

# *The Changing Face of ETL*

## *Event-Driven Architectures for Data Engineers*

@rmoff





























It used to be so simple





# More Sources



A high-angle, nighttime photograph of a complex multi-level highway interchange. The concrete overpasses and ramps are illuminated by streetlights, creating a series of curved, overlapping lines. Several cars with their headlights on are visible traveling along the different levels of the highway. In the background, the silhouettes of city buildings and a tall, white industrial tower are visible against the dark sky. The overall scene conveys a sense of constant movement and urban infrastructure.

# More Targets



The background is an abstract image featuring a dense, repeating pattern of hexagons. The pattern is composed of light gray hexagons with dark gray outlines, creating a mesh-like texture. The pattern is slightly distorted, giving it a three-dimensional, undulating appearance. The text "More Data" is centered over this pattern in a white, sans-serif font, enclosed within a semi-transparent gray rectangular box.

More Data





# Batches and Buckets





Applications  
**Respond**

→ an order was placed!

Analytics

**Tell Us What  
Happened**

→ how many orders  
were placed













# Events



An event is both:

- ★ Notification

- ★ State transfer



# A Customer Experience





# A Sensor Reading





# Events

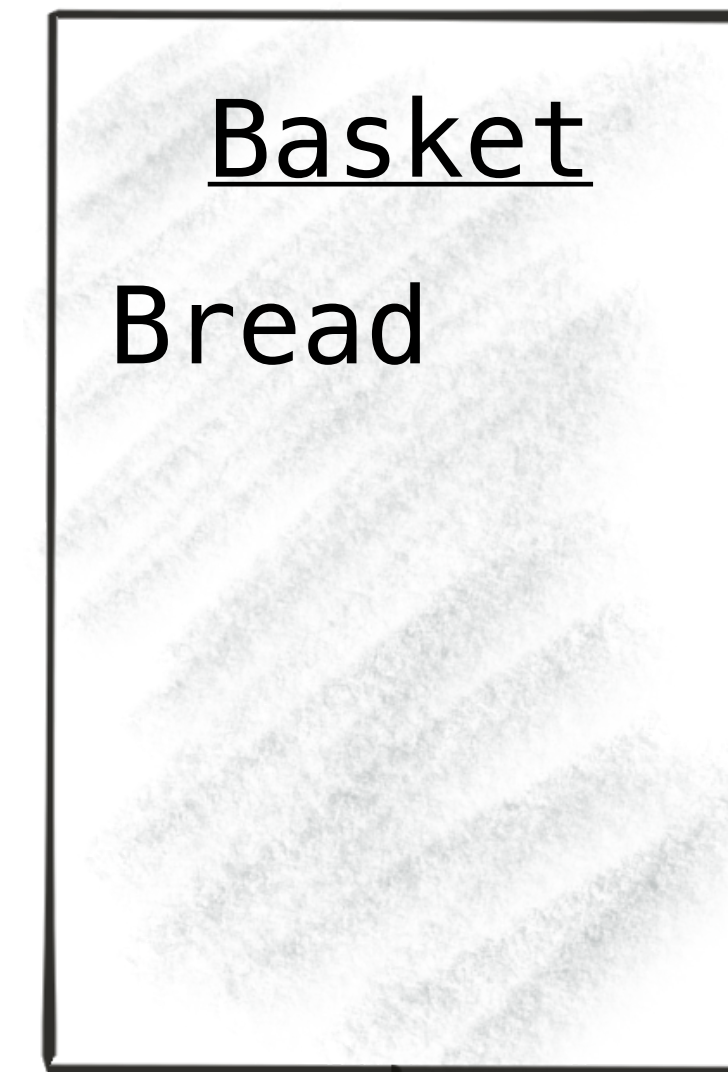
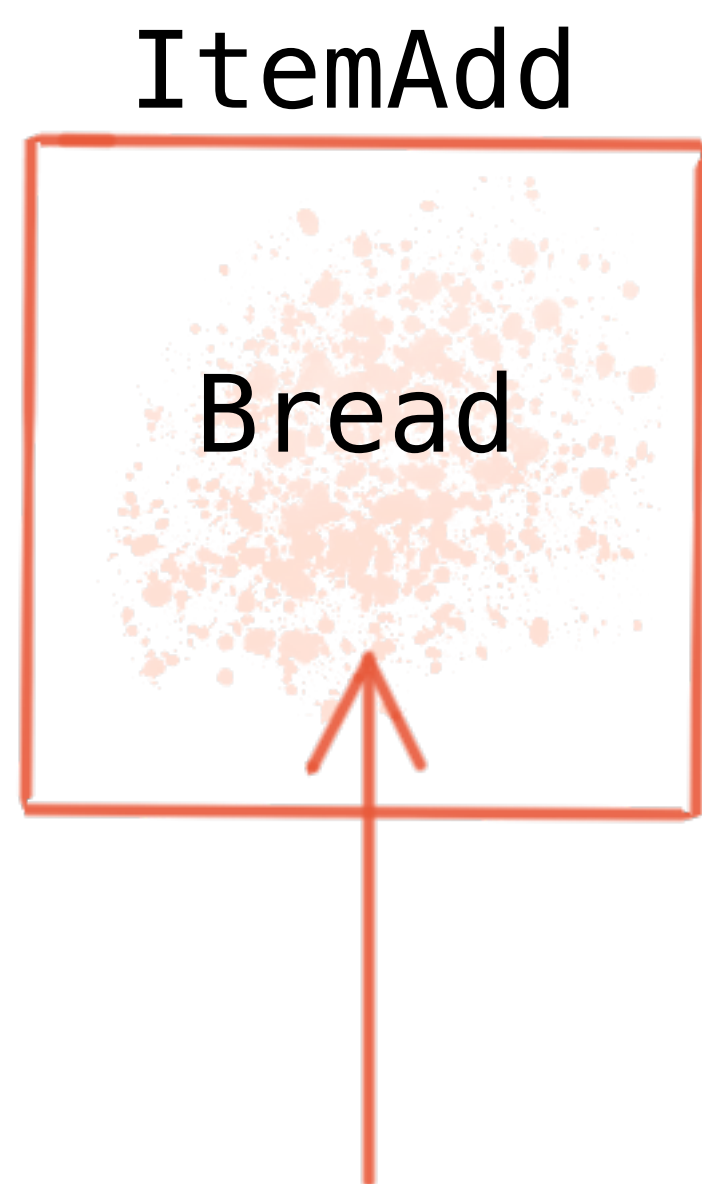
Basket

Bread

Canned  
Spaghetti

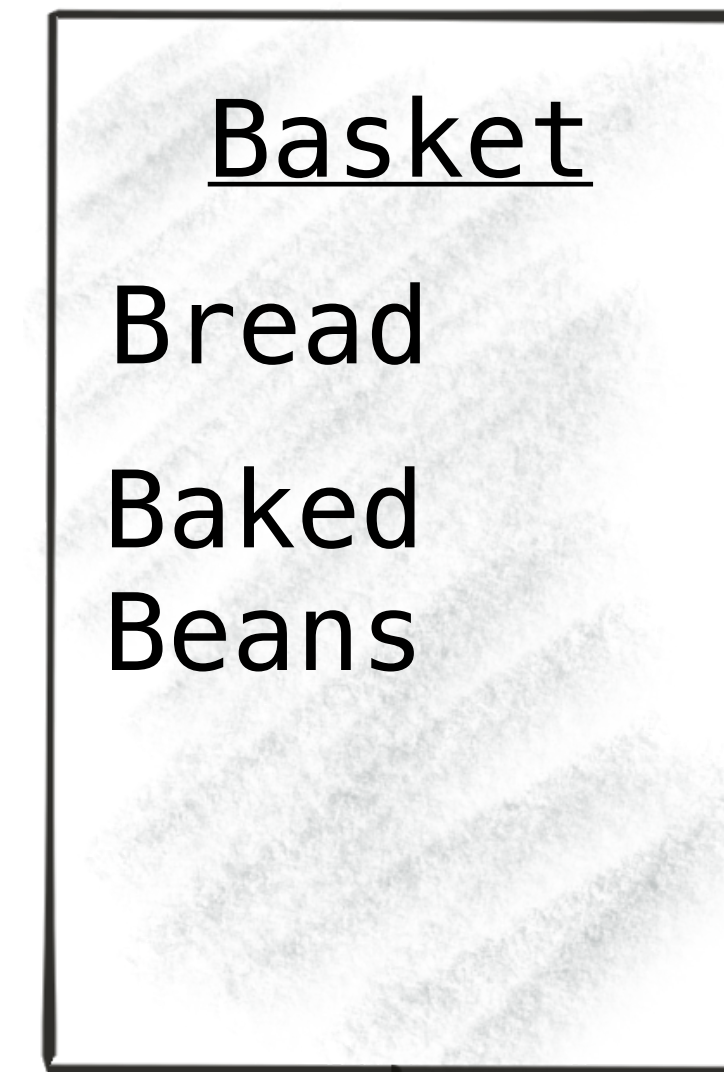
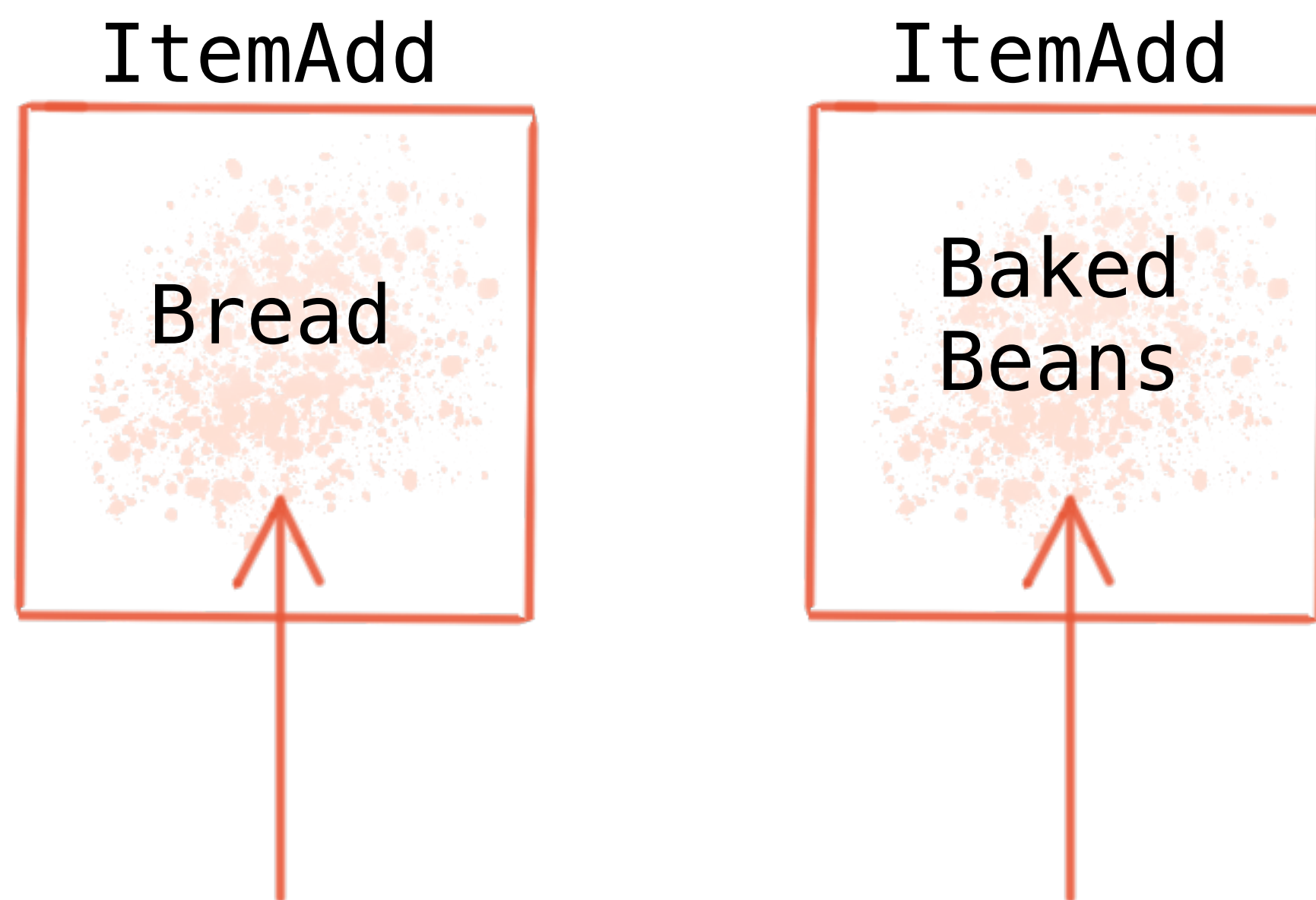


# Events



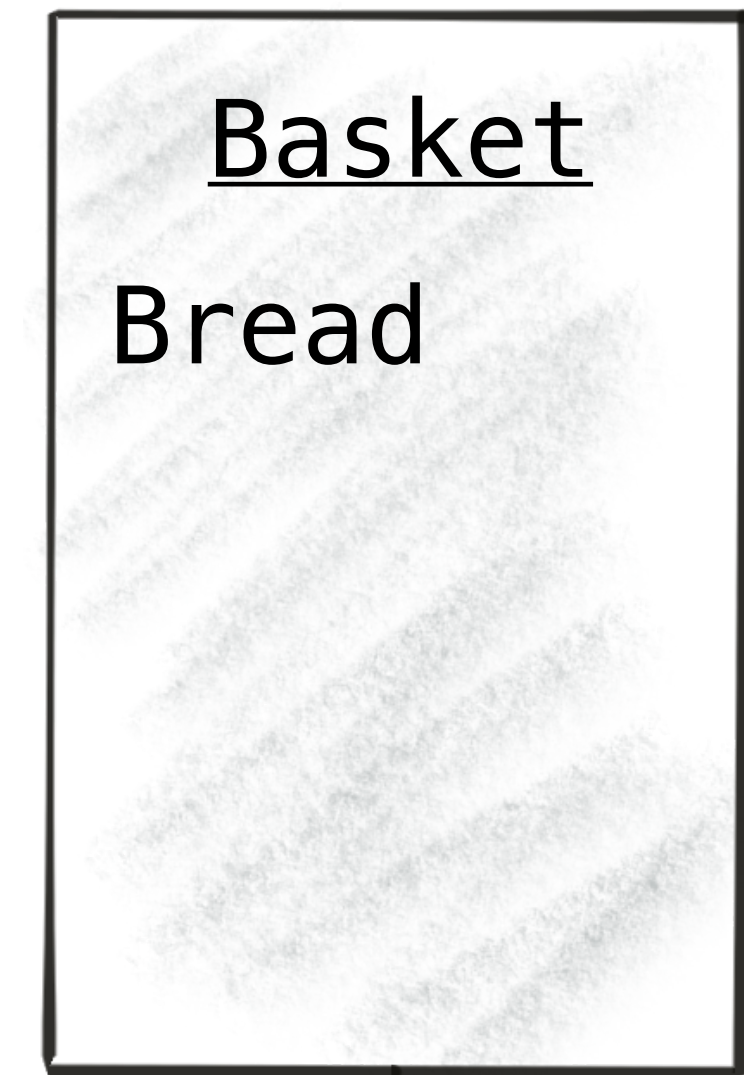
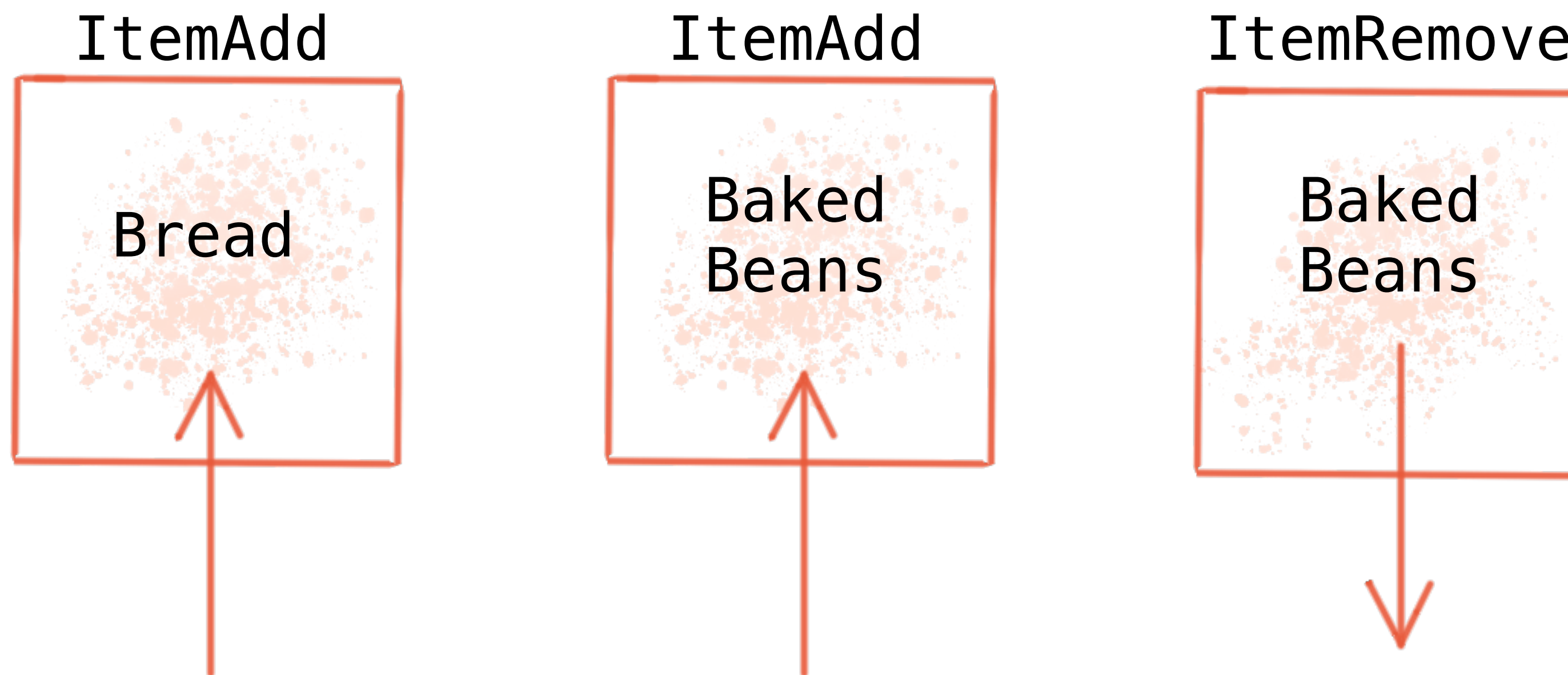


