any scale

Leverage any cluster type across your organization

Confluent Cloud | High-performance at GBps scale



- **Achieve sub-25 ms** latencies at massive scale
- **Unlimited** throughput and fanout
- **Infinite** retention





Démo : Démarrez avec Confluent Cloud en 5'

1. <https://status.confluent.cloud/> : Très haut niveau de SLA
2. Créer une instance Confluent Cloud : <http://confluent.cloud/login>
3. Générer les Clés (API) d'accès (mise à jour dans fichier jaas)
4. Montrer l'intégration avec les autres composants
5. Créer un topic : twitter-tweets
6. Dans le producteur Java : Changer le bootstrap du cluster
7. Publier les tweets dans twitter-tweets



Merci !

olaplace@confluent.io (+33 6 58 19 43 96)

<https://developer.confluent.io/>

<https://www.confluent.io/blog/>

<https://docs.confluent.io/current/>