



# The Streaming Data Platform

## as a heavy duty enterprise data marketplace

when to use it in your project. The case of MMA

# Hello



Denis Golovachev



# Grid Dynamics



SPB



Saint-Petersburg





# Software Architect

# Data Streaming



# The Streaming Data Platform as a heavy duty enterprise data marketplace

## When to use it in your project. The case of MMA



# **The Streaming Data Platform**

as a heavy duty enterprise data marketplace

## **When to use it in your project.**

### **The case of MMA**



# **The Streaming Data Platform**

as a heavy duty enterprise data marketplace

## **When to use it in your project.**

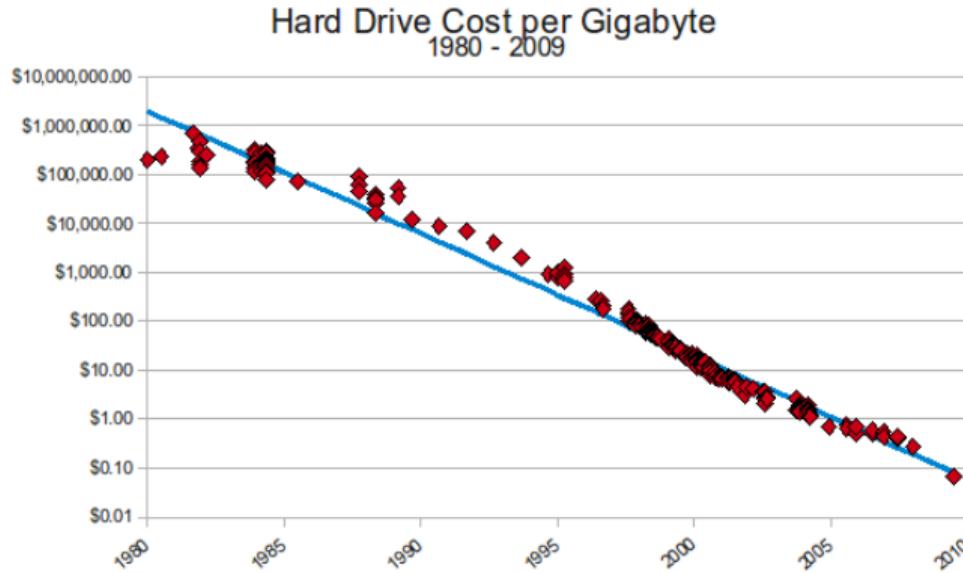
### The case of MMA

*MMA - Multichannel Marketing Automation*

## Small Introduction to



# Streaming Platforms



For the past 35+ years or so, hard drives prices are constantly dropping

## Backblaze Average Cost per Drive Size

By Quarter: Q1 2009 - Q2 2017

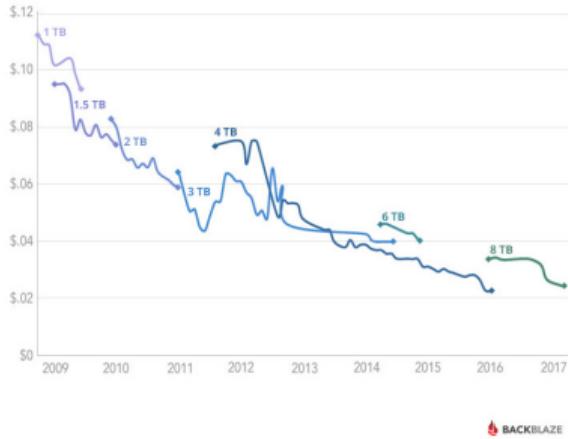


[<https://www.backblaze.com/blog/hard-drive-cost-per-gigabyte/>]

- 2010 year - countries could store everything. i.e. USA PRIZM
- 2015 - big companies could afford to store everything. i.e. Google
- 2019 - everyone!?

## Backblaze Average Cost per Drive Size

By Quarter: Q1 2009 - Q2 2017



[<https://www.backblaze.com/blog/hard-drive-cost-per-gigabyte/>]

So, the price is low, we store everything. [DRAFT] And i.e. if we're Facebook it's a tough question whether we should create a task for admins to create a script for housekeeping or just buy one more drive for storage.

# Information = Money

But could we earn **more** money with all of this information we collecting.



That's what basically we do in our BigData field



The world's most valuable resource is no longer oil, but data

# Knight

Do you know this company?

# Knight

The Knight Capital Group was an American global financial services firm engaging in market making, electronic execution, and institutional sales and trading.

With its high-frequency trading algorithms Knight was the largest trader in U.S.