# Events



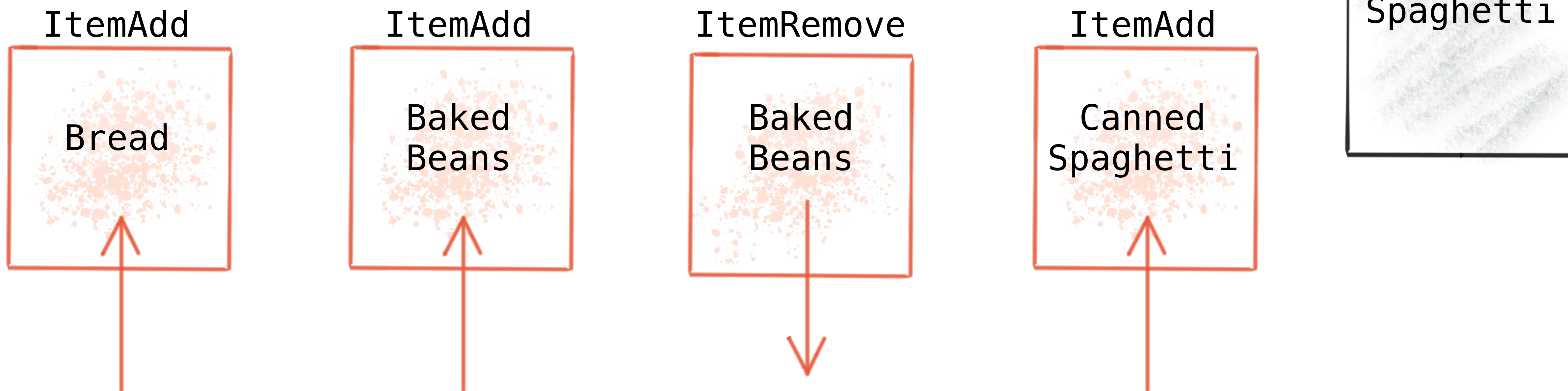


# Events



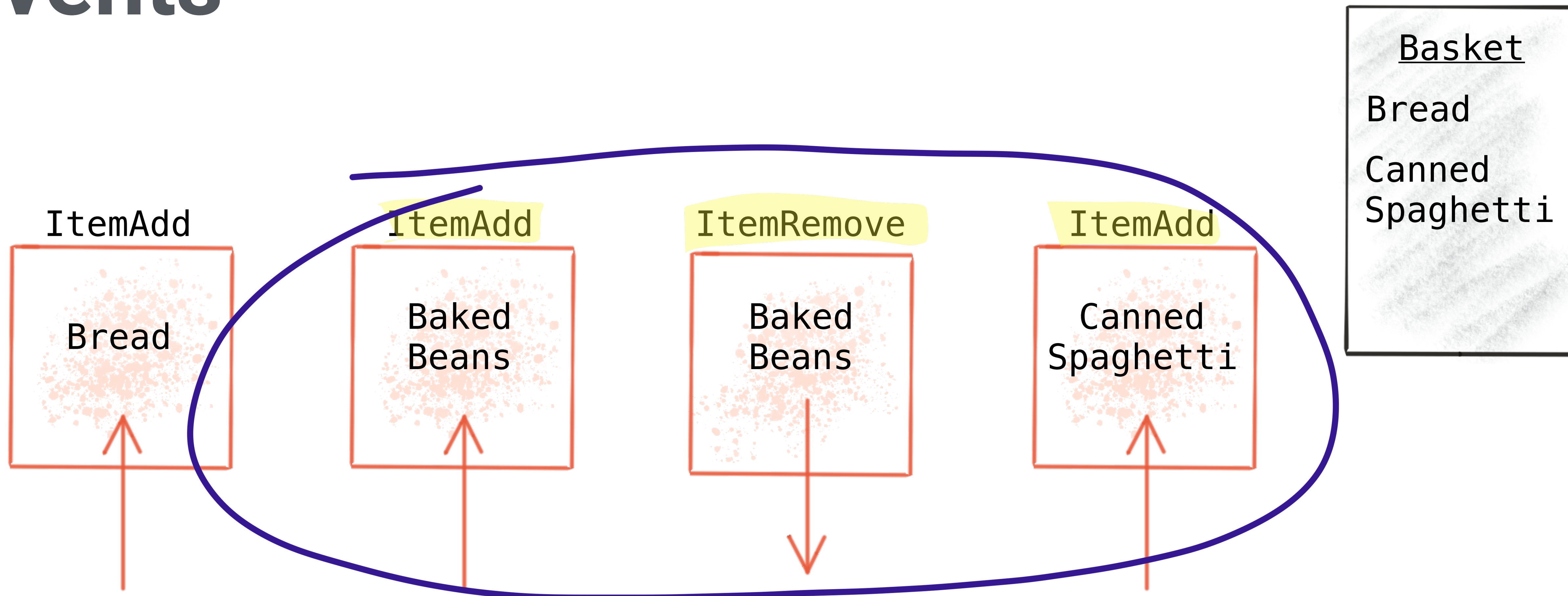


# Events





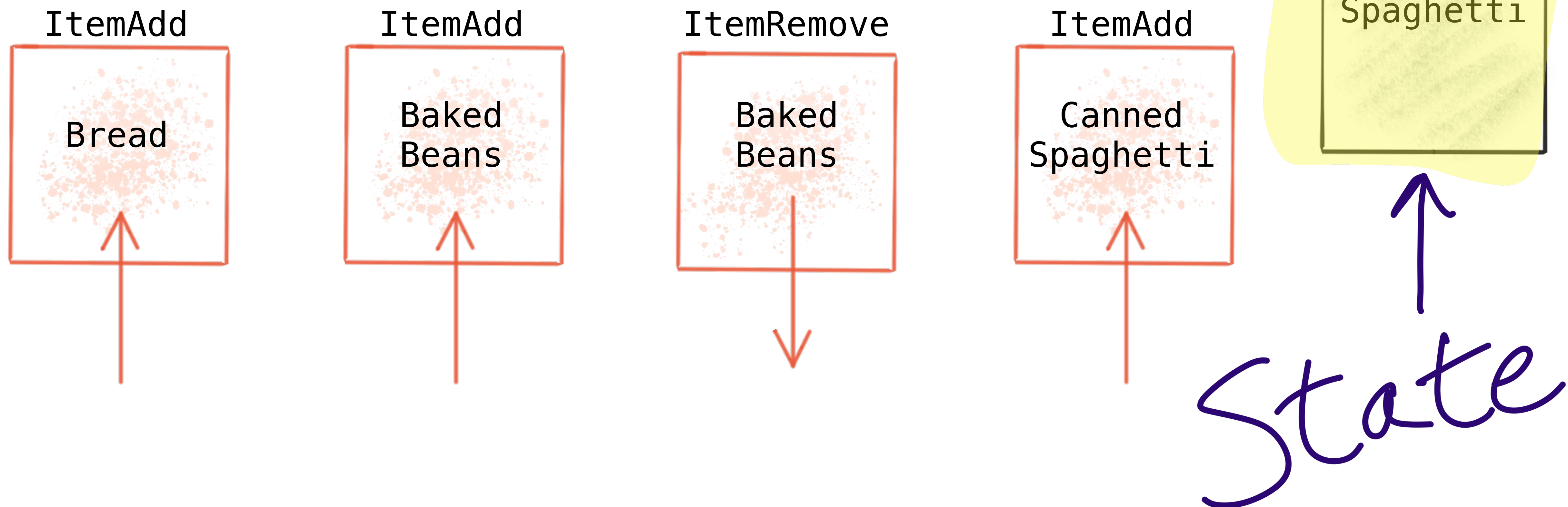
# Events



Behaviour



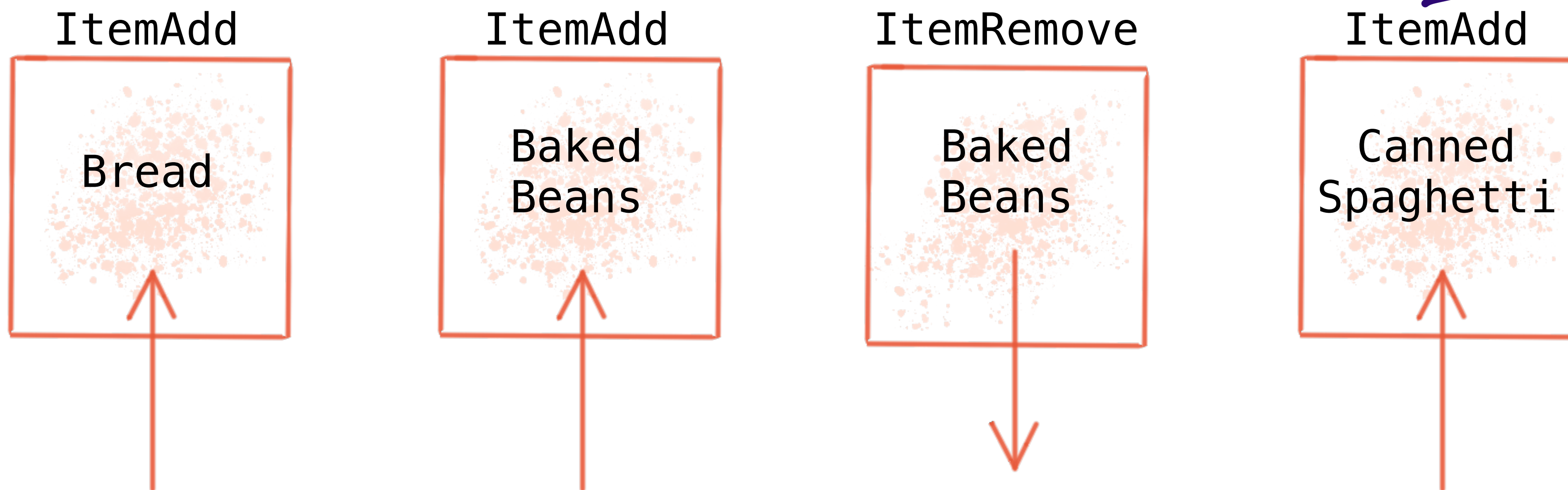
# Events





# Events

Event Stream



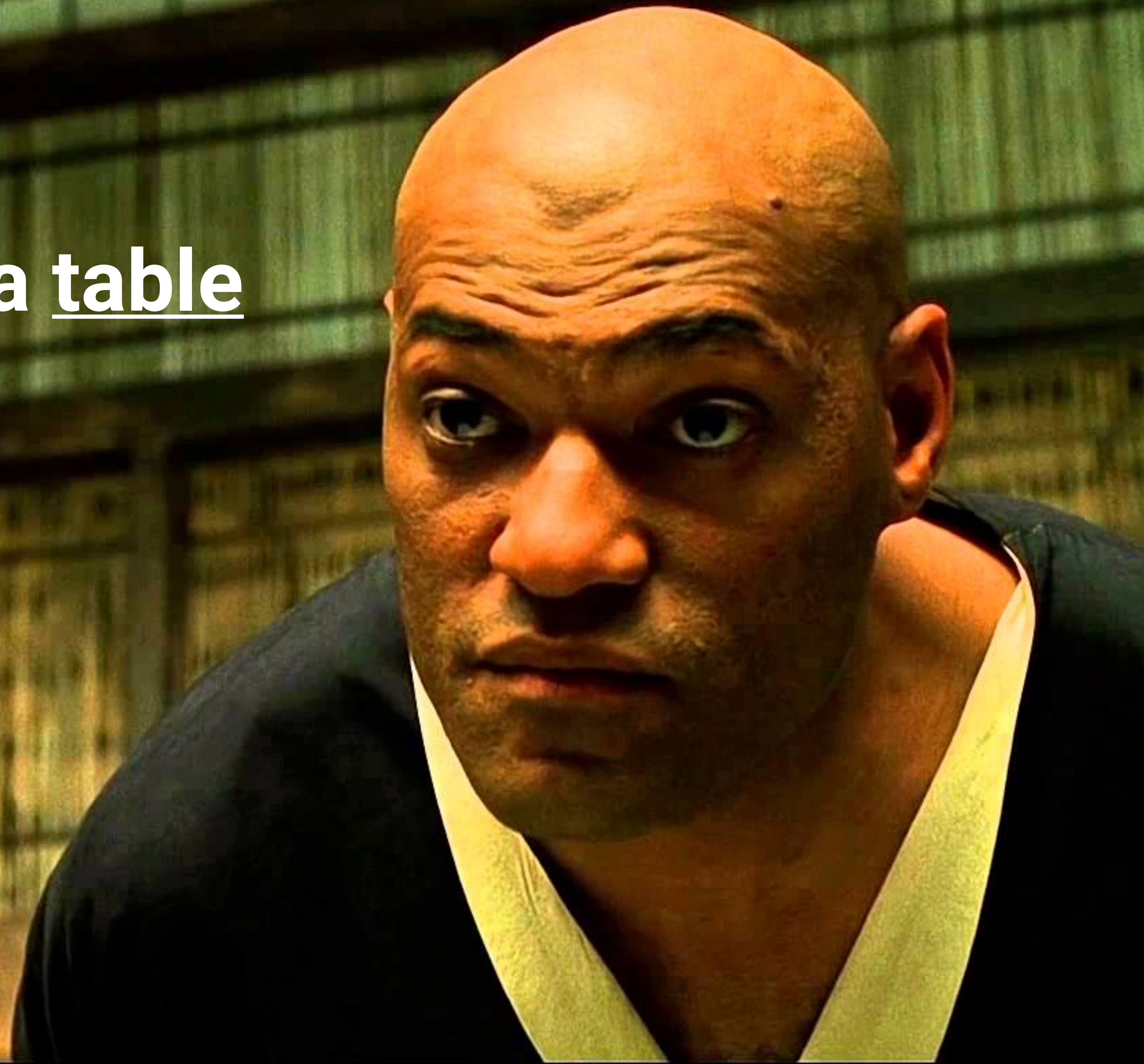


# Databases





**Do you think that's a table  
you are querying?**





# The Stream/Table Duality

*Stream*

Time



Account ID	Amount
12345	+ €50
12345	+ €25
12345	-€60

*Table*

Account ID	Balance
12345	€50

Account ID	Balance
12345	€75

Account ID	Balance
12345	€15



A night photograph of a campsite. A small, glowing tent is illuminated from within, casting a warm light on the surrounding grass and bushes. In the background, a large, dark tree stands prominently. The sky is filled with stars, and the Milky Way is visible, arching across the upper portion of the image. The overall scene is dark and serene, with the tent's light providing a focal point of warmth.

**The truth is the log.**

**The database is a cache  
of a subset of the log.**

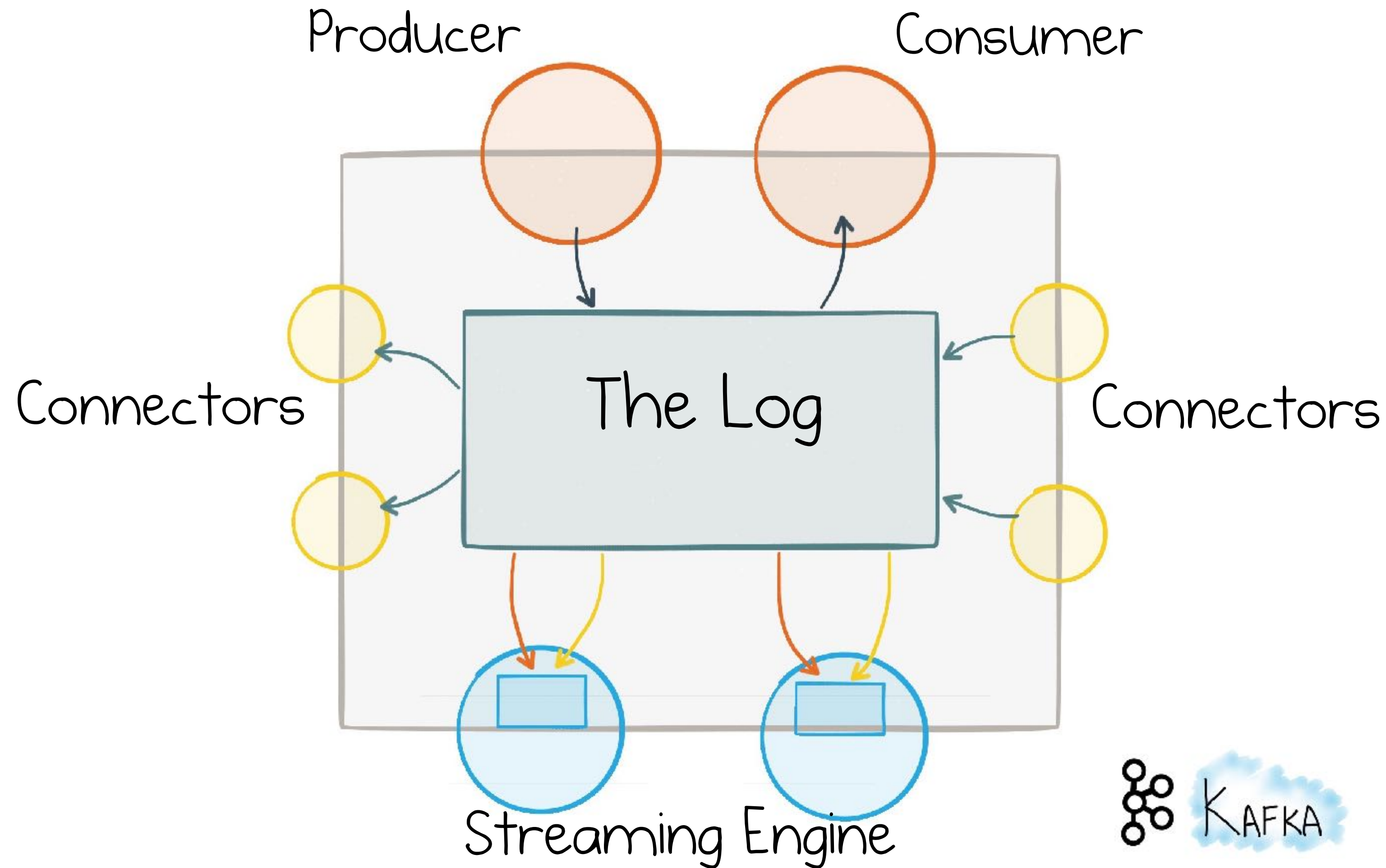
**—Pat Helland**

**Immutability Changes Everything**

[http://cidrdb.org/cidr2015/Papers/CIDR15\\_Paper16.pdf](http://cidrdb.org/cidr2015/Papers/CIDR15_Paper16.pdf)

