



# 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

# Hello



Roman Zakharov [I don't have his photo, but this is quite close]



SPB



Saint-Petersburg

SPB



Krakov





# Software Architect







# Business Analytic

# 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

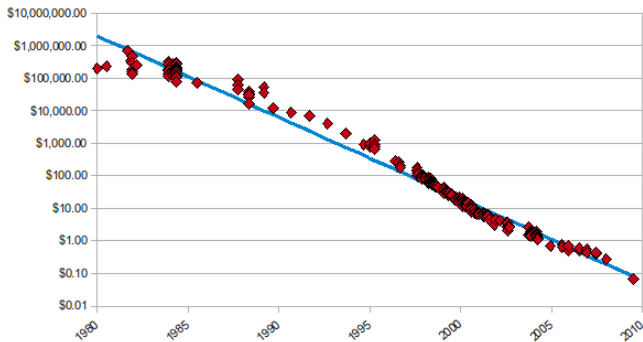
*MMA - Multichannel Marketing Automation*

# Small Introduction to



## Streaming Platforms

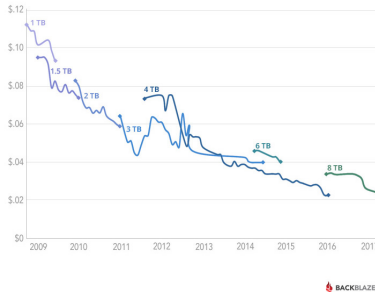
## Hard Drive Cost per Gigabyte 1980 - 2009



For the past 35+ years or so, hard drives prices have dropped, from around \$500,000 per gigabyte in 1981 to less than \$0.02 per gigabyte today

## 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 afford to 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

# 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	052	052	052	02	TECHIC	100	100	100	00	UMSH	000
000	00	SUNCON	000	000	000	00	TEKALA	110	103	103	93	UNZA	000
000	00	SUNCRN	000	000	000	00	TEHCO	000	000	000	00	UPA	000
100	00	SUNRISE	144	144	144	01	TGL	000	000	000	00	UTAMA	000
140	00	SUNT-U	000	000	000	00	TGUAN	007	007	007	00	UTUSAN	000
000	00	SUNT-W	000	000	000	00	THGROUP	110	107	109	22	VSIND	000
100	00	SUNTECH	100	100	100	01	THIN	000	000	000	00	WATTA	000
000	00	SUREMAX	000	000	000	00	TIENHAW	120	120	120	02	WCT	000
000	00	SURIA	034	034	034	0100	TIMMELL	155	150	152	99	WEMBLEY	000
100	00	SYSTEM	000	000	000	00	TINT	000	000	000	00	WIDETEC	000
000	00	T CORP	100	100	100	00	TINT-W	000	000	000	00	WONG	110
000	04	T OCEAN	000	000	000	00	TONYPAK	000	000	000	00	WOODLAN	000
144	0100	T STORE	000	000	000	00	TOYOCEH	000	000	000	00	WOVENTX	000
000	00	T.CAP	010	010	010	11	TRANCAP	000	000	000	00	WATER	000
000	00	T.CAP-W	000	000	000	00	TRU-W	000	000	000	00	WACABLE	000
000	00	TAJO	053	050	051	03	TRUTECH	000	000	000	00	YAHONG	000
101	01	TAJO-W	000	000	000	00	TSH	210	210	210	01	YCS	000
000	00	TAKAFUL	000	000	000	00	TSH-W	000	000	000	00	YCS-W	000
000	00	TAKASO	113	112	112	10	TSUPER	000	000	000	00	YECHIU	000
000	00	TAMADAN	000	000	000	00	TTRES	000	000	000	00	YEE LEE	000
036	36	TAN&TAN	000	000	000	00	U-WOOD	000	000	000	00	YINSON	000
015	20	TAP	100	100	100	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	UN DOVE	000	000	000	00	YONGTAI	000
000	00	TEGUAN	113	113	113	01	ULSON	000	000	000	00	YONGTAI	000
000	01	TECHVEN	000	000	000	00						YONGTAI	000

Humans still watch the systems, but **the computers move far too quickly for us to react to everything they do** - and at Knight Capital, the computer glitch meant the company was making trades it didn't intend to make. That's how to lose almost half a billion dollars in a little over half an hour.