# Knight

- BigData
- Lot of analytics

# Knight

On August 1st, 2012, Knight Capital deployed a new software update to their production servers. They switched it on and immediately they started losing literally \$10 million [£6.4m] a minute. And this went on for 45 minutes. At the end of it all they wound up having lost \$440 million.

000	00	SUNCITY	012	012	012	02	TECHIC	000	000	000	00	UMSH	000
000	00	SUNCON	000	000	000	00	TEXALA	110	101	103	93	UNZA	000
000	00	SUNCRM	000	000	000	00	TENCO	000	000	000	00	UPA	000
100	00	SUNRISE	111	111	111	01	TGL	000	000	000	00	UTAMA	000
100	00	SUNT-U	000	000	000	00	TGUAN	001	001	001	00	UTUSAH	000
000	00	SUNT-W	000	000	000	00	THGROUP	170	167	165	22	VSIIND	000
000	00	SUNTECH	000	000	000	01	TMHIN	000	000	000	00	WATTA	000
000	00	SUREMAX	000	000	000	00	TIDWAH	120	120	120	02	WCT	000
000	00	SURIA	01Y	017	013	012	TIMMELL	155	150	152	93	WEMBLEY	000
100	00	SYSTEM	000	001	001	00	THTP	000	000	000	00	WIETEC	000
000	00	T CORP	000	000	000	00	THTT-W	000	000	000	00	WONG	110
099	04	TIDECEAN	000	000	000	00	TONYPAK	000	000	000	00	WOODLAH	000
200	011	T STORE	000	000	000	00	TOYOCEN	000	000	000	00	WOVENTH	000
000	00	T.CAP	010	010	010	01	TRANCAP	000	000	000	00	WATER	000
000	00	T.CAP-W	000	000	000	00	TRU-W	000	000	000	00	WCABLE	000
000	00	TAJO	053	050	051	03	TRUTECH	000	000	000	00	YAHORG	000
000	01	TAJO-W	000	000	000	00	TSB	210	210	210	01	YCS	000
000	00	TAKAFUL	000	000	000	00	TSB-W	000	000	000	00	YCS-W	000
000	00	TAKASO	173	172	172	10	TSUPER	000	000	000	00	YECIU	000
000	00	TAHADAH	000	000	000	00	TTRES	000	000	000	00	YEE LEE	000
036	26	TAN-TAN	000	000	000	00	U-WOOD	000	000	000	00	YINSON	000
015	20	TAP	000	000	000	00	UBB	000	000	000	00	YLI	000
000	00	TAS	000	000	000	00	UCPRES	000	000	000	00	YOKO	000
125	02	TCL	000	000	000	00	UH DOVE	000	000	000	00	YONGTAI	000
000	00	TECGUAN	113	113	113	01	ULBON	000	000	000	00	ZOUR	000
000	01	TECHVEN	000	000	000	00	ULBON	000	010	010	00	ZYNG	000

**Humans still watch the systems, but the computers move far too quickly for us to react to everything they do.**



# My previous project

OK Sir...  
Here you go...

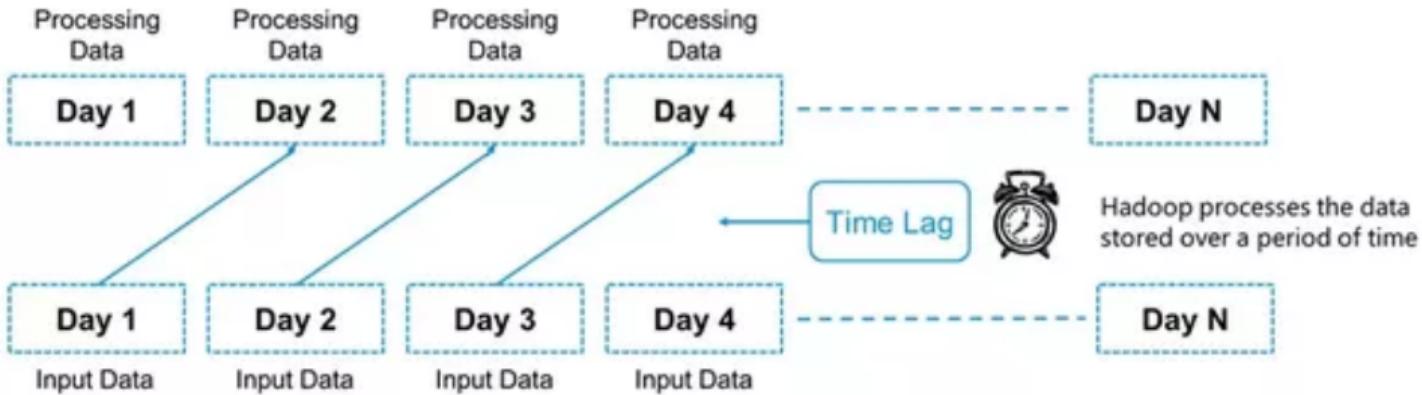
...but don't you know  
you can do all this  
much more easily  
online.