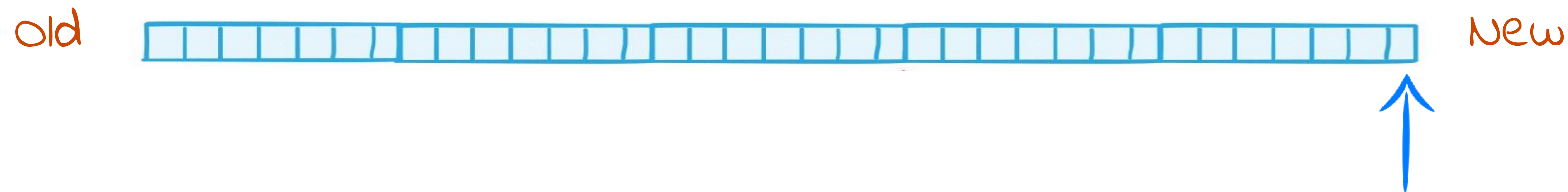


# What is an Event Streaming Platform?





# Immutable Event Log



Messages are added at the end of the log

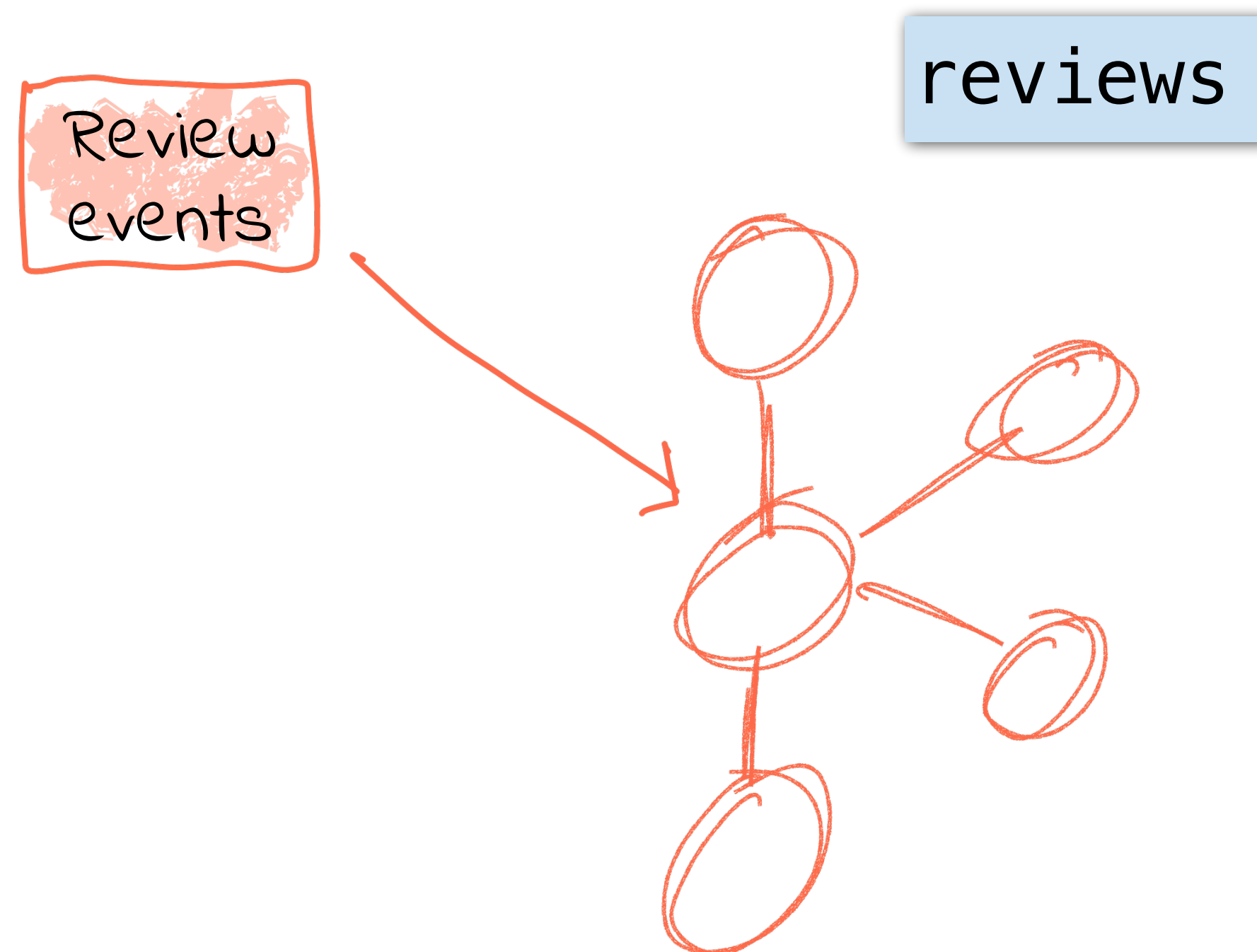




**This is  
Something  
New**

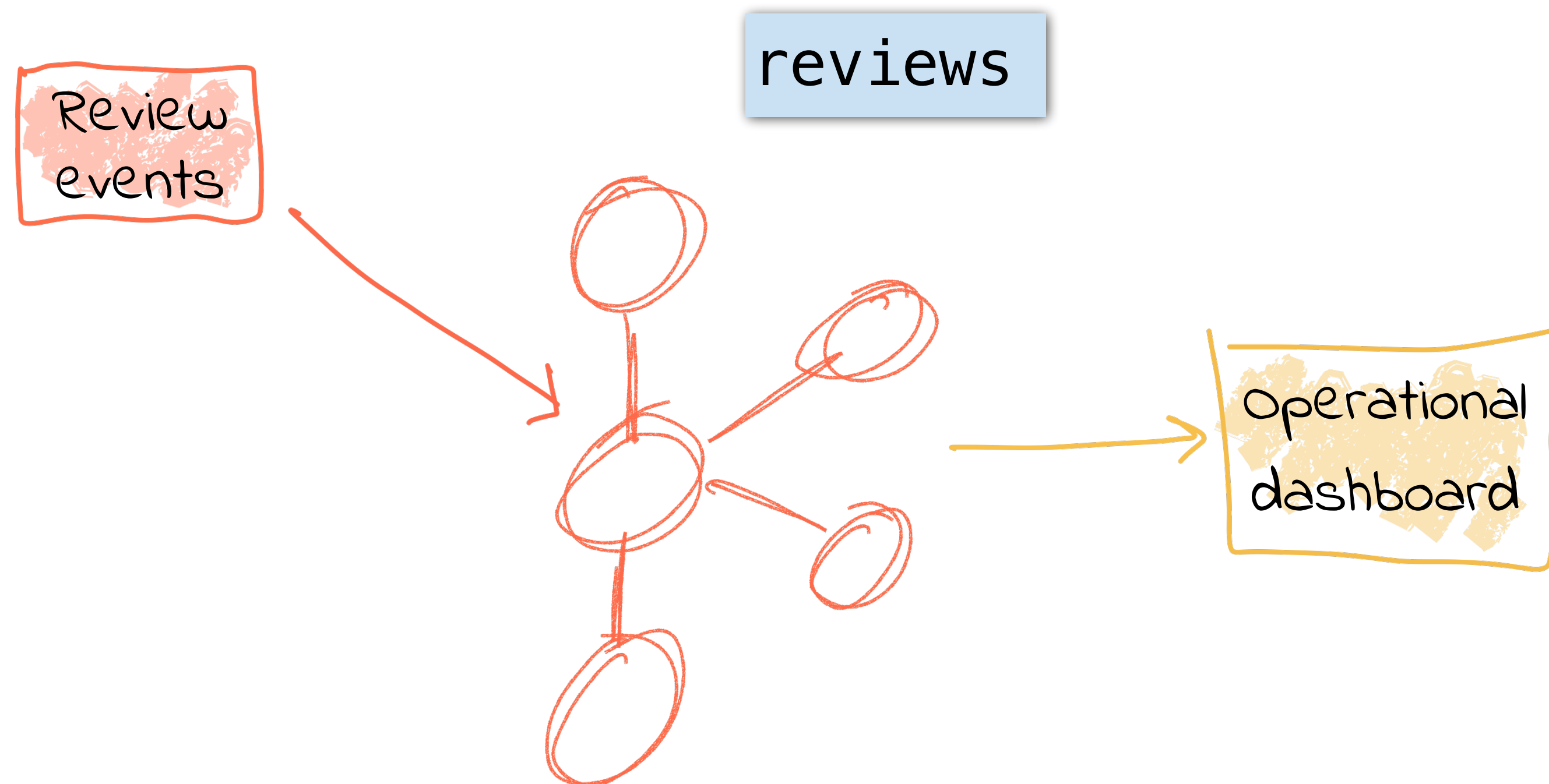


# Events in Action



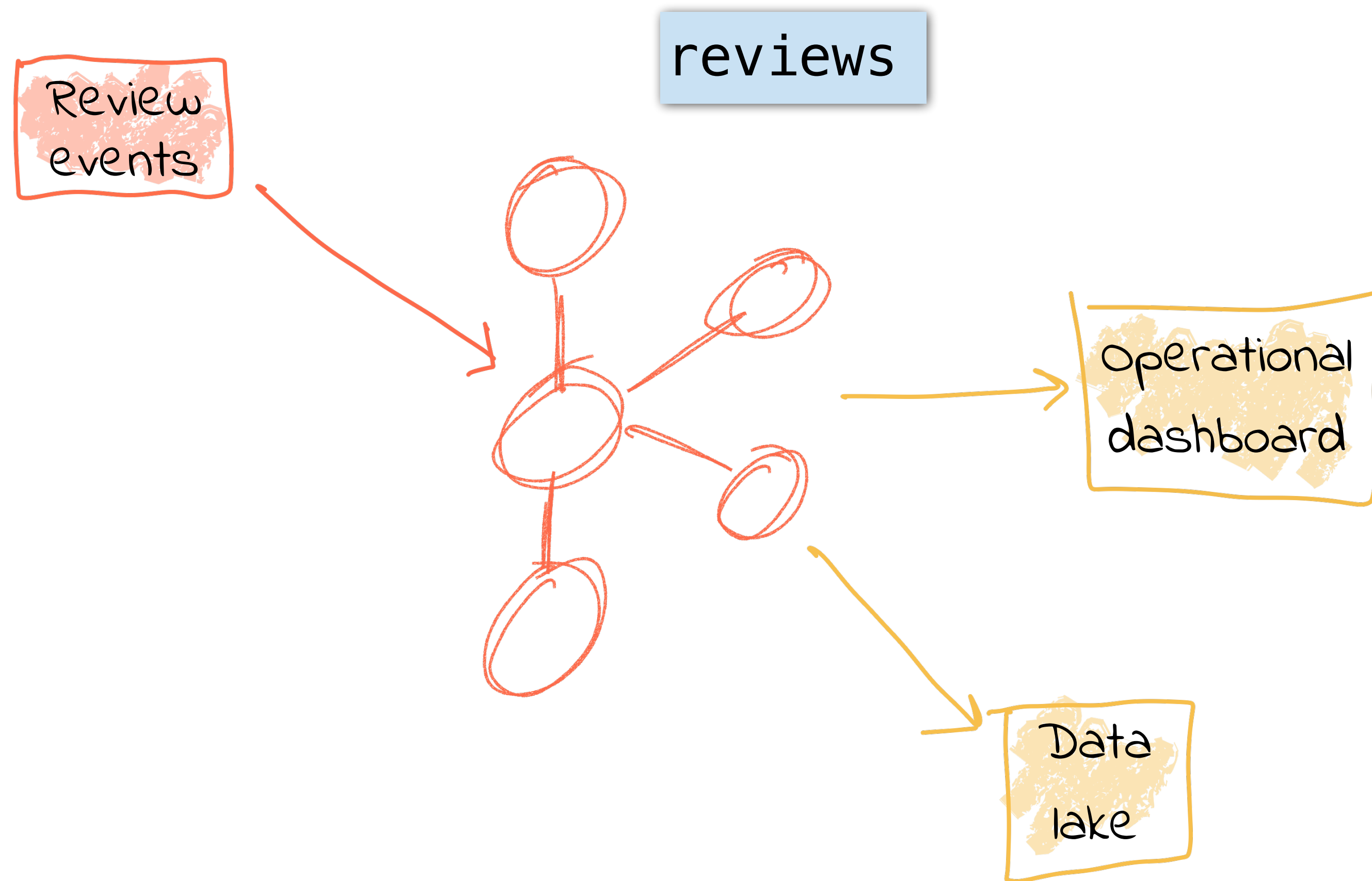


# Events in Action



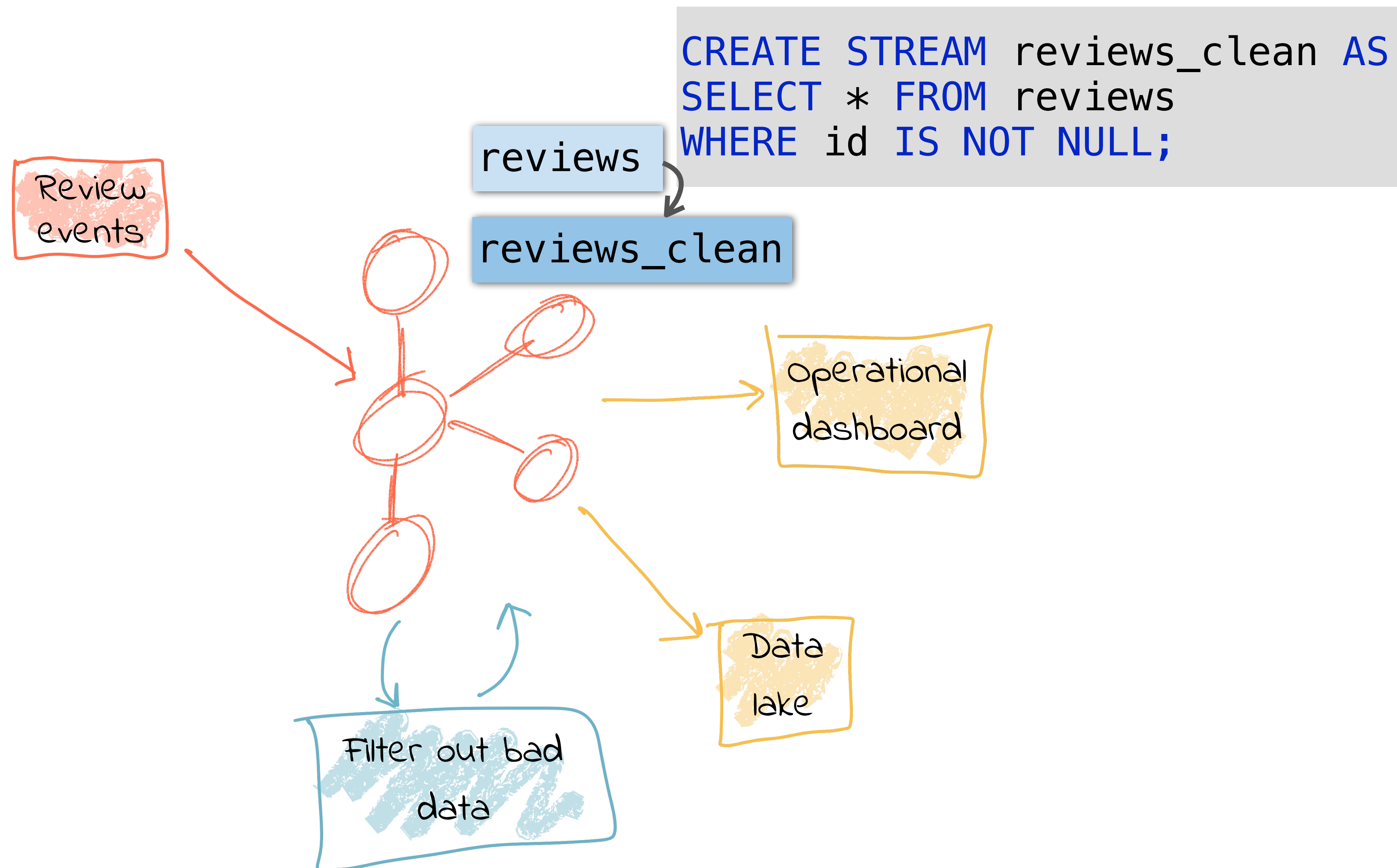


# Events in Action



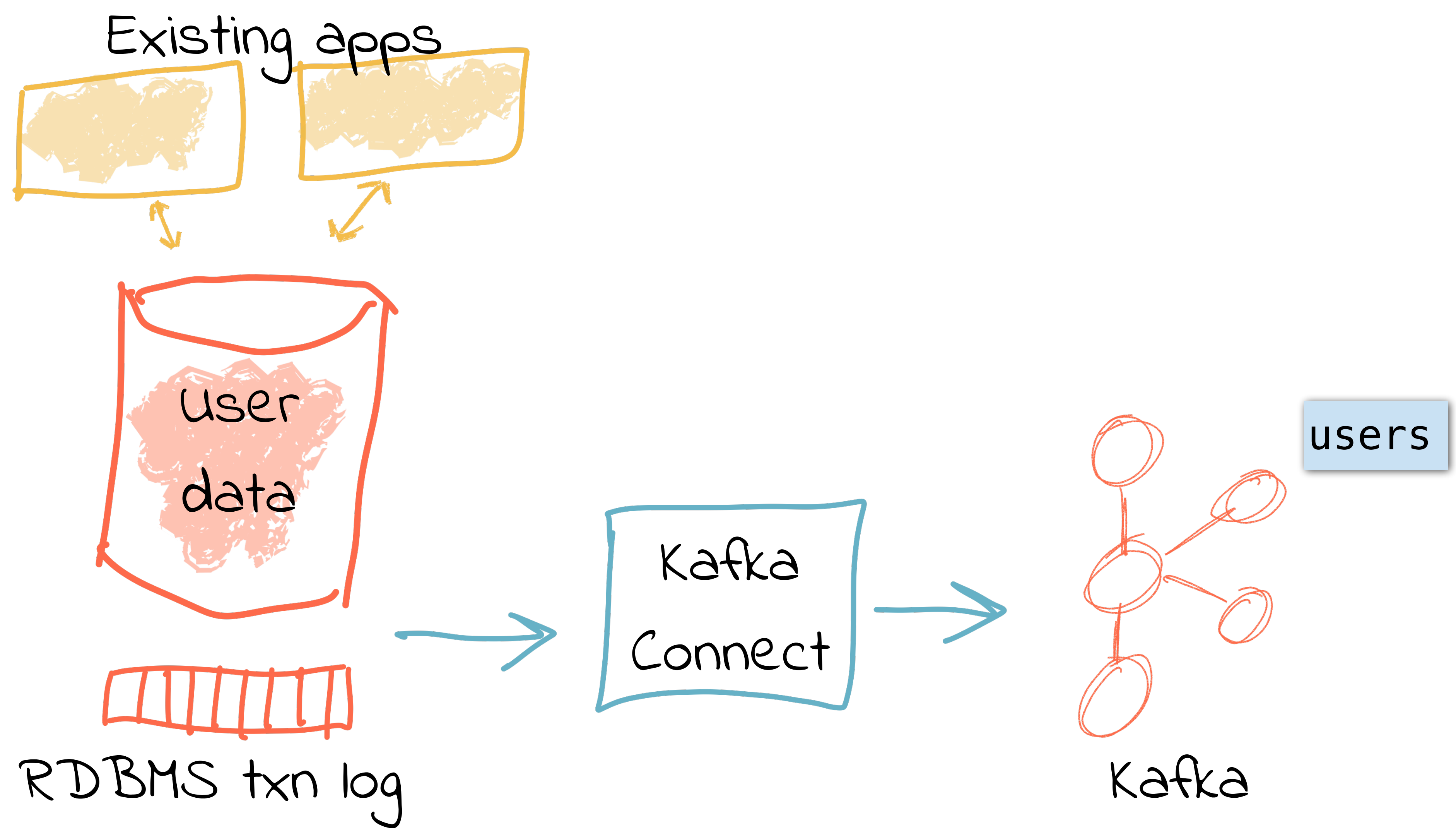


# Events in Action



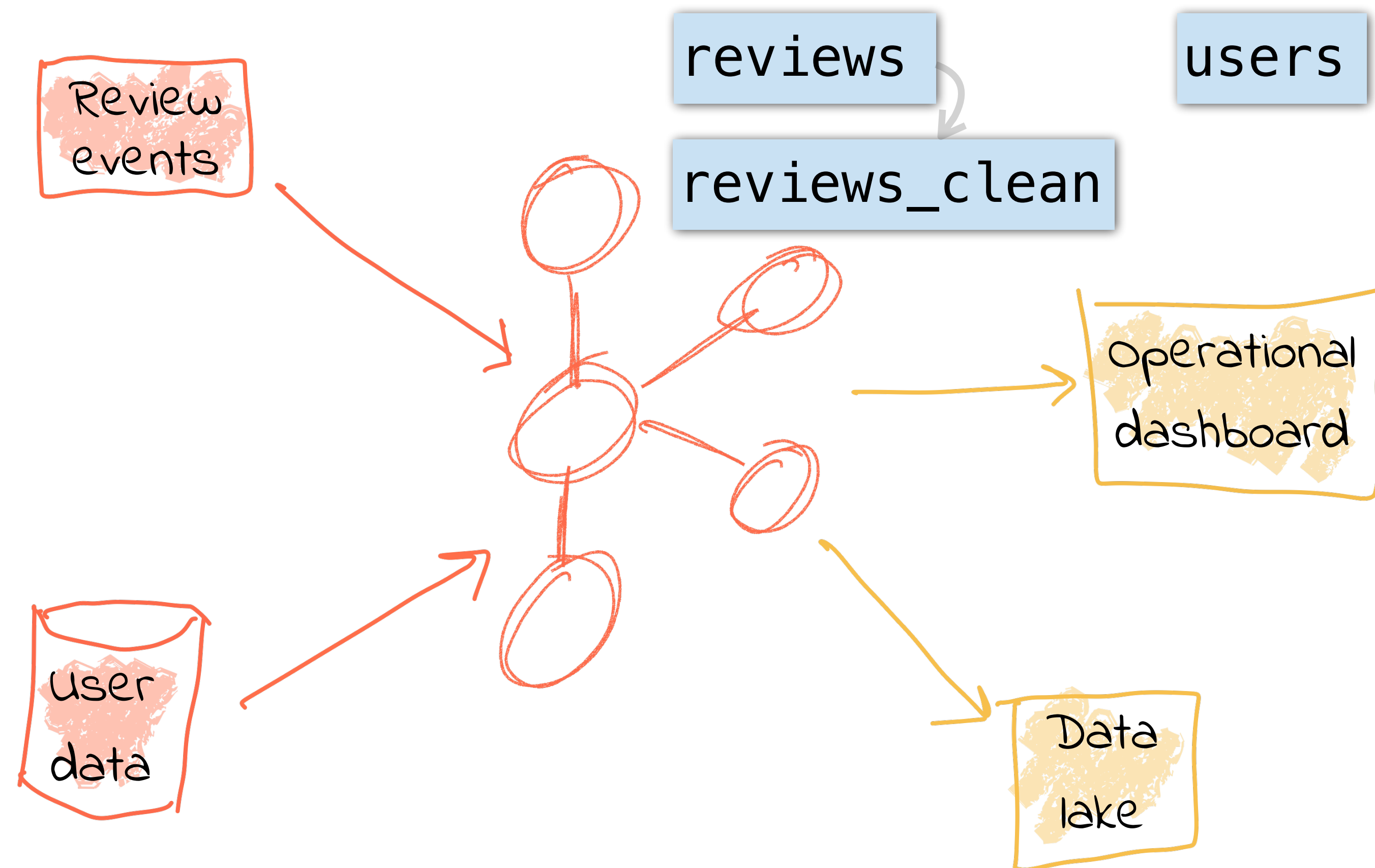