## Fraud Detection System

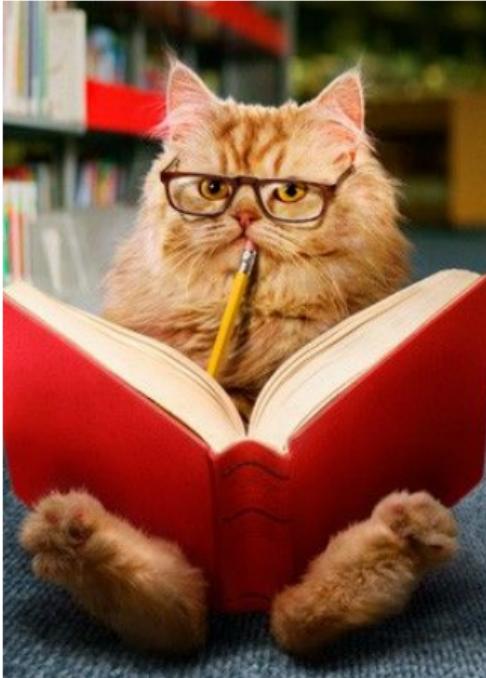


## Processing Data Using MapReduce

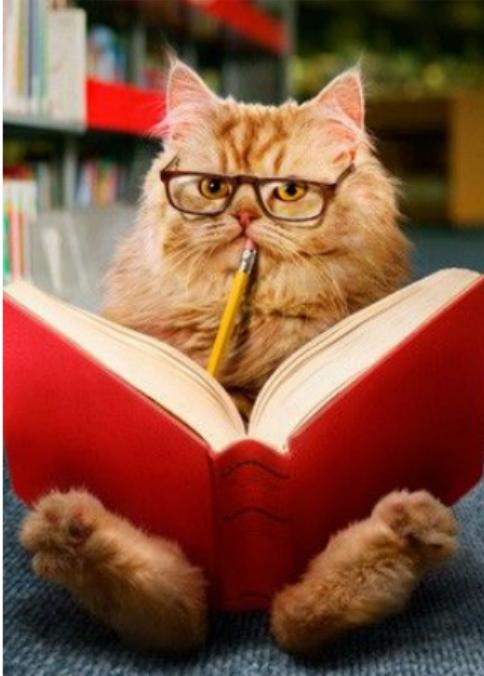


- Collect - 24hrs
- Process - 6hrs
- Block Fraudsters - 10 minutes