# Events in Action





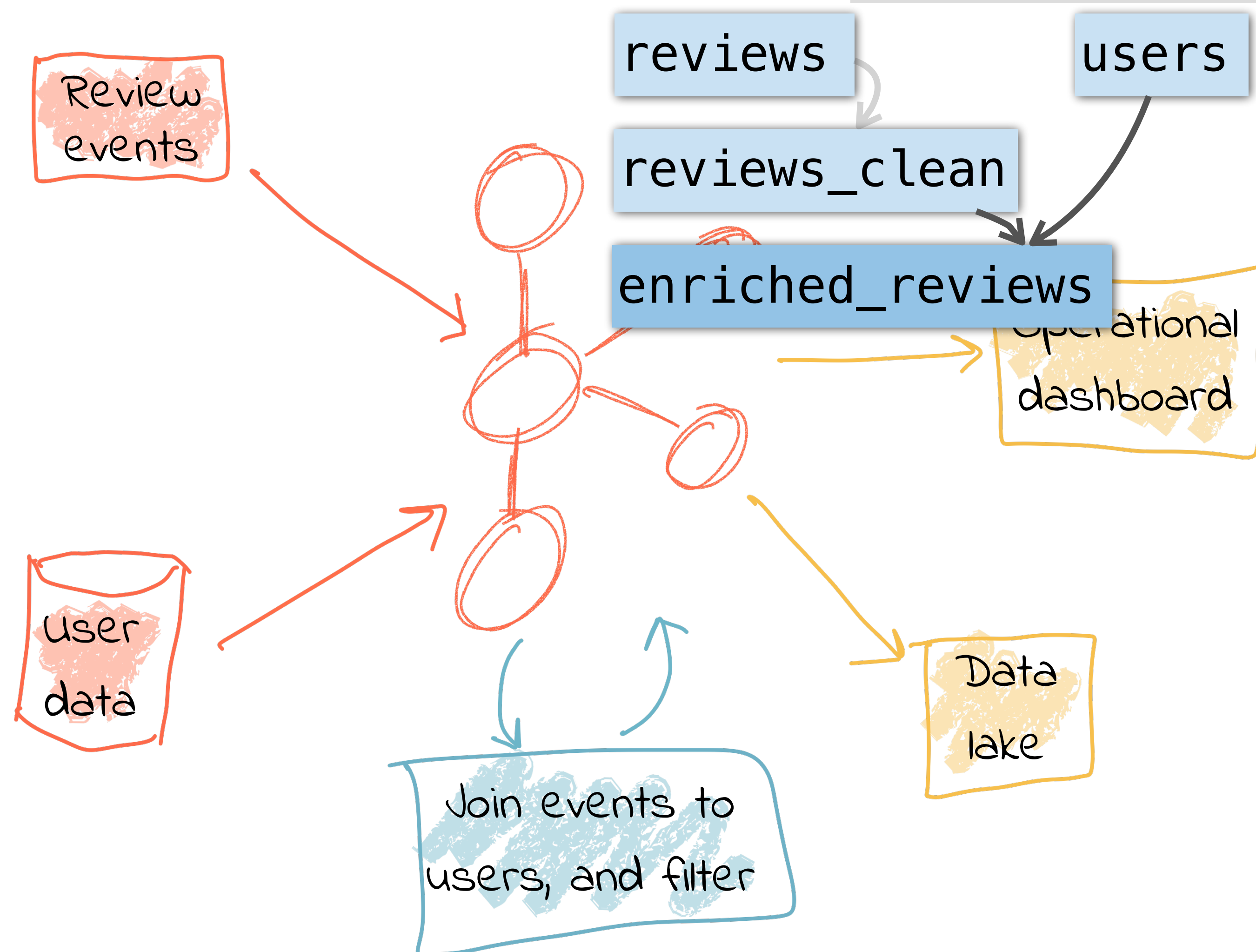
# Events in Action





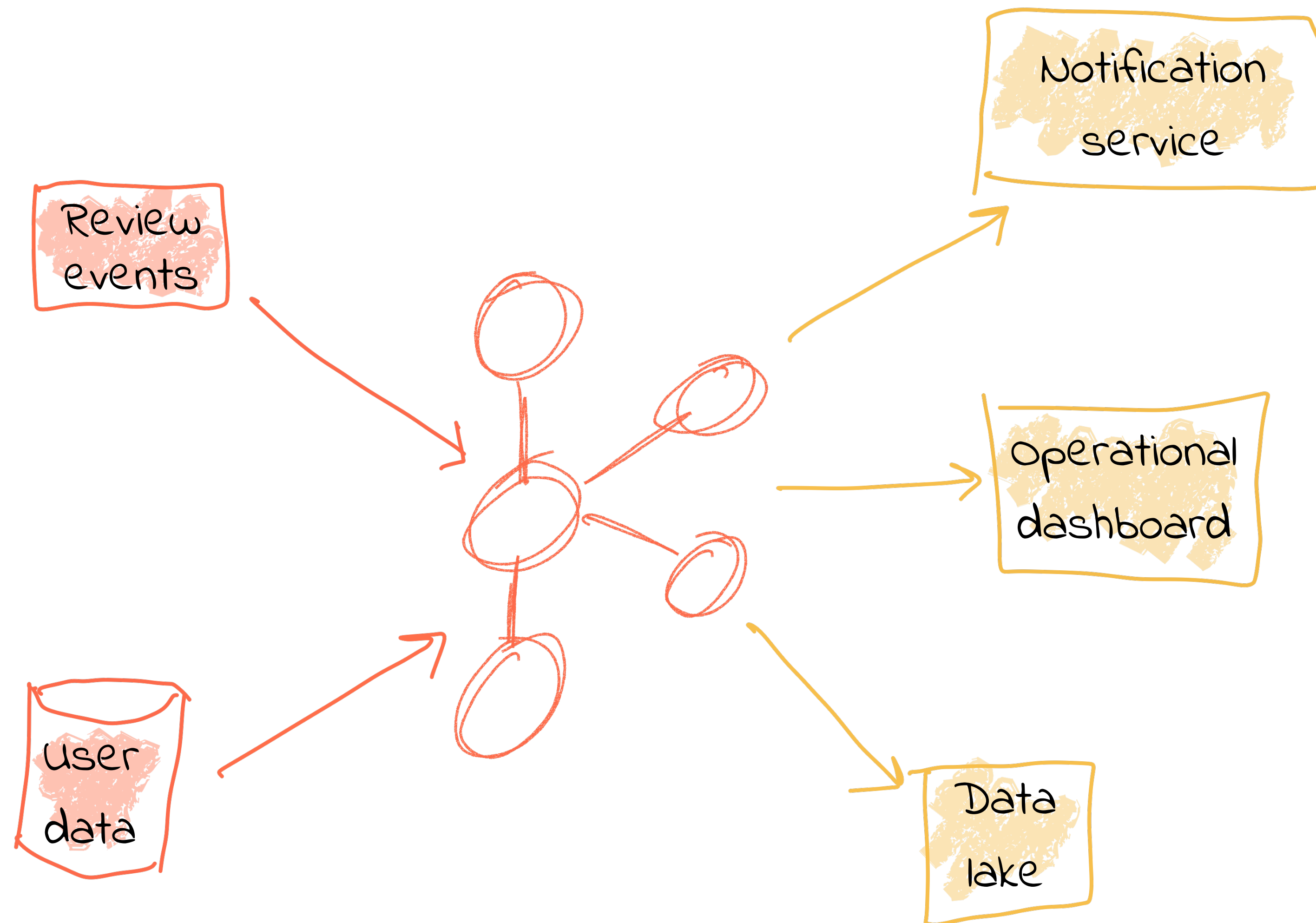
# Events in Action

```
CREATE STREAM enriched_reviews AS  
SELECT * FROM reviews_clean r  
      INNER JOIN users u  
      ON r.userid=u.userid;
```





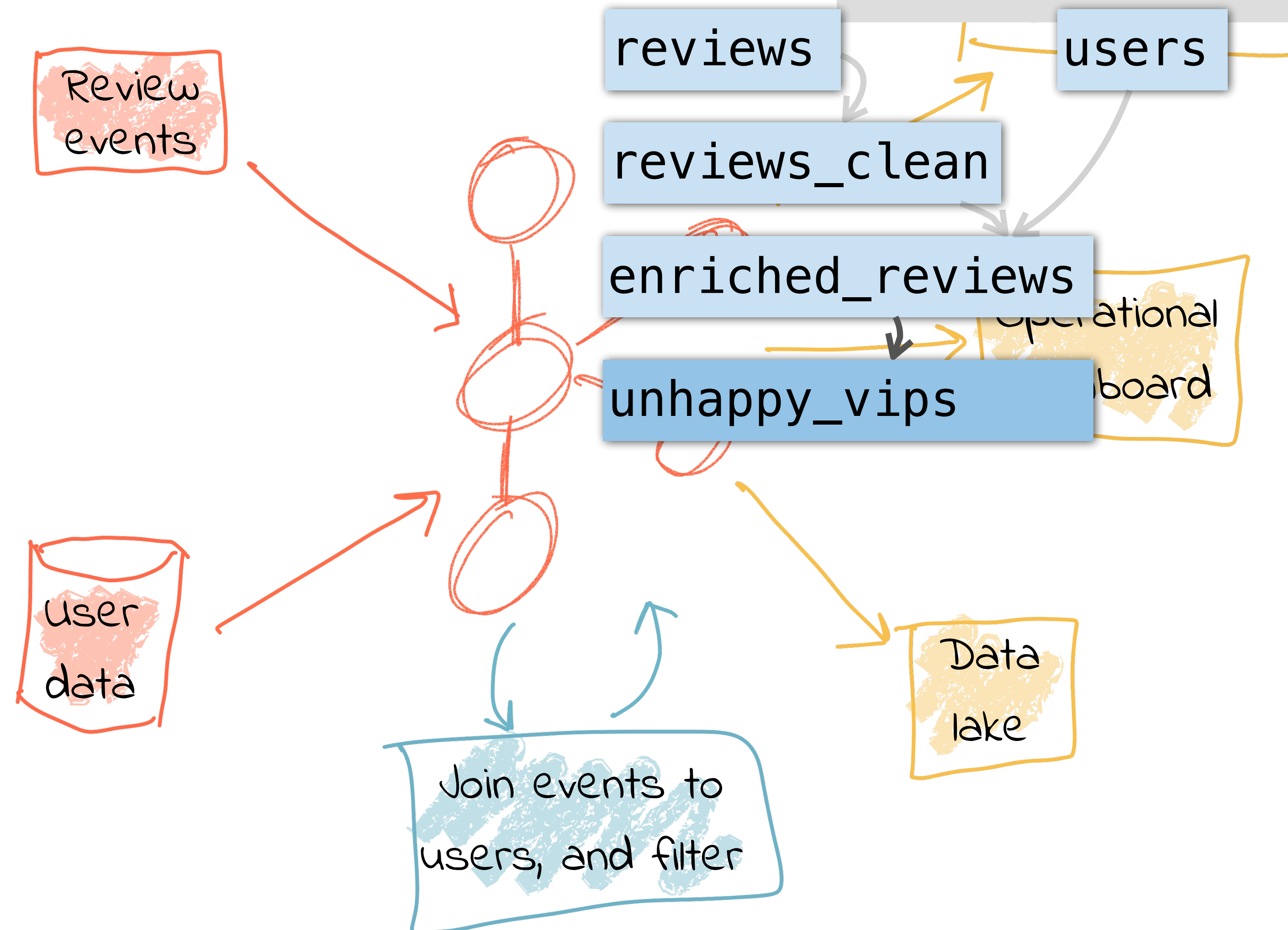
# Events in Action





# Events in Action

```
CREATE STREAM unhappy_vips AS  
SELECT * FROM enriched_reviews  
WHERE rating < 3  
AND status = 'Platinum';
```



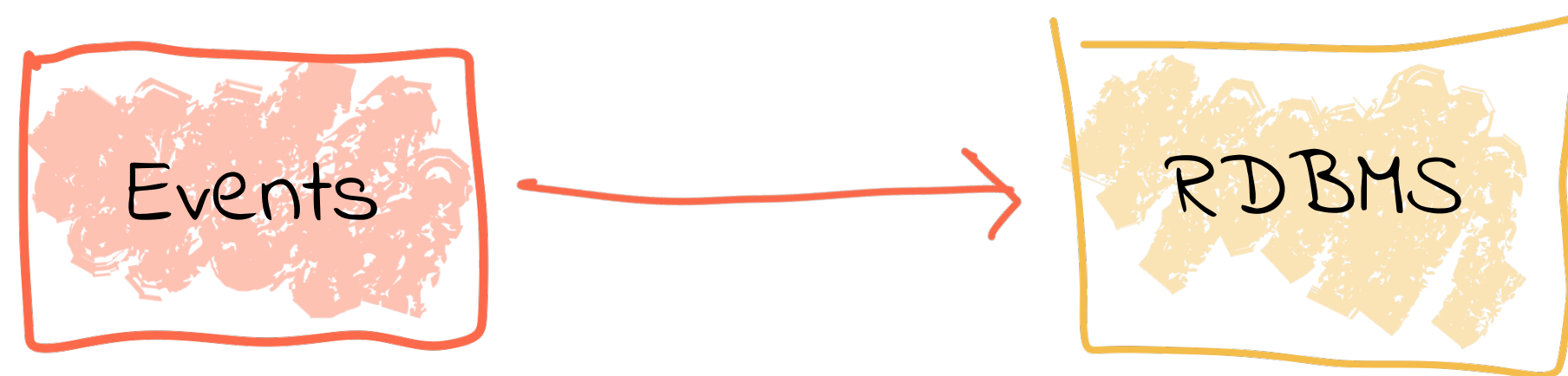


# The Power of an Event-Driven Architecture



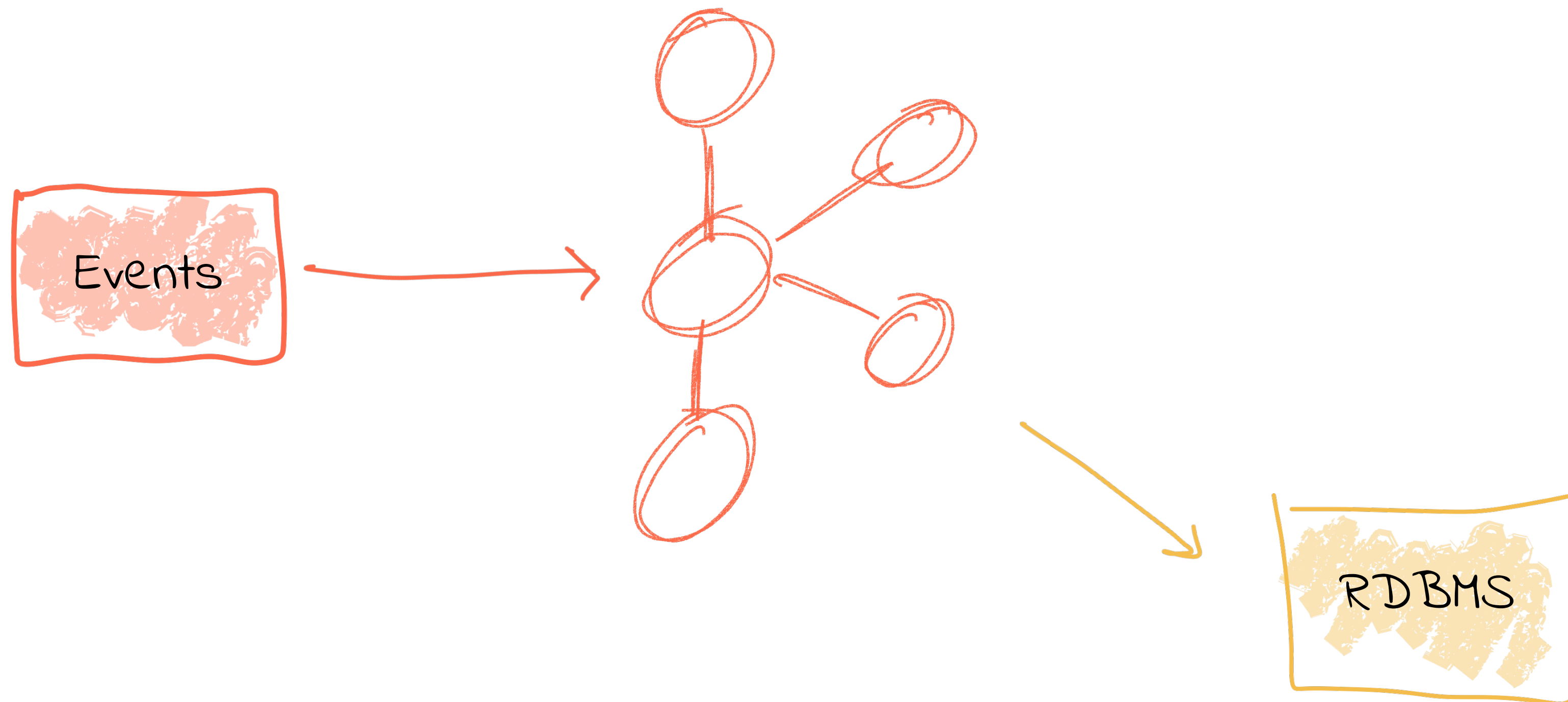


# Not Everything is a Nail



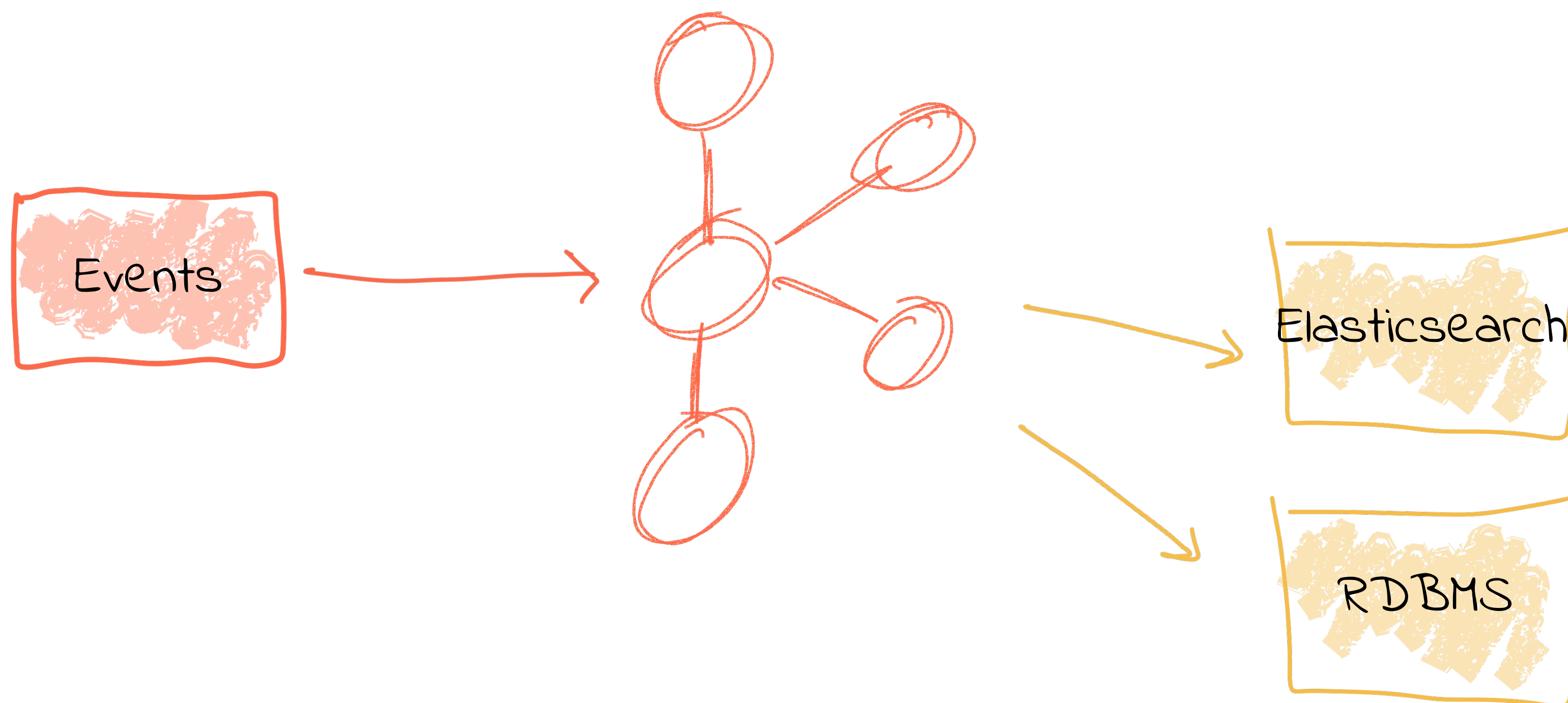


# Not Everything is a Nail



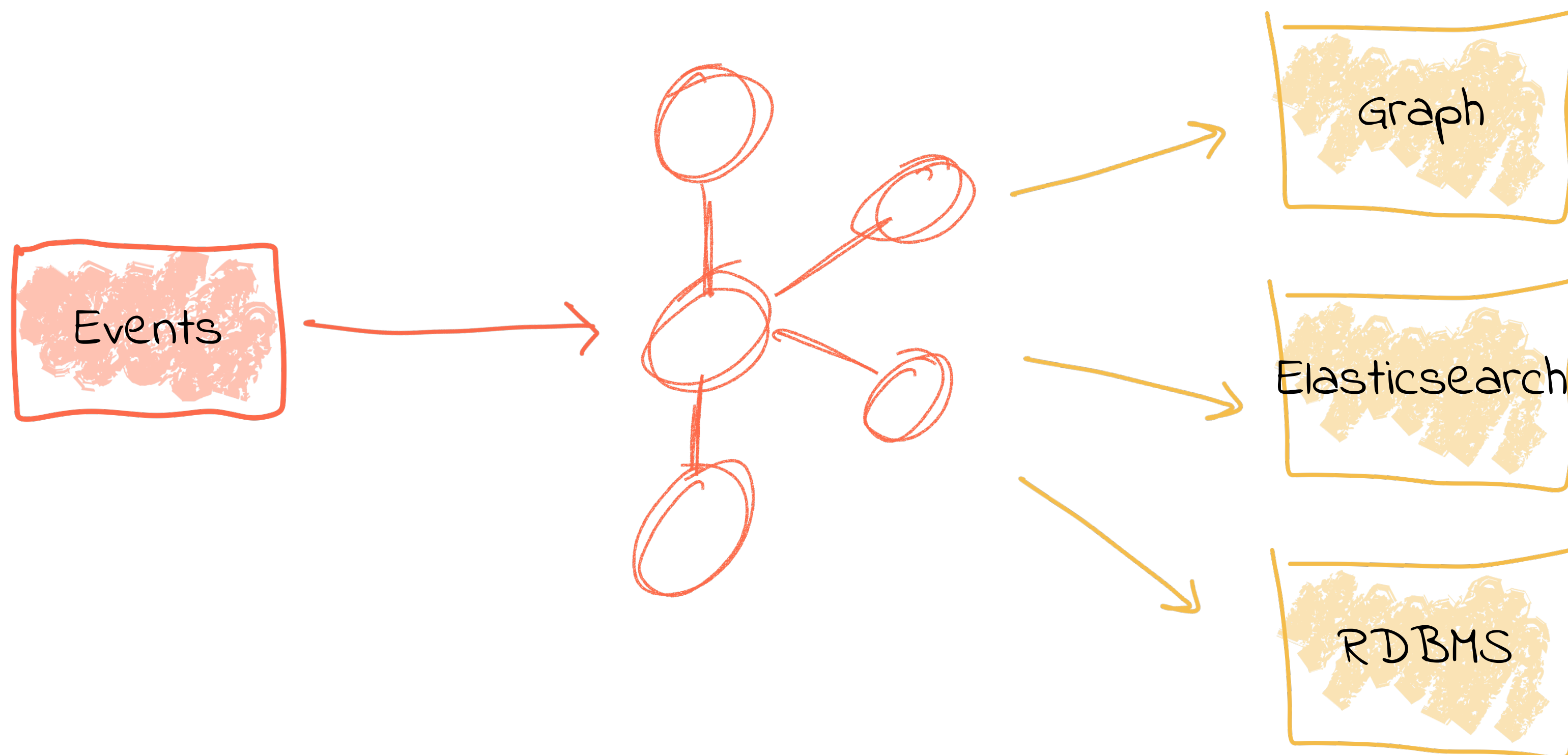


# Not Everything is a Nail



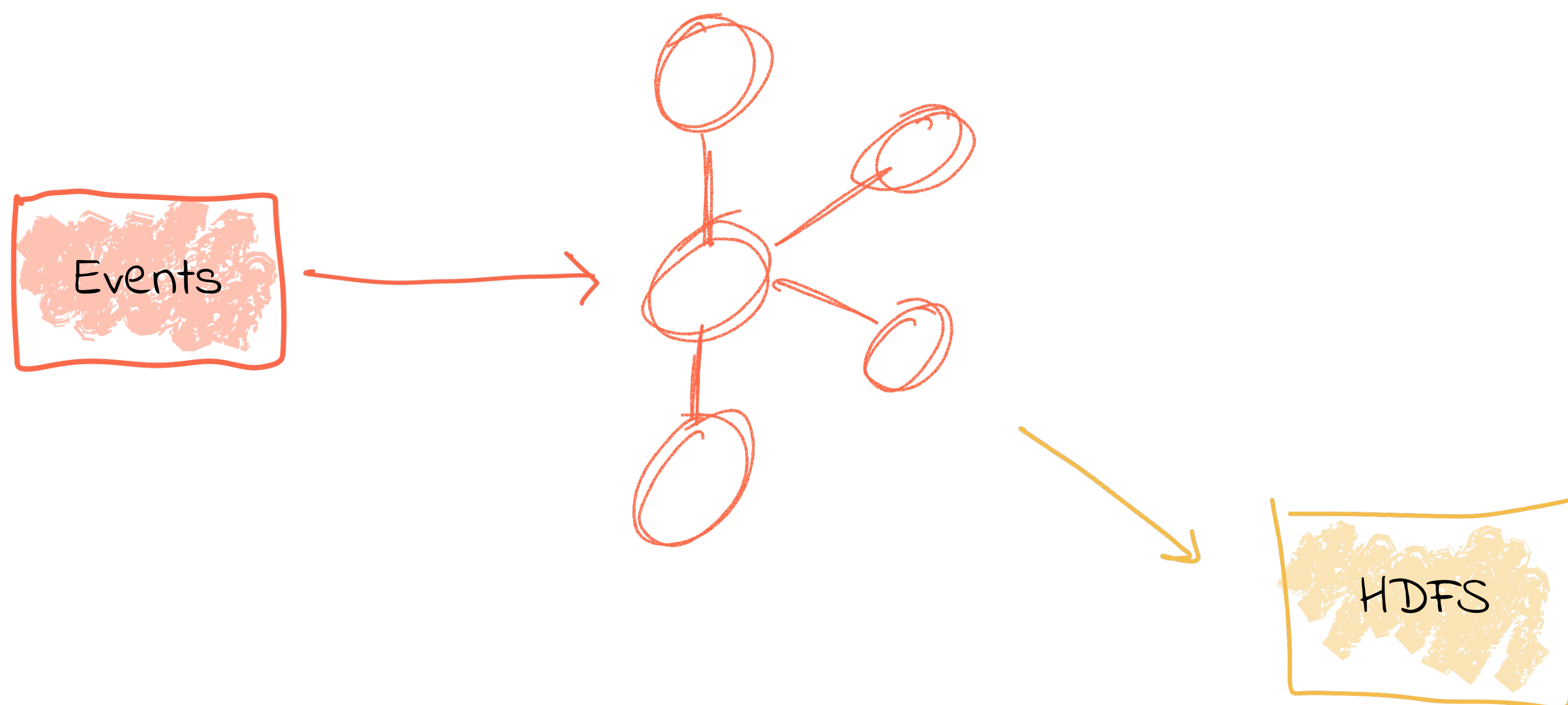


# Not Everything is a Nail



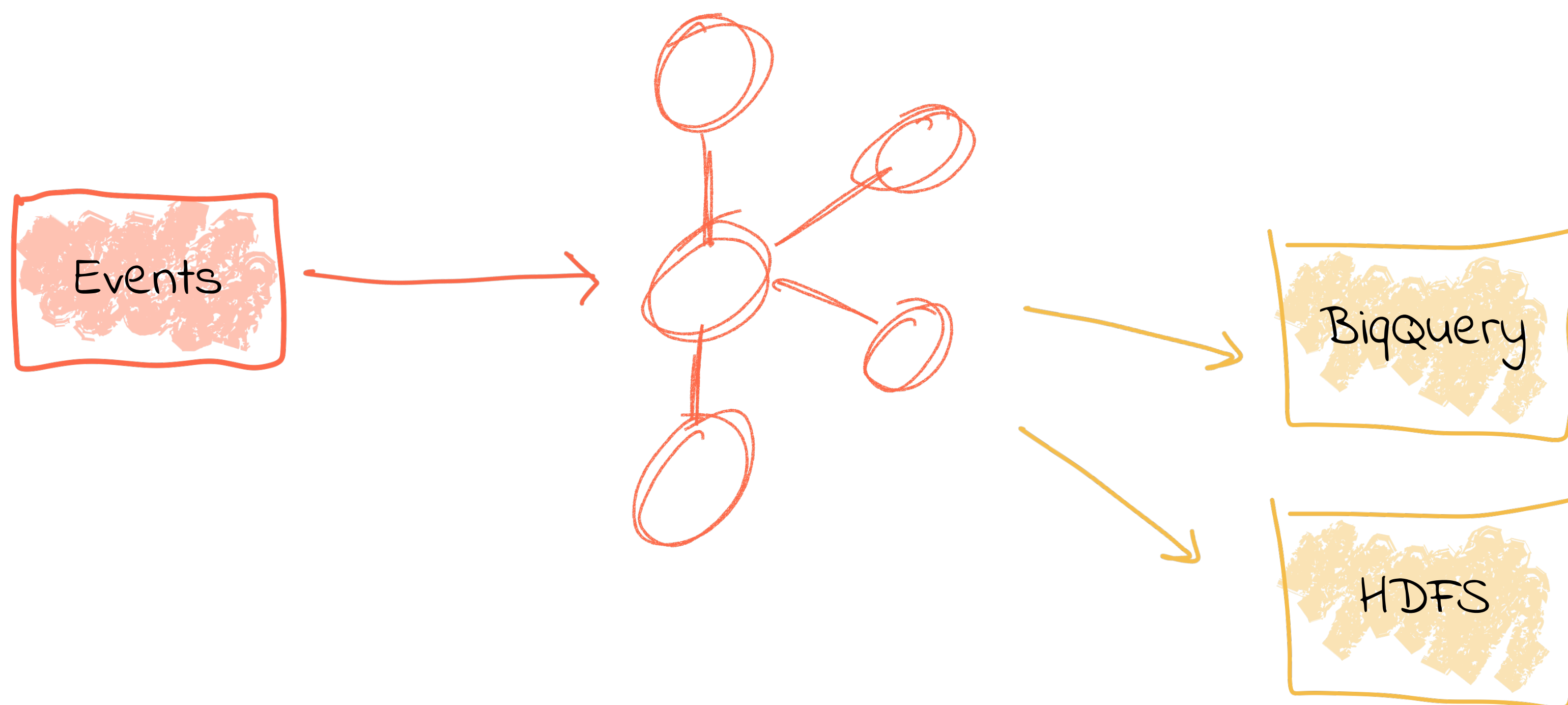


# Side-by-Side Tech Evaluation



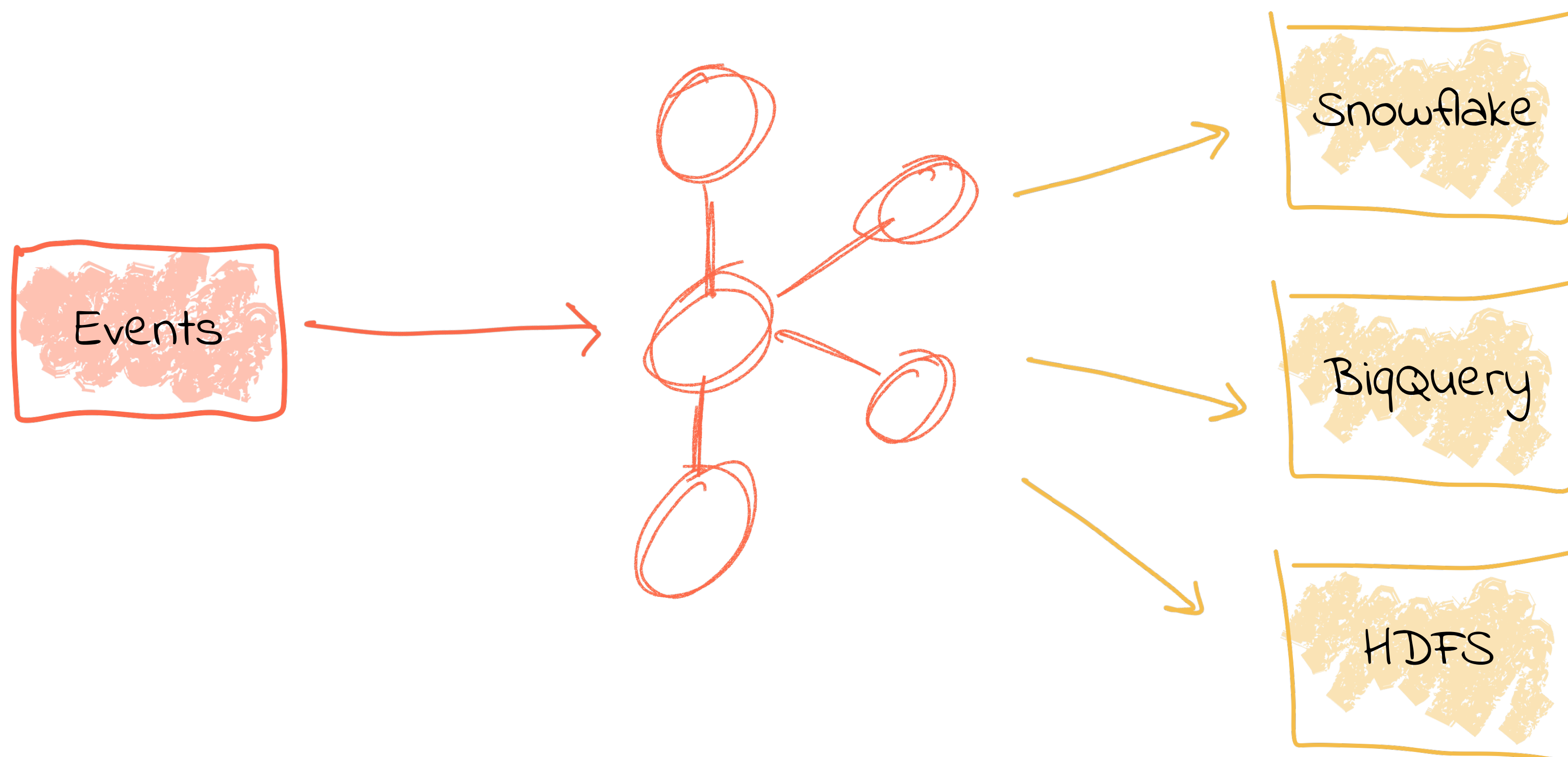


# Side-by-Side Tech Evaluation



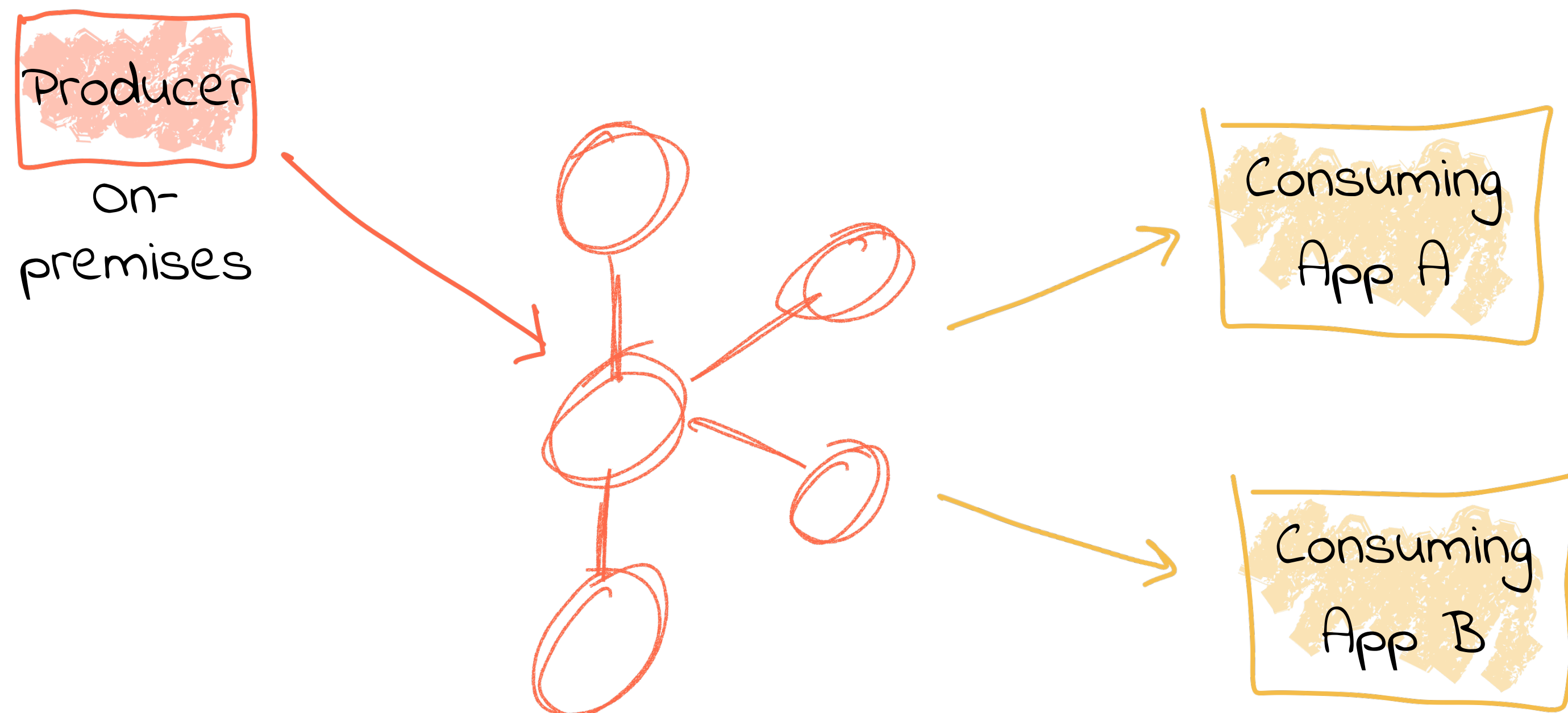


# Side-by-Side Tech Evaluation



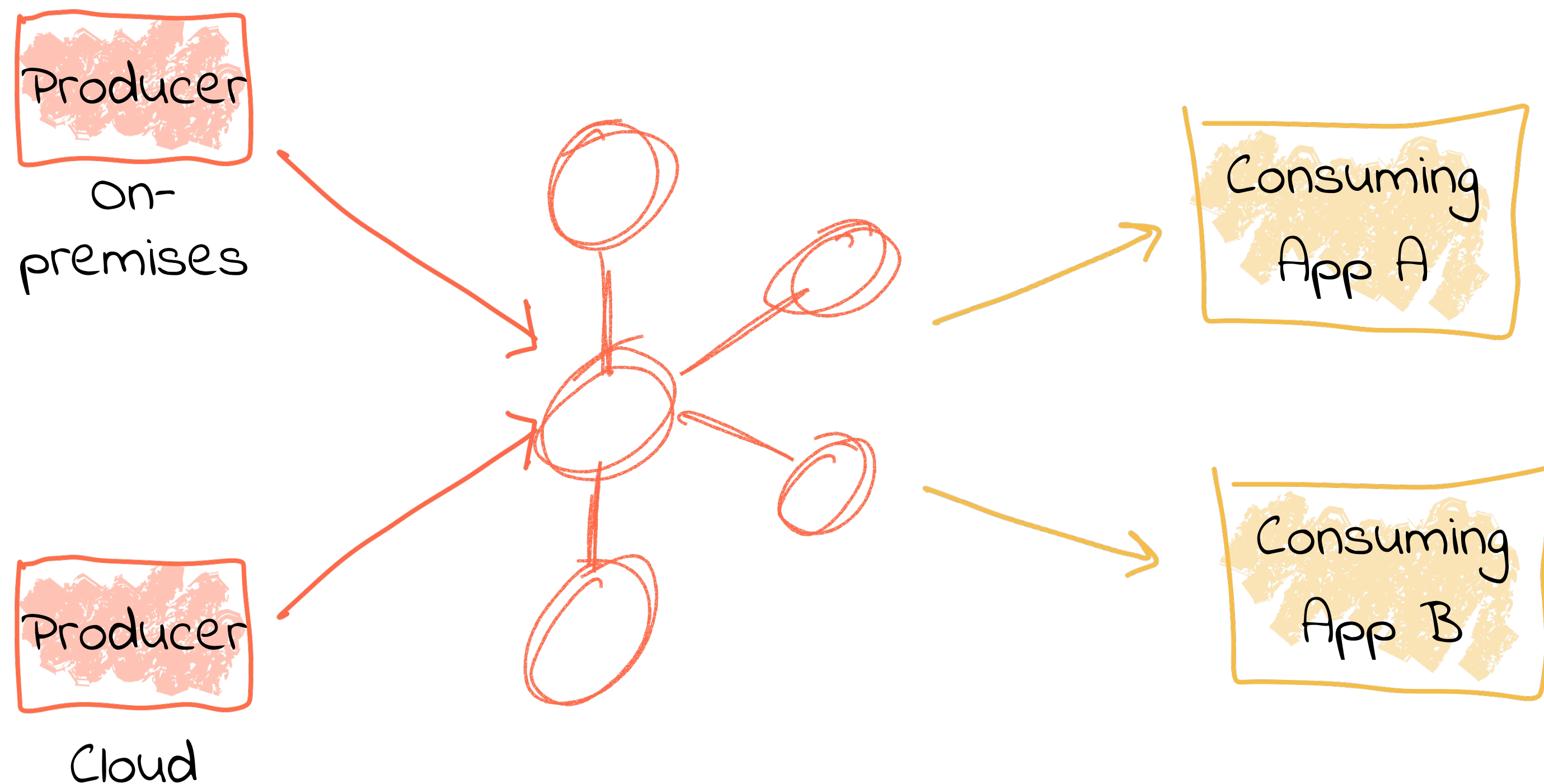