# We're loosing money!



# We're loosing money!



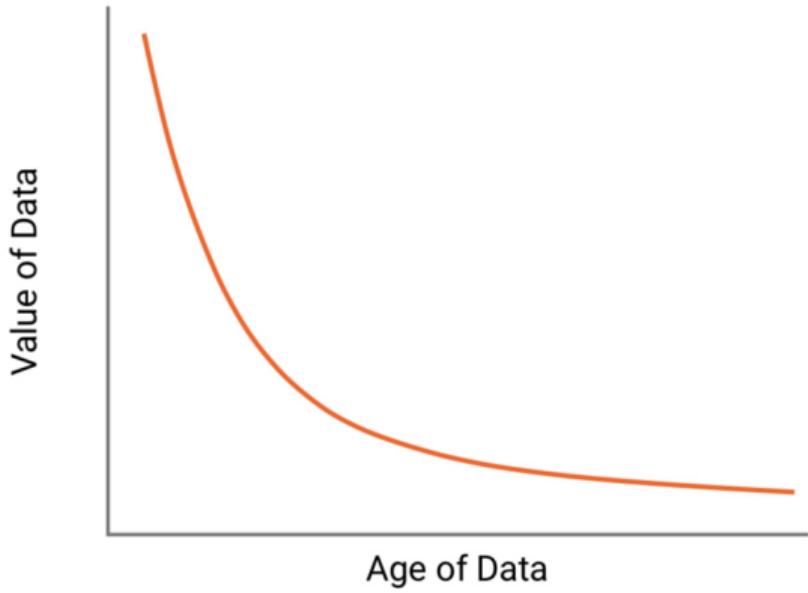
30hr lag → 70 000\$ per day

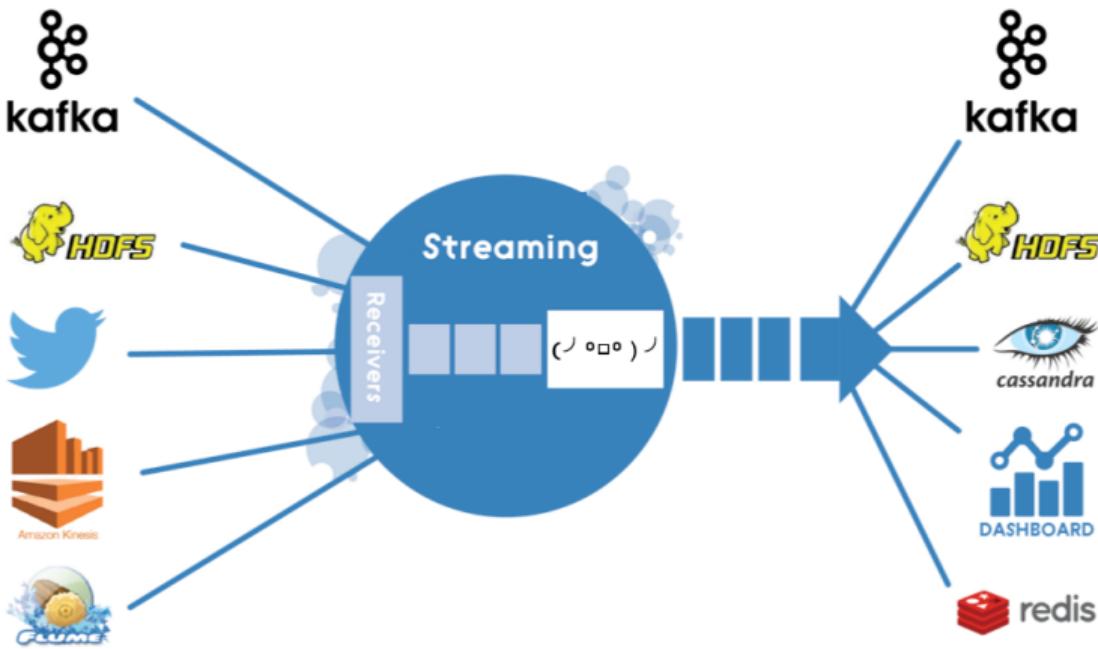
# What could be improved here?