# Evolve Data Sources



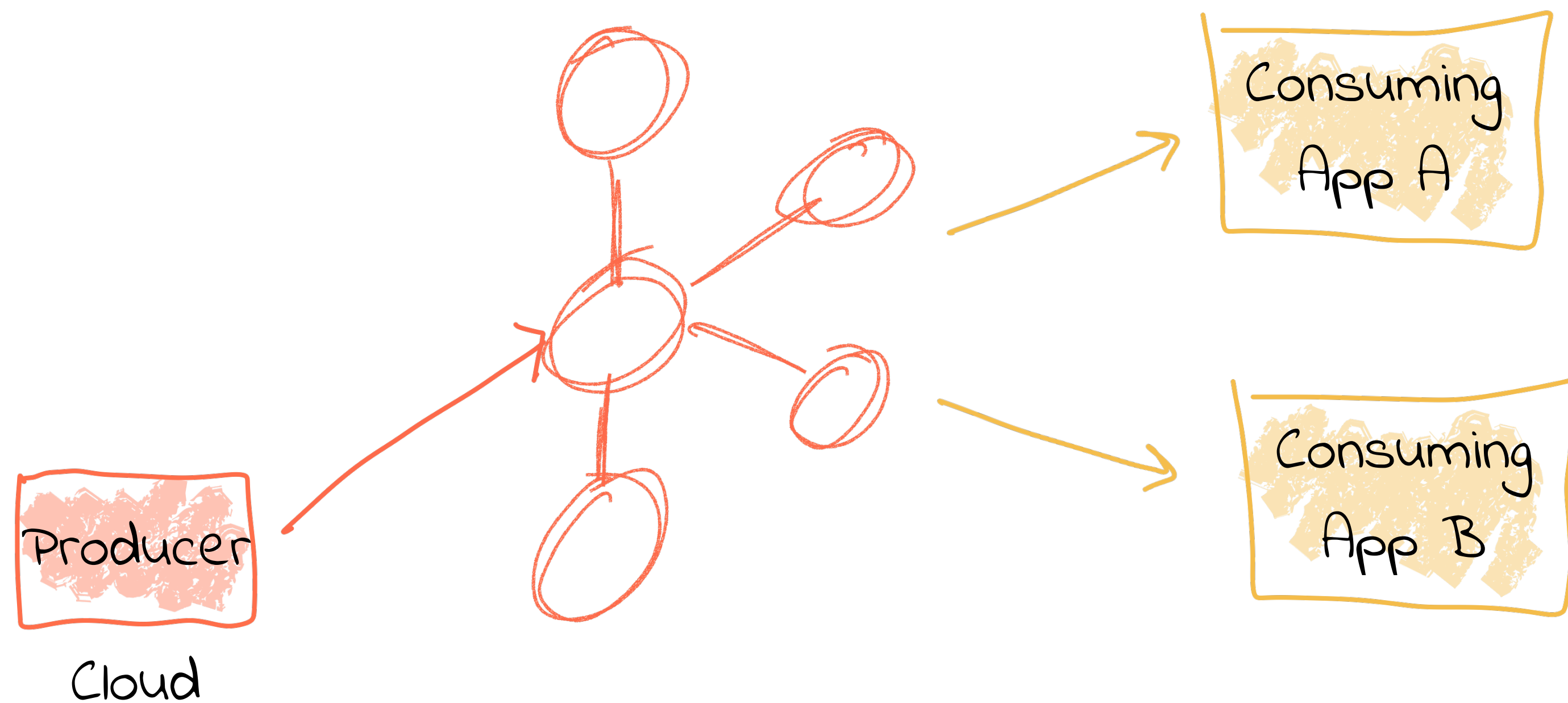


# Evolve Data Sources





# Evolve Data Sources





# Tight Coupling != Flexible



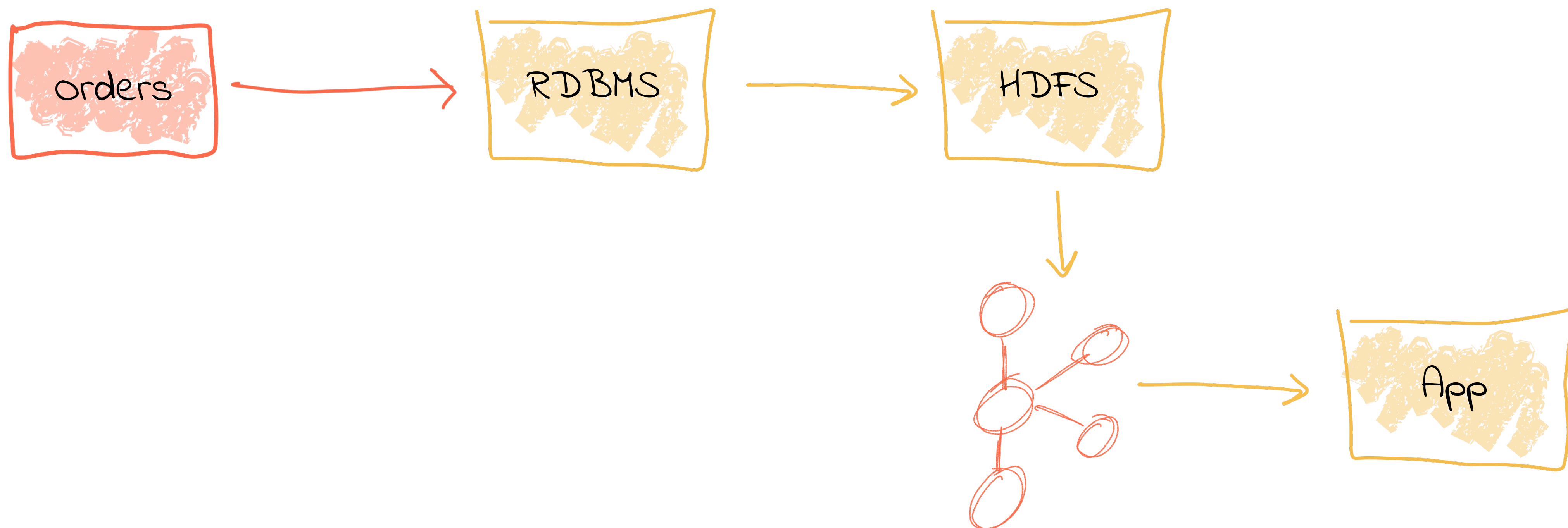


# Tight Coupling != Flexible



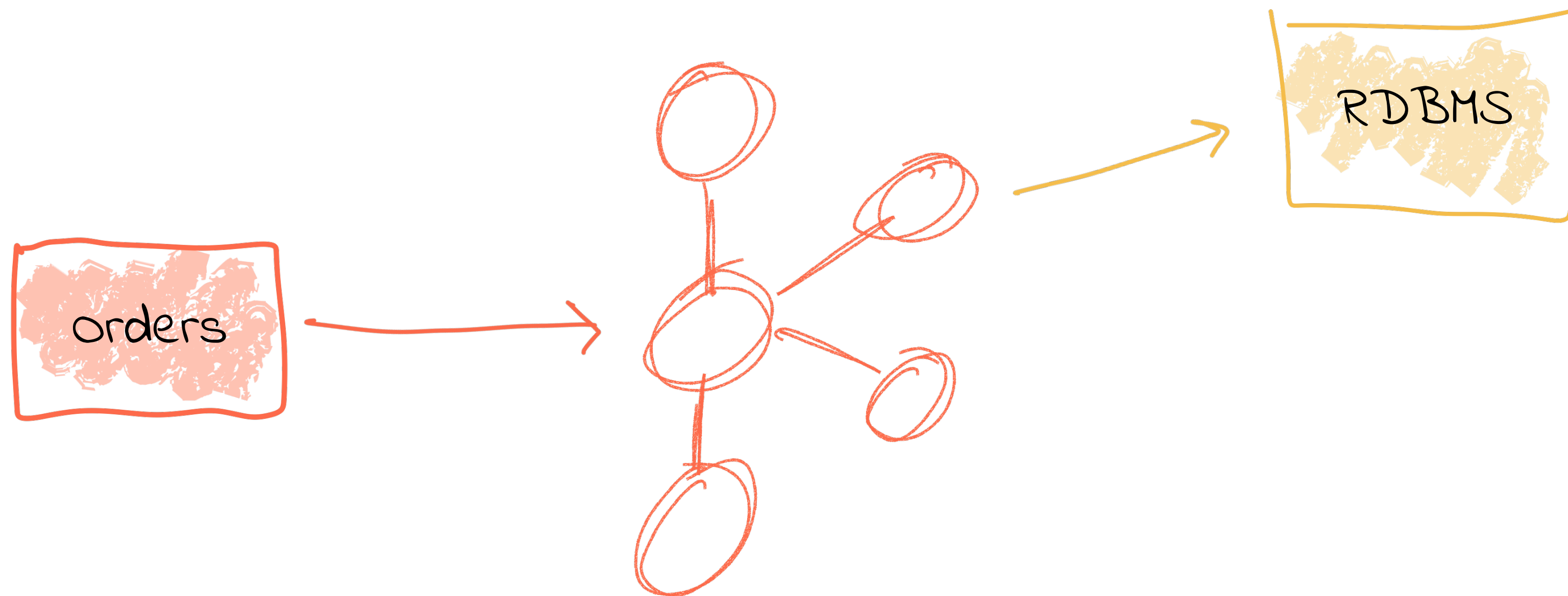


# Tight Coupling != Flexible



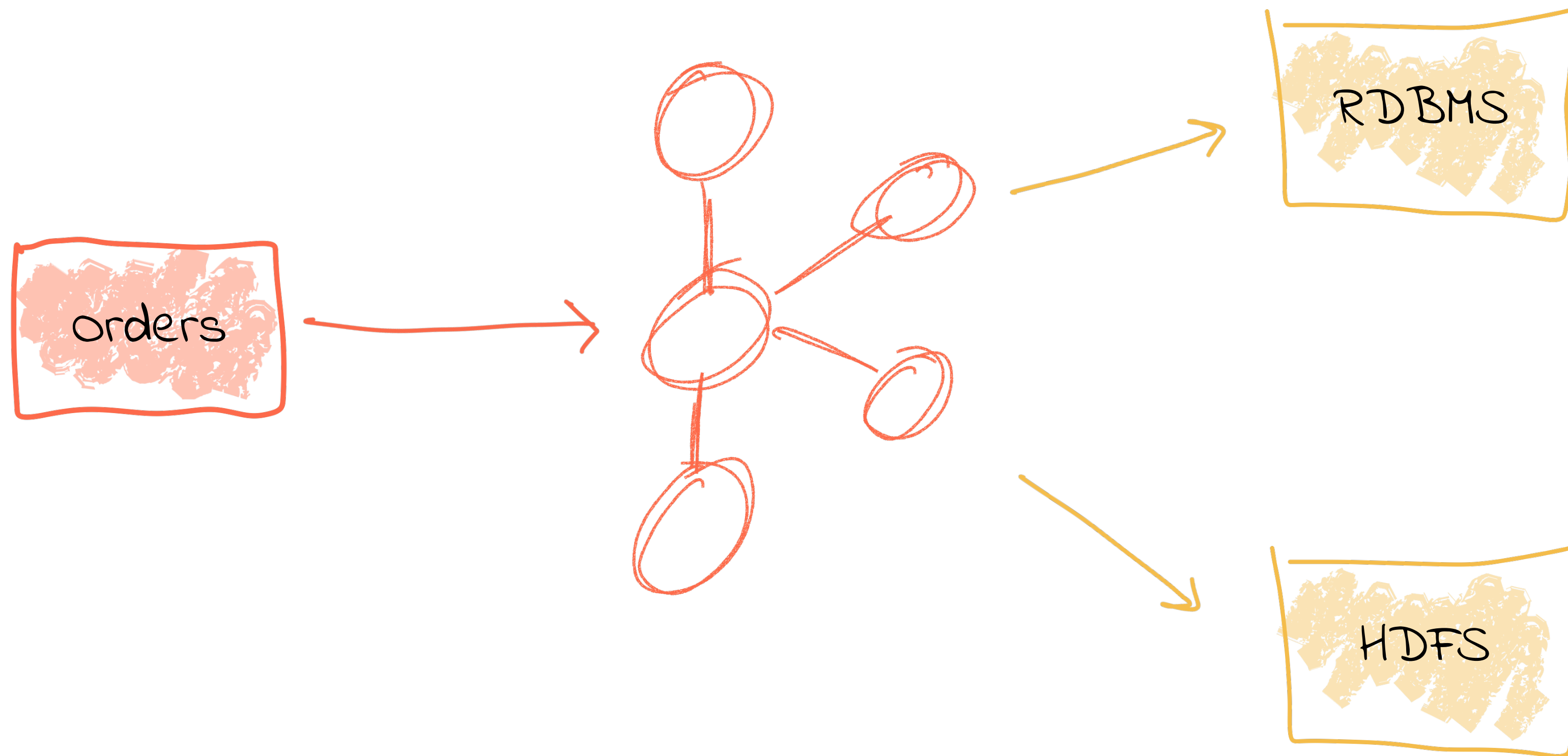


# Loose Coupling == Freedom to Evolve



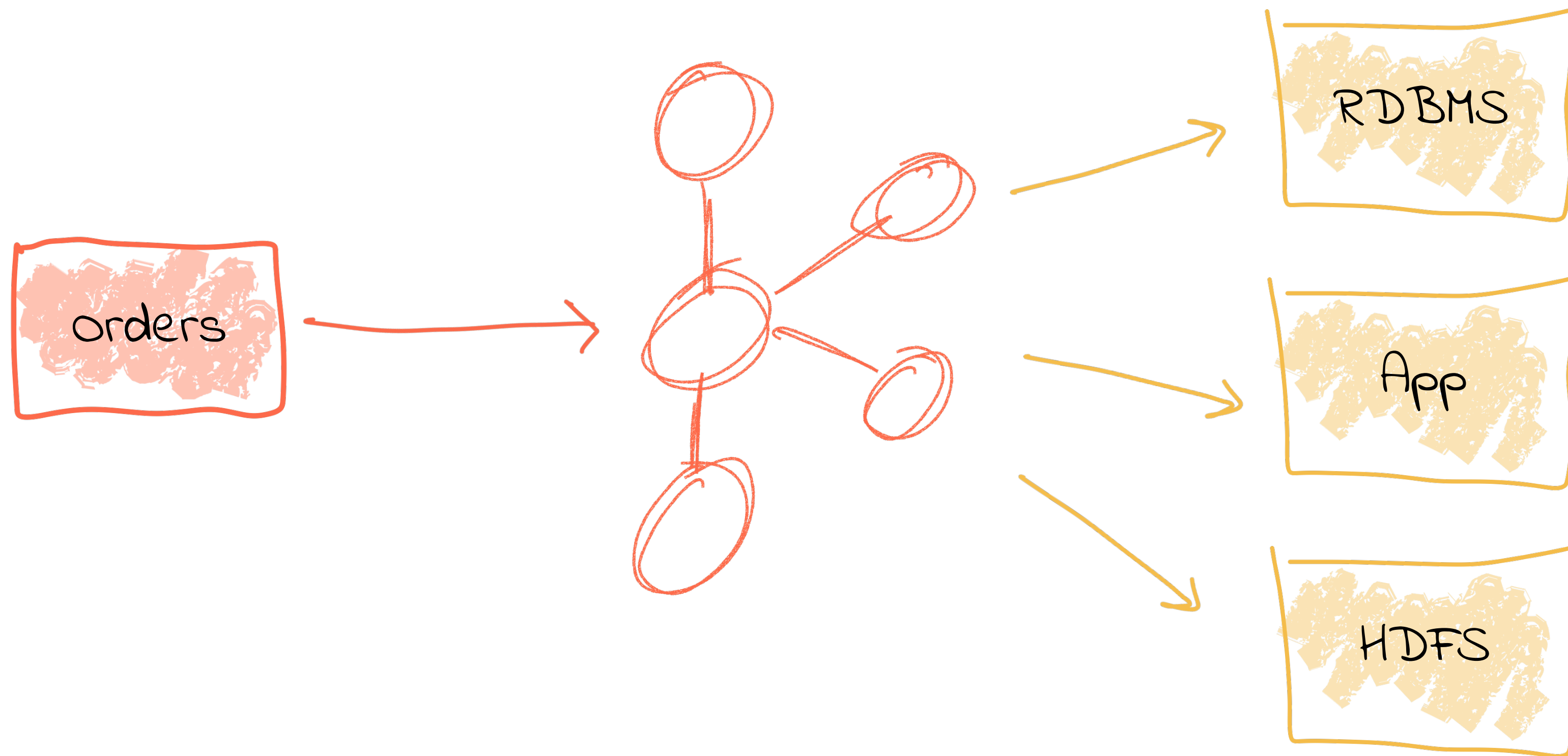


# Loose Coupling == Freedom to Evolve



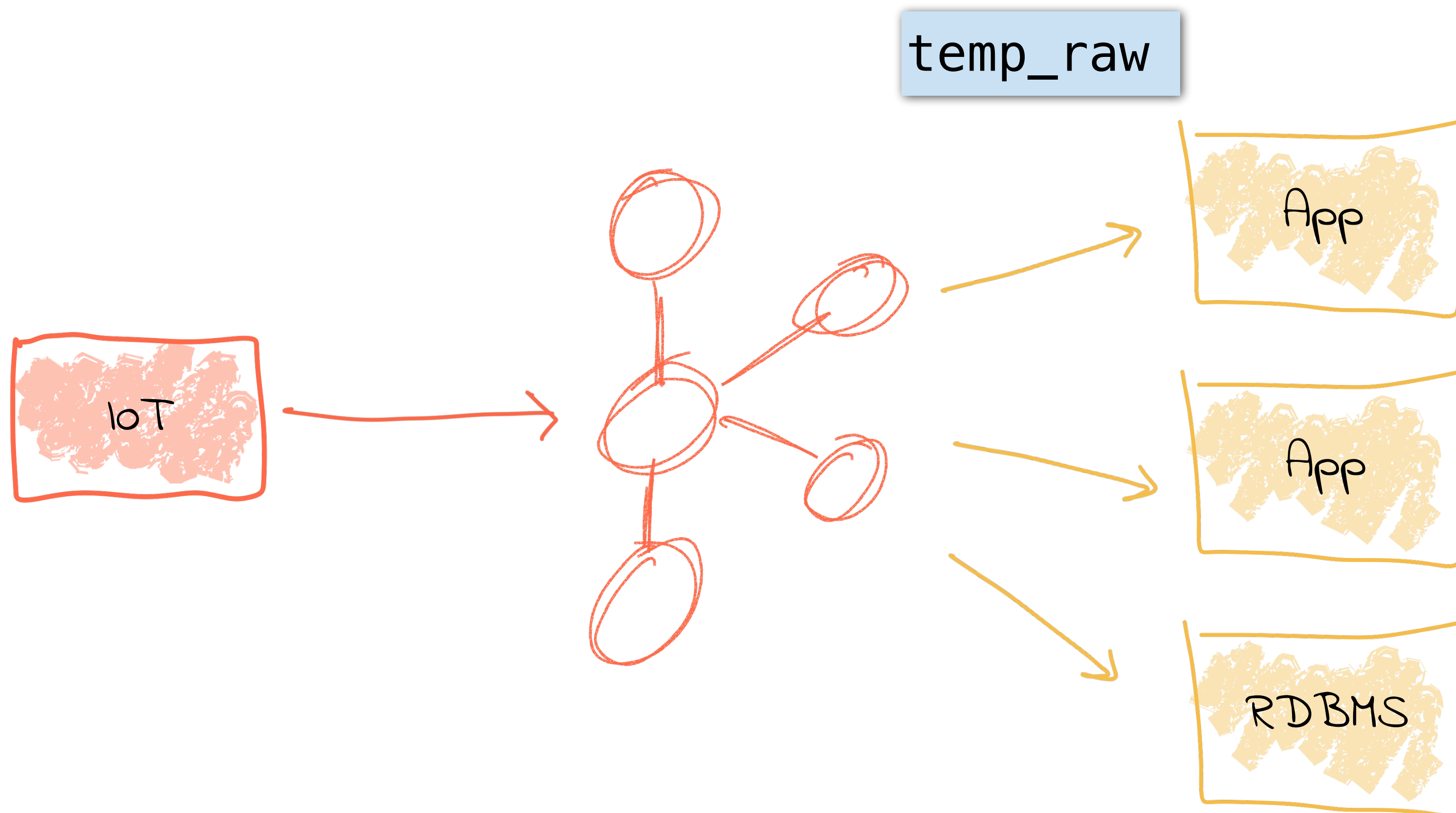


# Loose Coupling == Freedom to Evolve





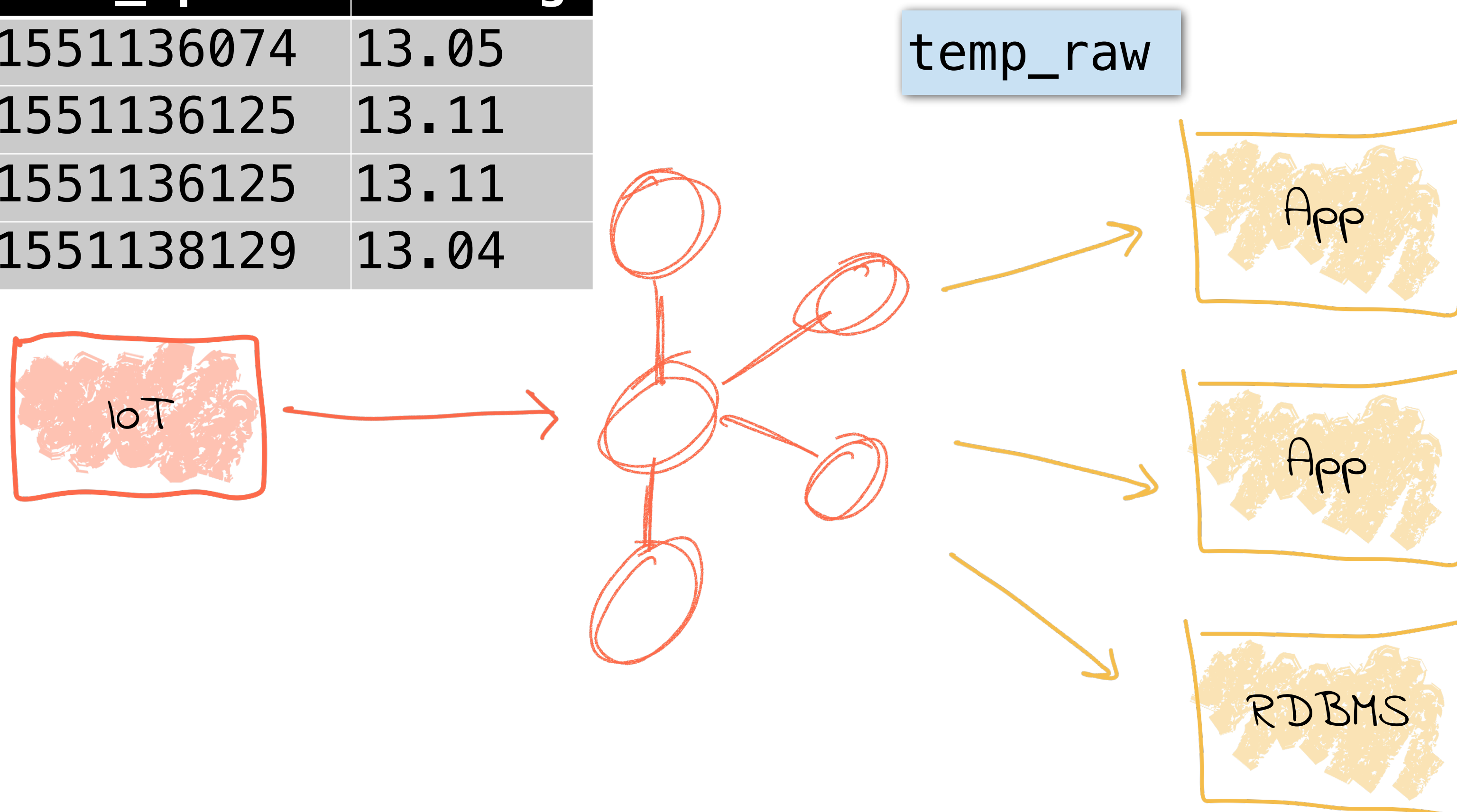
# Transform Once, Use Many: Data Cleansing