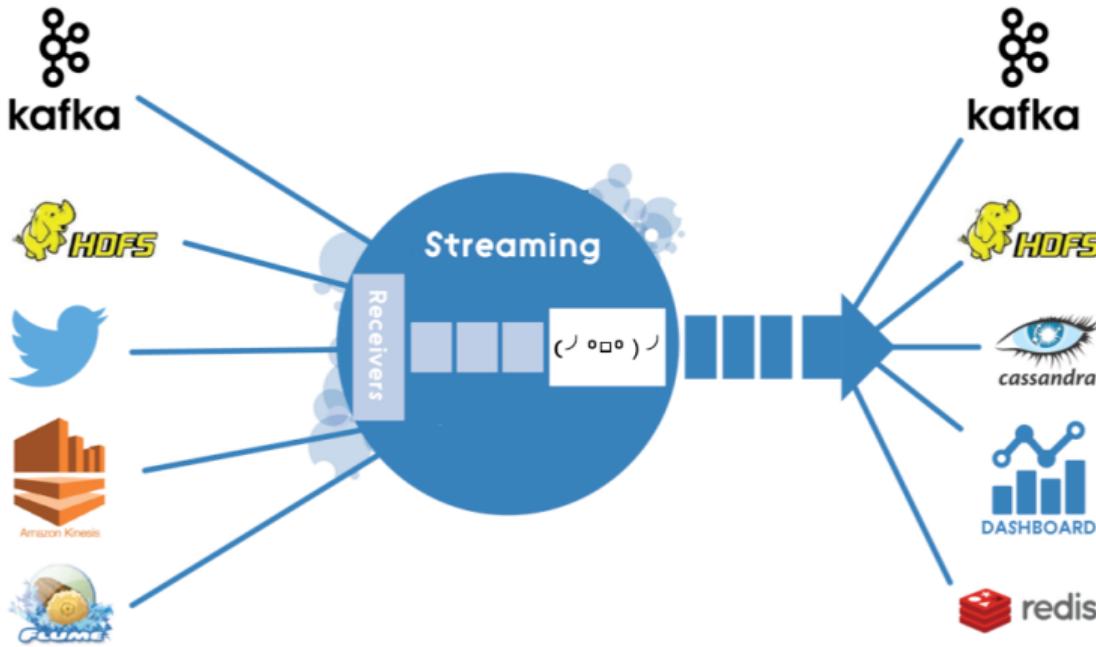


# Reaction Time!









30hr → 10mins  
And Fraudsters were Disappointed

# Knight



Was not so lucky

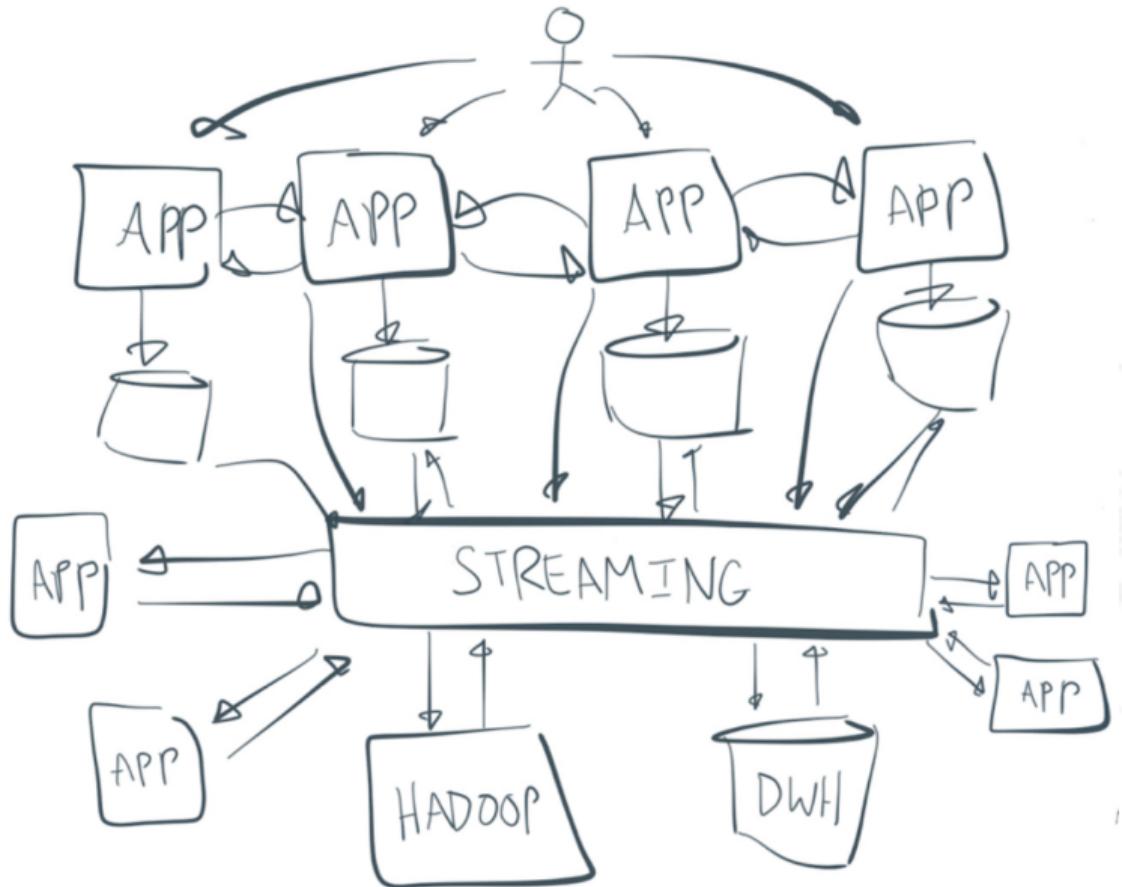
# GOTTA GO FAST!



- Reduce reaction time
- Minimize risk surface
- Compete for best offers in market
- Give out customers what they need right in time
- ...