# Transform Once, Use Many: Data Cleansing

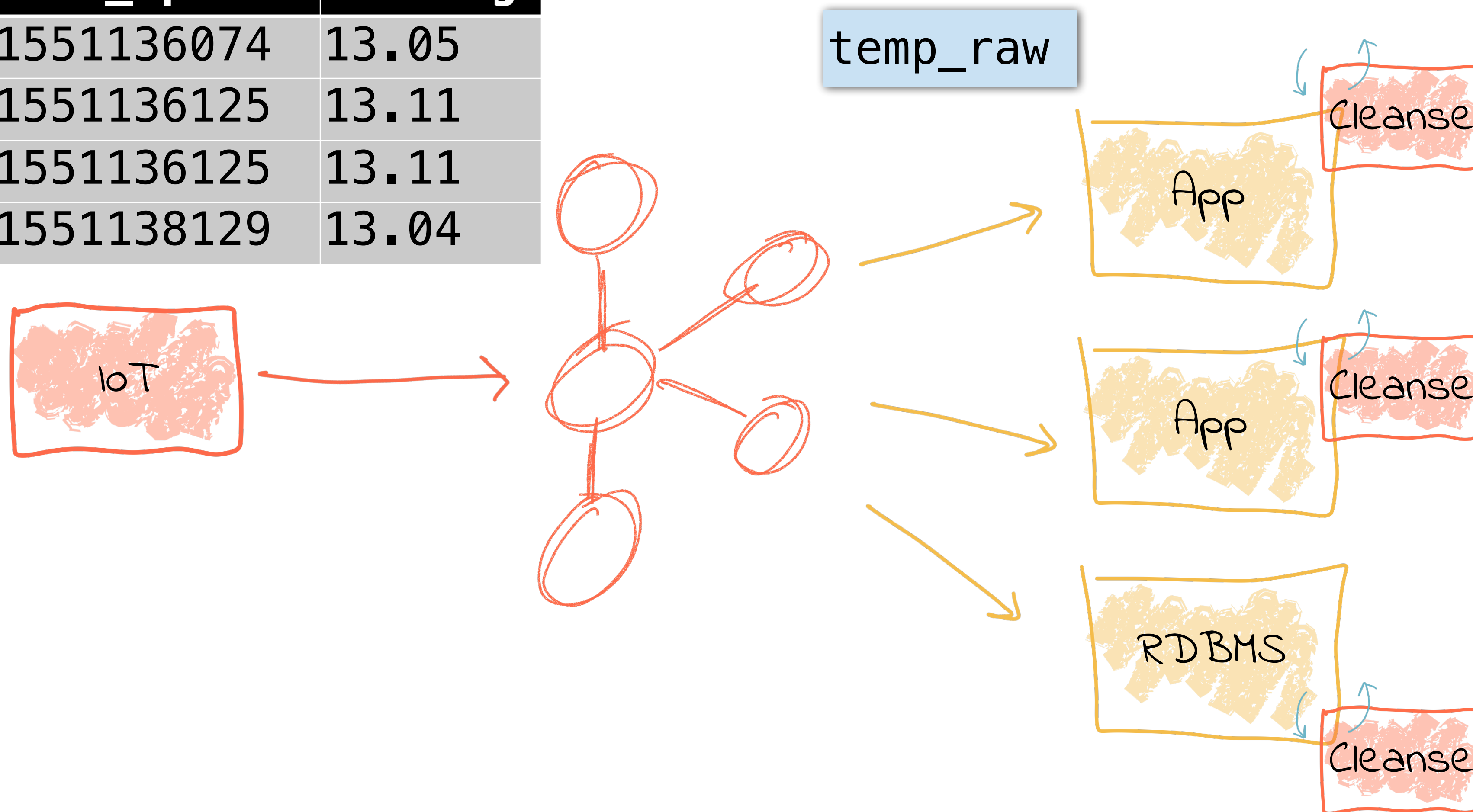
sensor_id	time_epoch	reading
42	1551136074	13.05
42	1551136125	13.11
	1551136125	13.11
42	1551138129	13.04





# Transform Once, Use Many: Data Cleansing

sensor_id	time_epoch	reading
42	1551136074	13.05
42	1551136125	13.11
	1551136125	13.11
42	1551138129	13.04



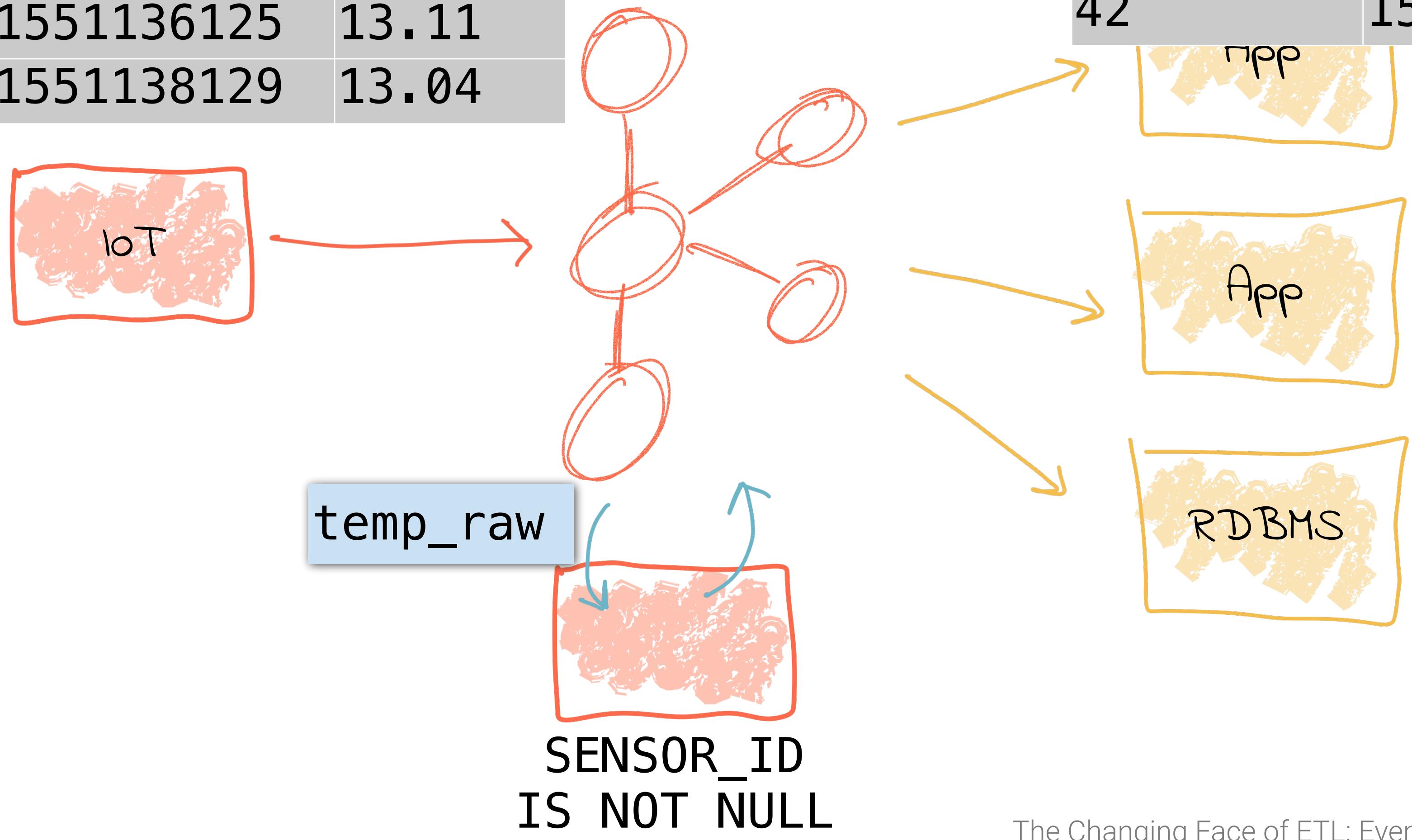


# Transform Once, Use Many: Data Cleansing

sensor_id	time_epoch	reading
42	1551136074	13.05
42	1551136125	13.11
	1551136125	13.11
42	1551138129	13.04

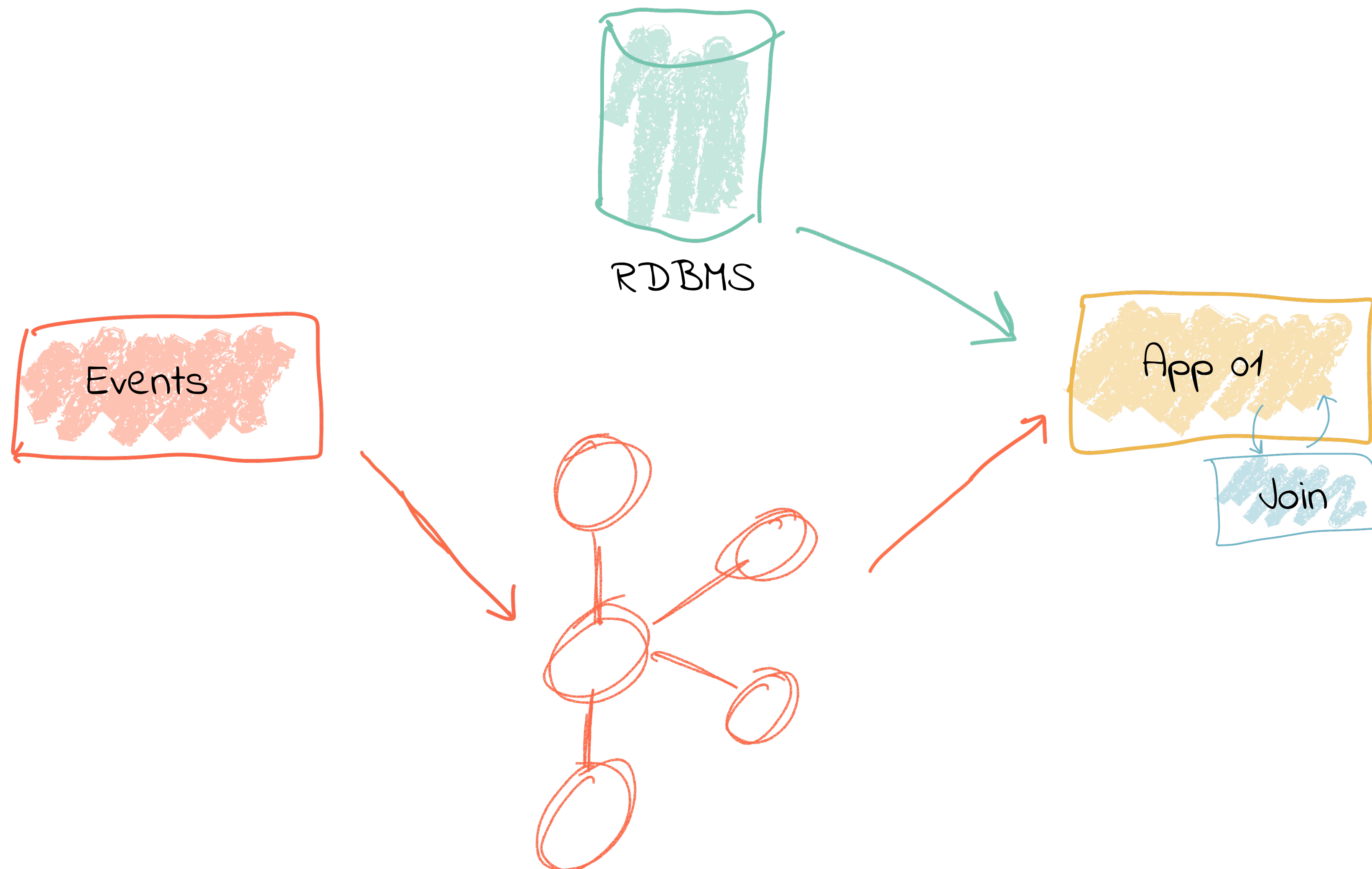
temp\_clean

sensor_id	time_epoch	reading
42	1551136074	13.05
42	1551136125	13.11
42	1551138129	13.04