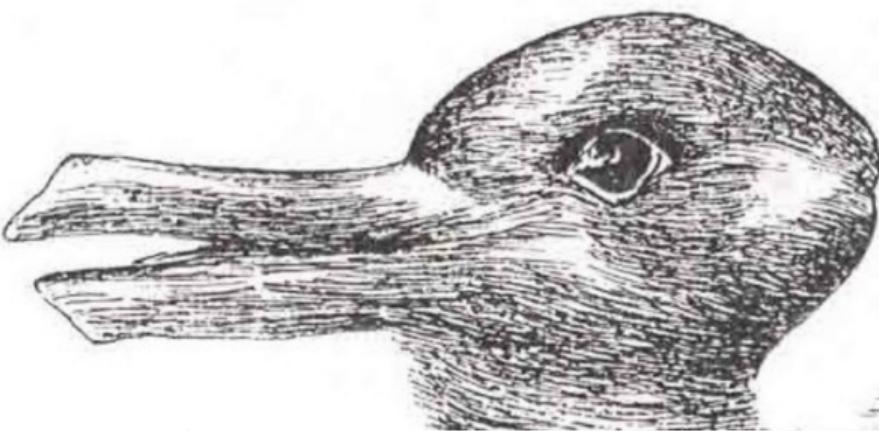


# Stream- ing Data Platform

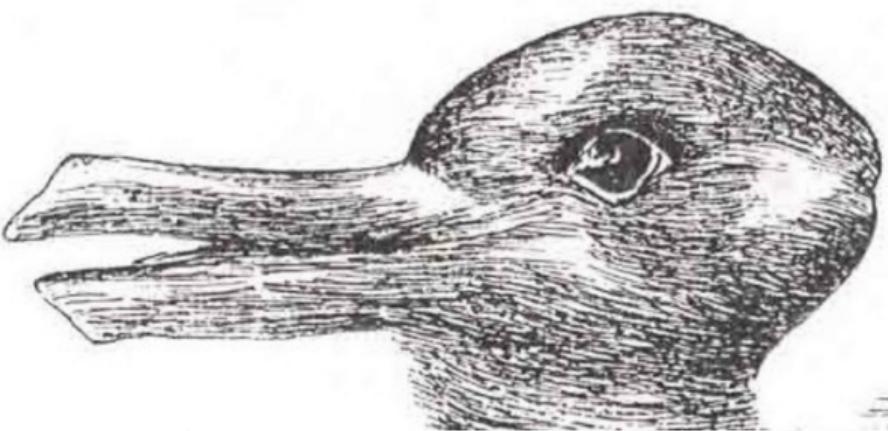


# Streaming

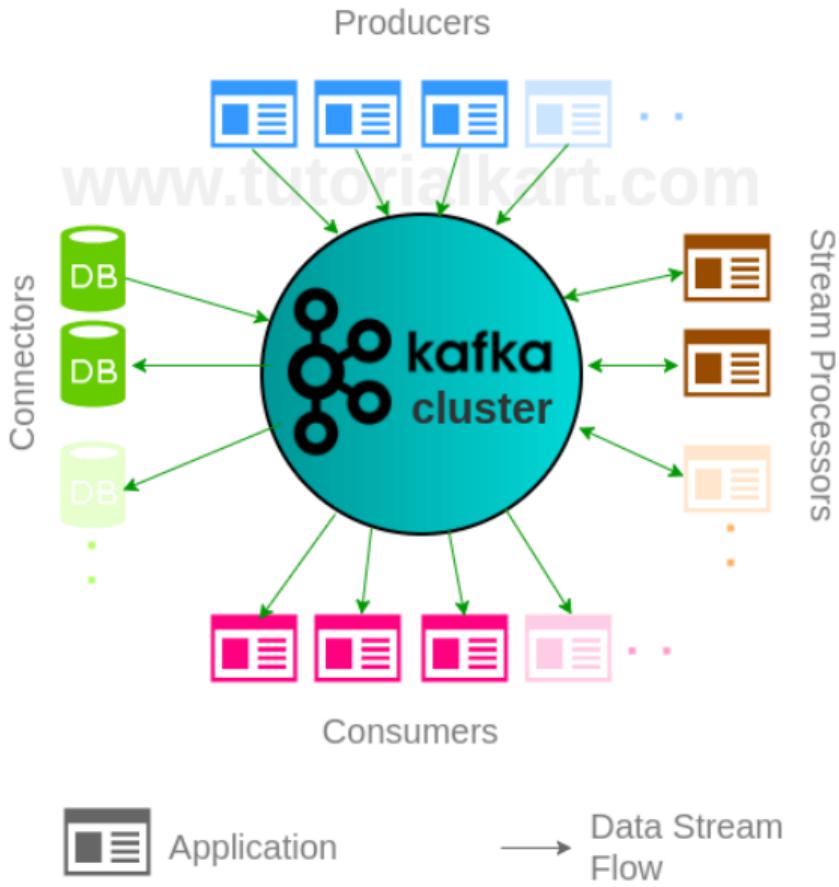
## Data Platform

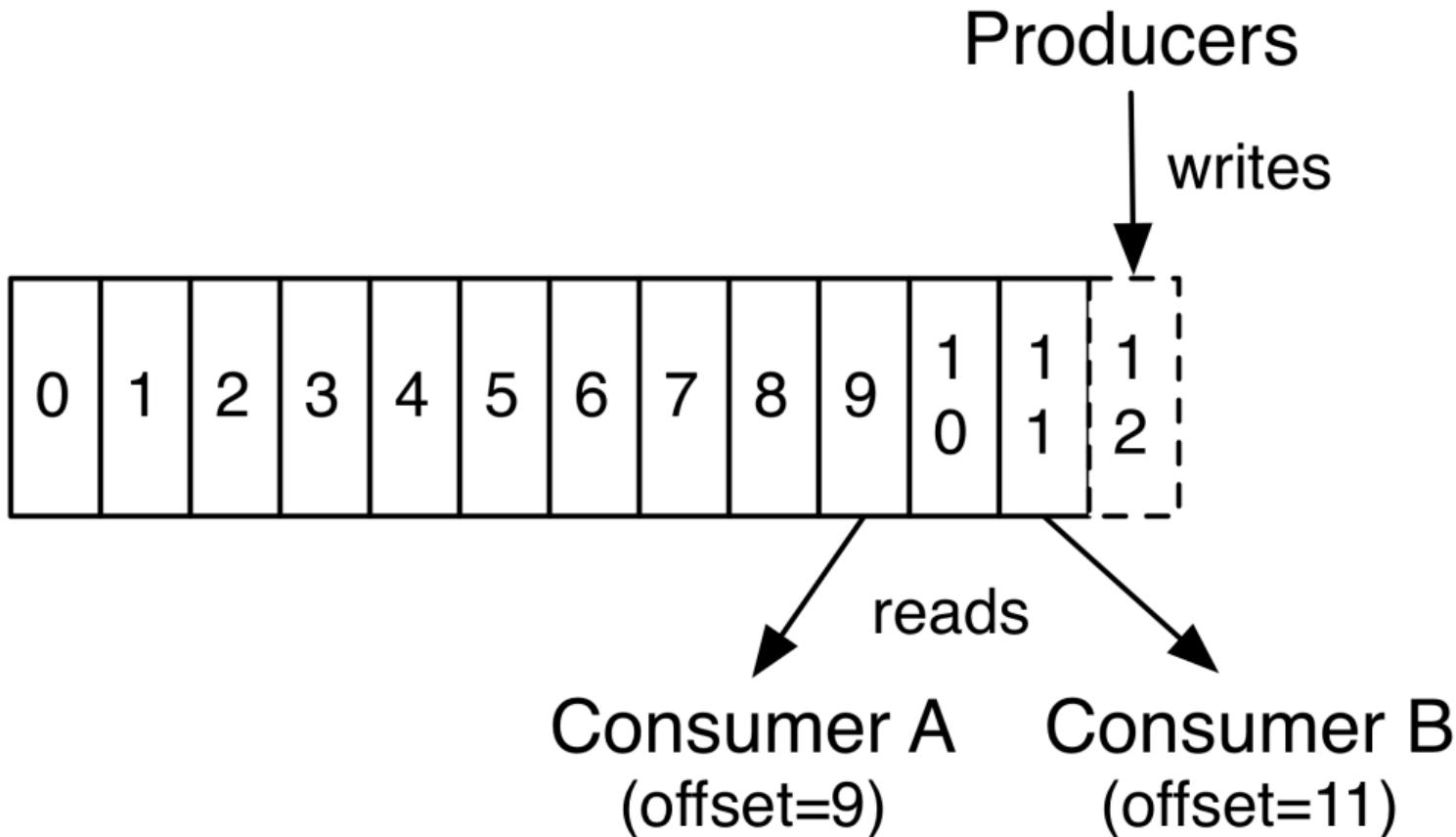


**Yet another ESB!?**

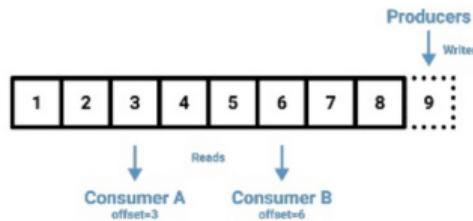


**Yet another ESB!?**





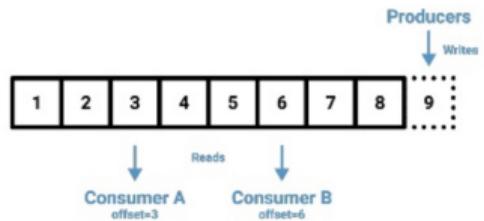
## Event Streaming Paradigm



- Don't route or just processing, it's storage and processing.

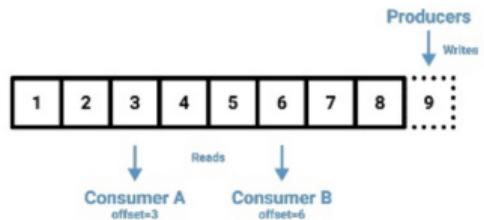


## Event Streaming Paradigm



- Don't route or just processing, it's storage and processing.
  - Allows downtime for maintenance

## Event Streaming Paradigm



- Don't route or just processing, it's storage and processing.
- Allows downtime for maintenance
- Backpressure
- Built-in scalability
- Dumb pipes, smart endpoints



Informatica



They are friends!  
Kafka is scalable, reliable, **but**:

- Integration with legacy



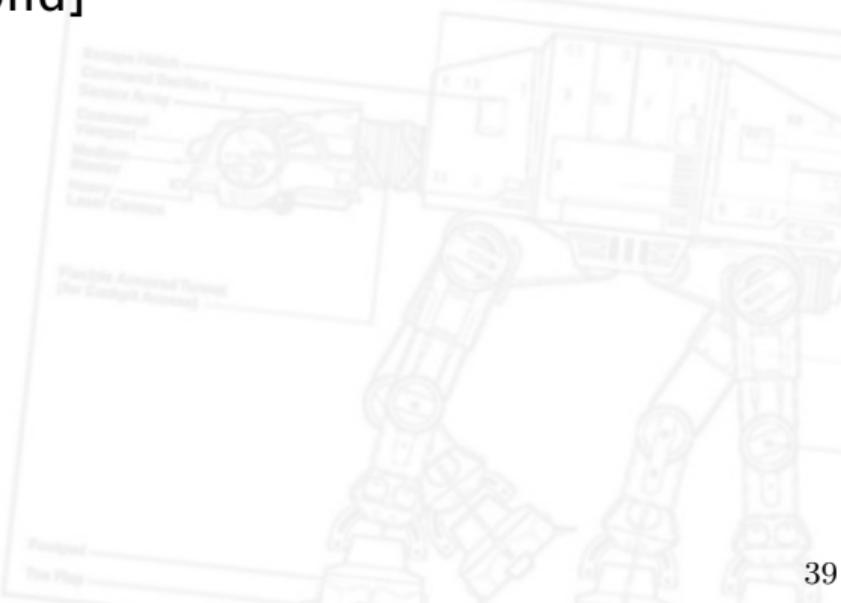
Informatica



ActiveMQ

They are friends!  
Kafka is scalable, reliable, **but**:

- Integration with legacy
- Fast message delivery [Less than a second]





Informatica



solace

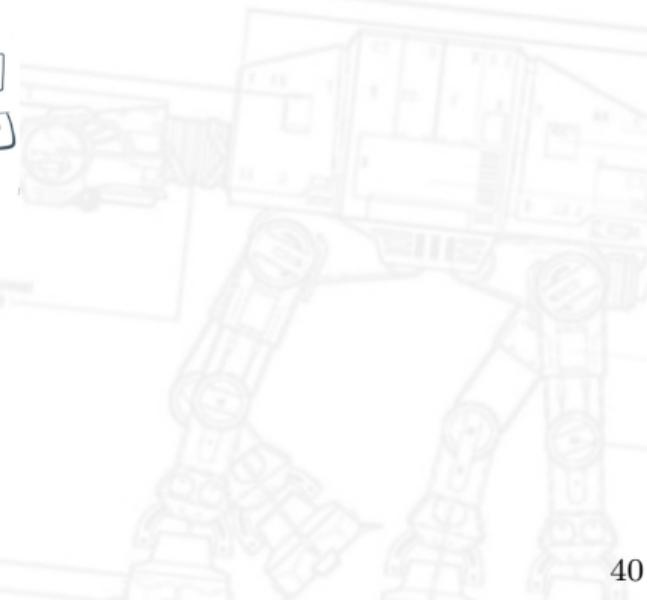
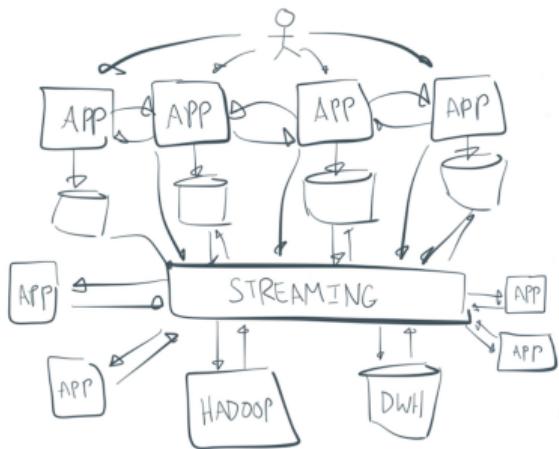
RabbitMQ

ActiveMQ

They are friends!

Kafka is scalable, reliable, **but**:

- Integration with legacy
- Fast message delivery [Less than a second]
- Synchronous Point to Point messaging
- Complex Routing
- Default protocol is proprietary



## References:

- [https://en.wikipedia.org/wiki/Knight\\_Capital\\_Group](https://en.wikipedia.org/wiki/Knight_Capital_Group)
- <https://www.bbc.com/news/magazine-19214294>
- <https://www.backblaze.com/blog/hard-drive-cost-per-gigabyte/>