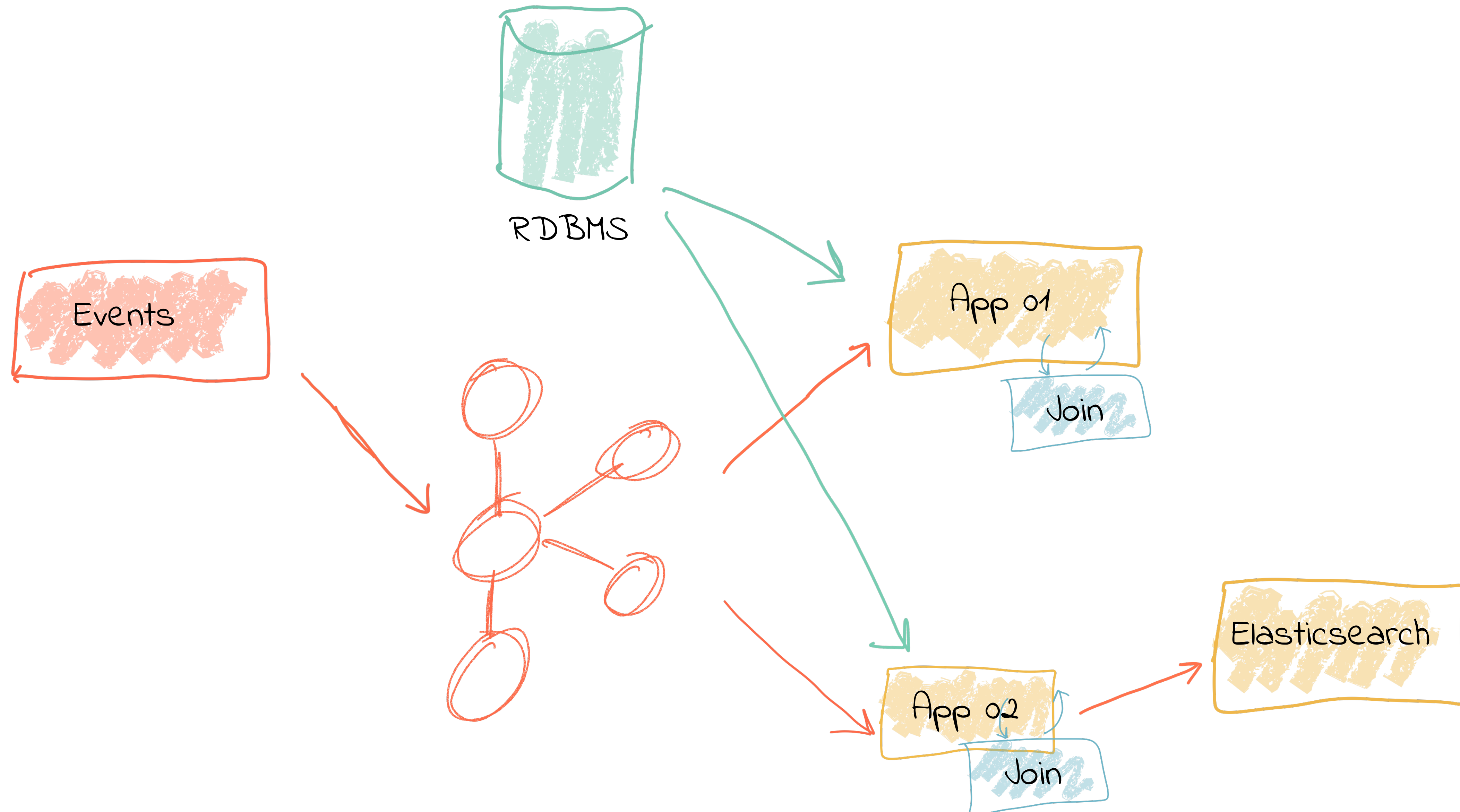


# Transform Once, Use Many: Data Enrichment



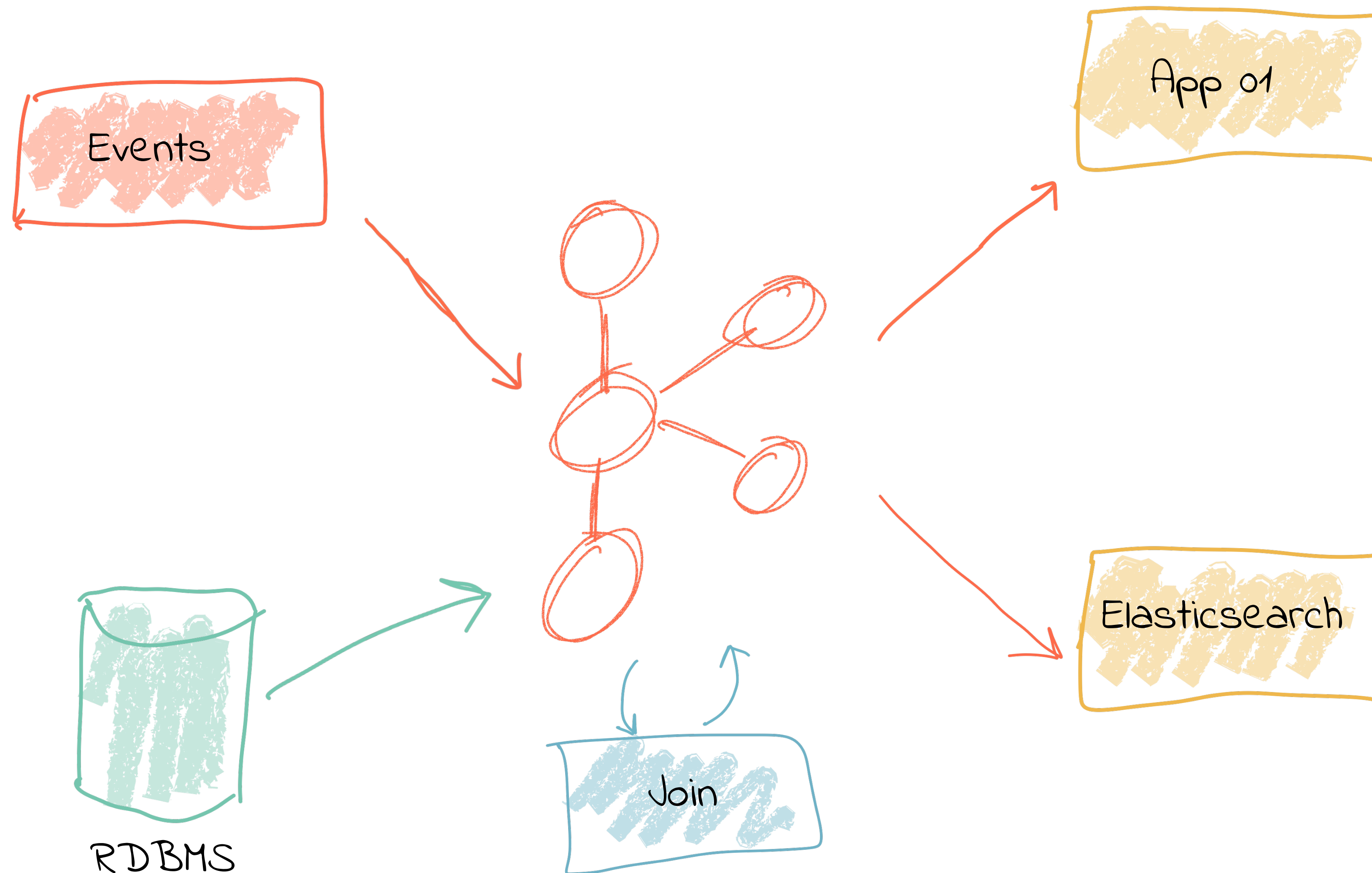


# Transform Once, Use Many: Data Enrichment





# Transform Once, Use Many: Data Enrichment





*Events model  
the real world*

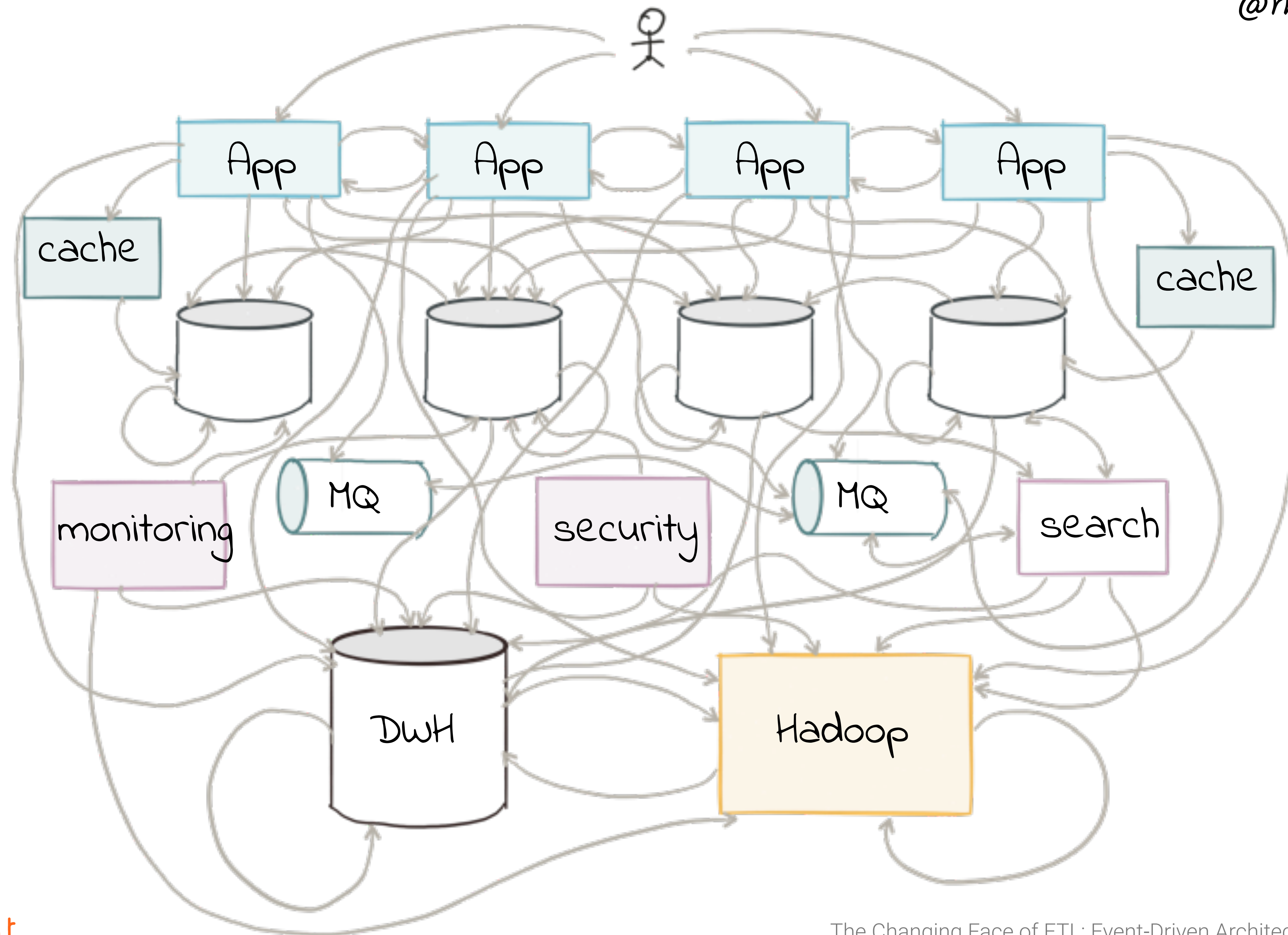




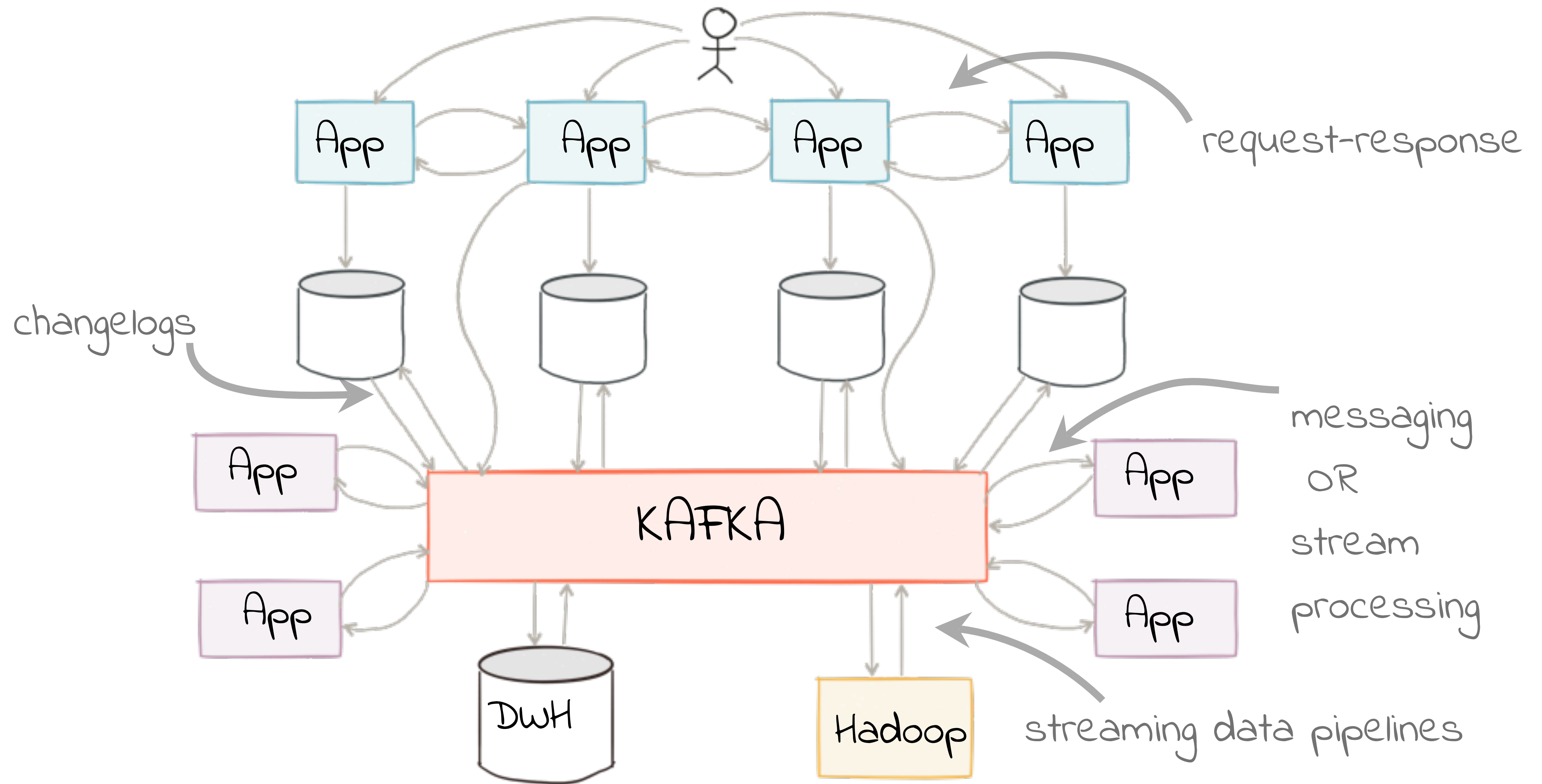
# Say NO to brittle pipelines













The background of the entire image is a photograph of the Golden Gate Bridge in San Francisco, taken during sunset or sunrise. The bridge's iconic red-orange towers and suspension cables are visible, with the sun low on the horizon behind the bridge, creating a warm, golden glow. The sky is a mix of soft pinks, oranges, and blues. The bridge spans across the water, with hills visible in the distance.

# NEXT STOP ***SAN FRANCISCO*** SEPT 30-OCT 1

**CONFLUENT COMMUNITY DISCOUNT CODE**

**KS19Meetup**

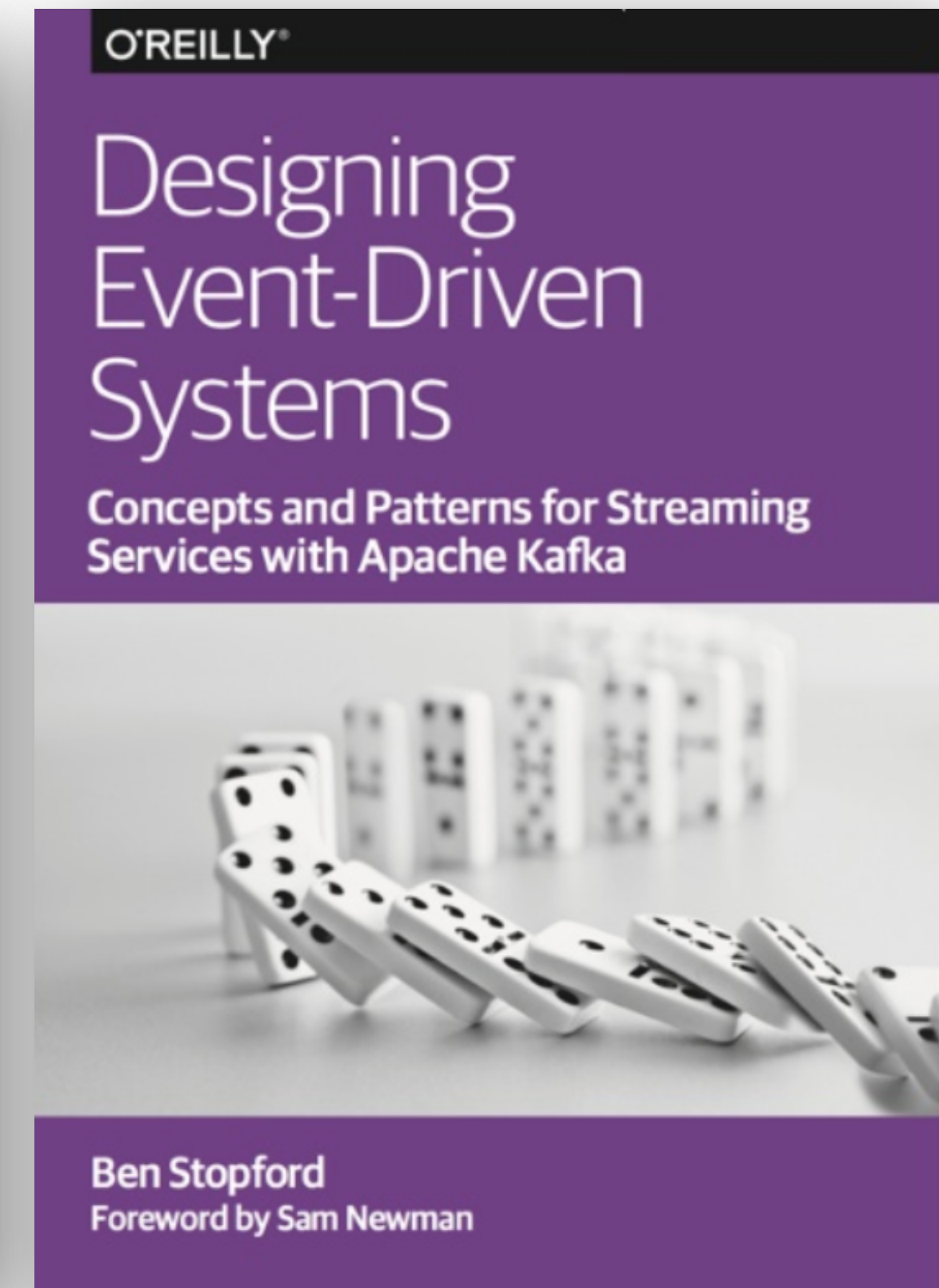
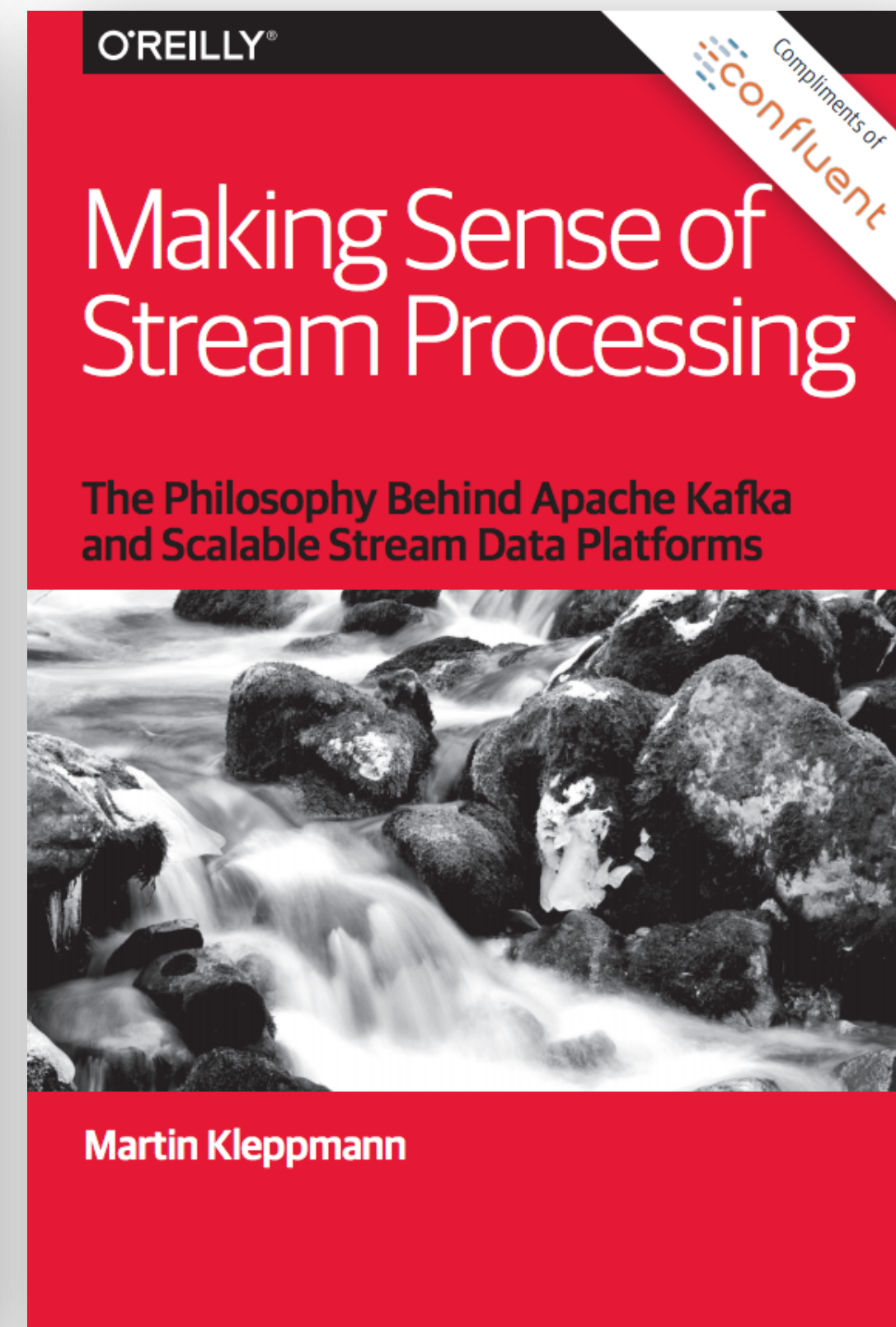
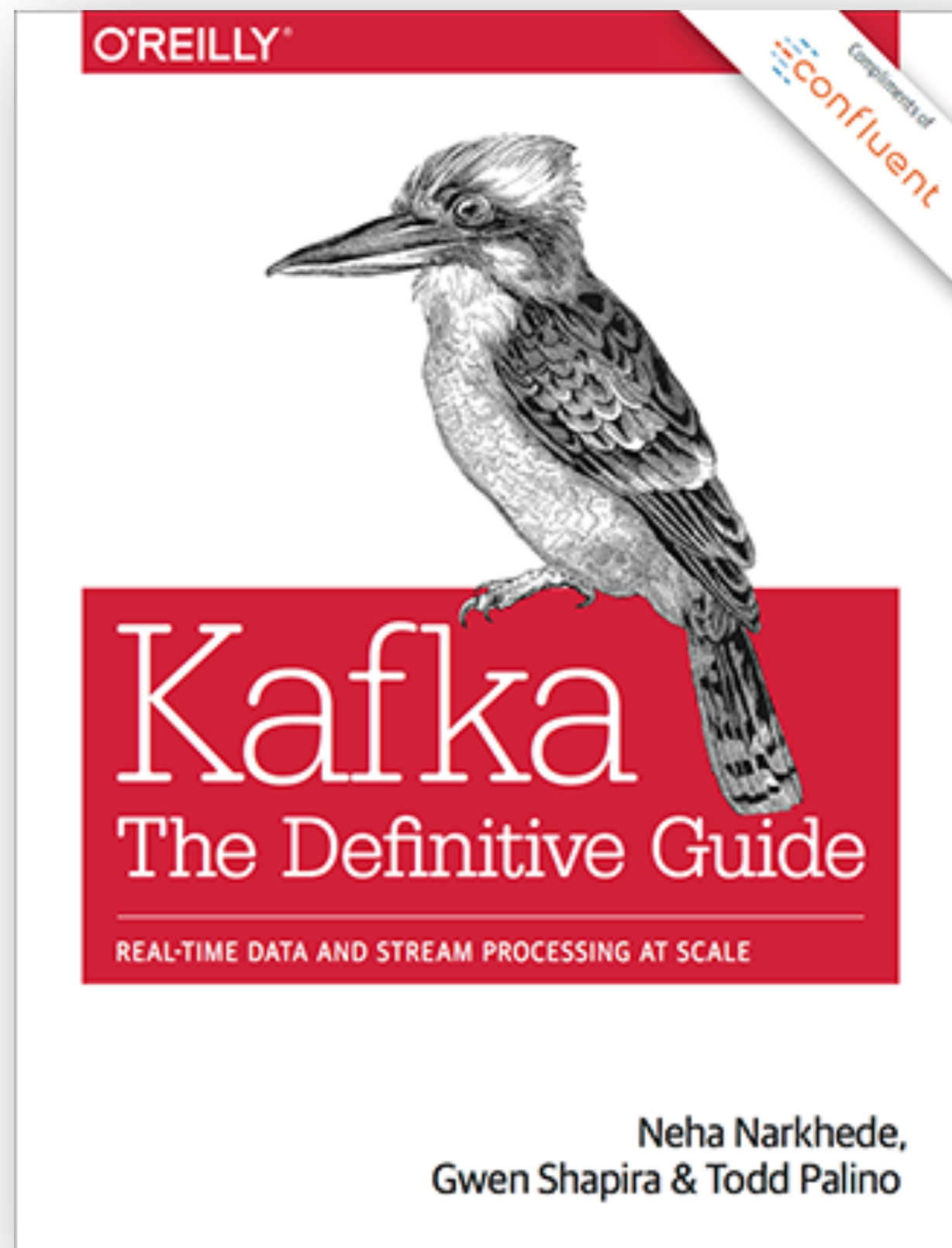
**25% OFF\***

kafka  
summit

\*Standard Priced Conference pass



# <http://cnfl.io/book-bundle>





# confluent.io/download

<http://cnfl.io/book-bundle>

<http://cnfl.io/slack>

@rmoff



**#EOF**