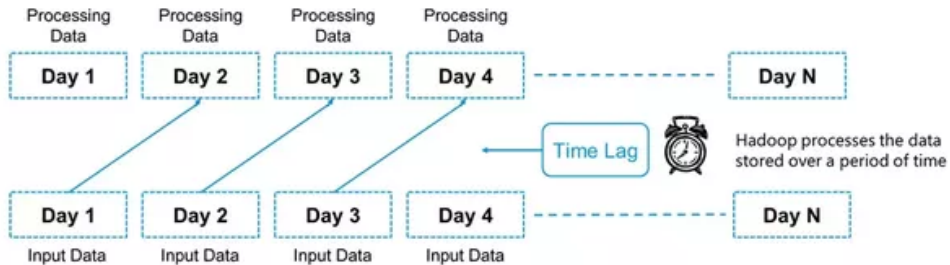
My previous project



# Fraud Detection System

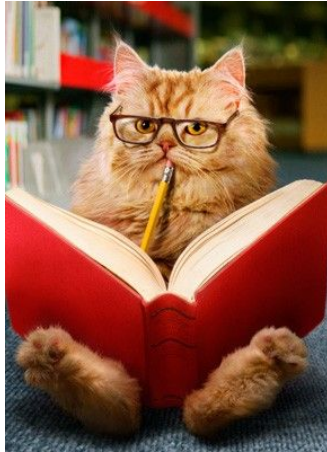


## Processing Data Using MapReduce

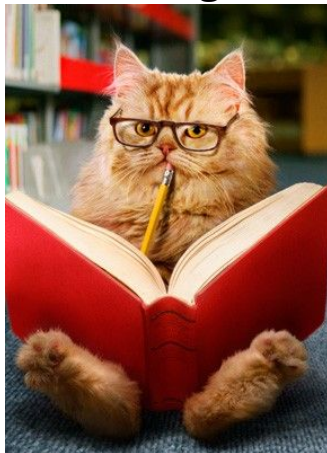


- Collect - 24hrs
- Process - 6hrs
- Block Fraudsters - 10 minutes

# We're loosing money!



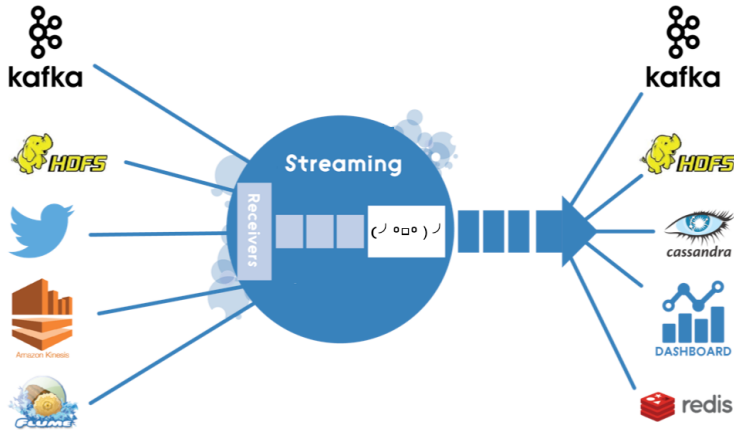
We're loosing money!



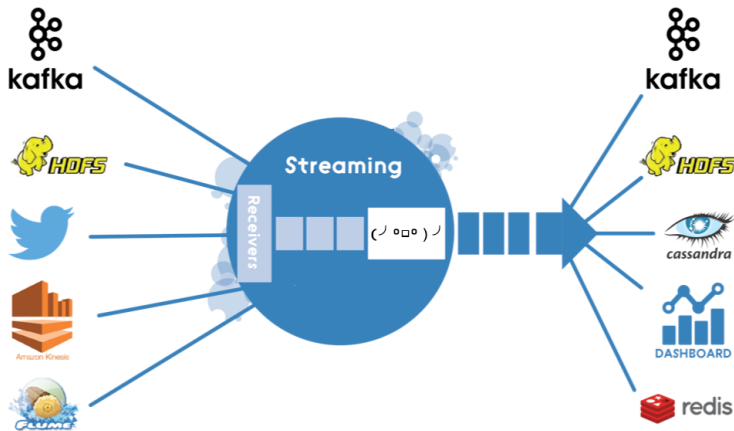
30hr lag  $\rightarrow$  70 000\$ per day

# What could be improved here?

# Reaction Time!







30hr → 10mins  
And Fraudsters were Disappointed

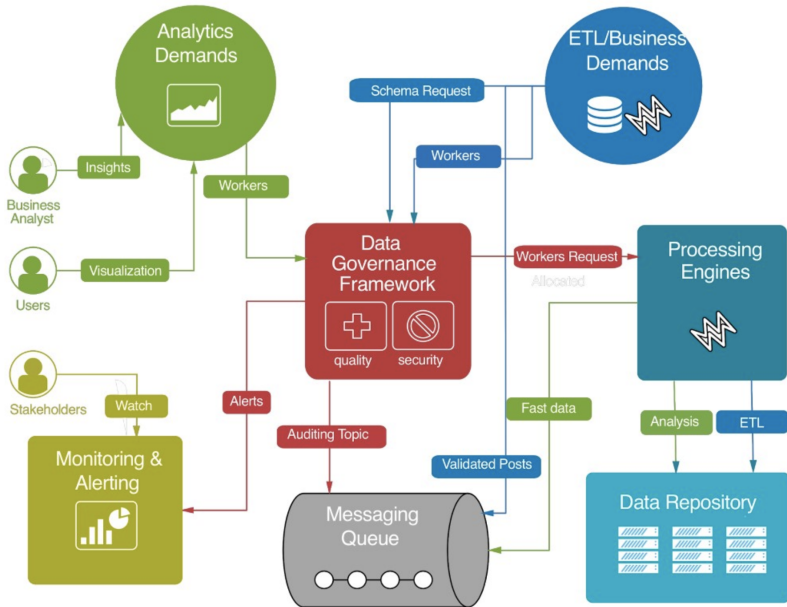
# Knight

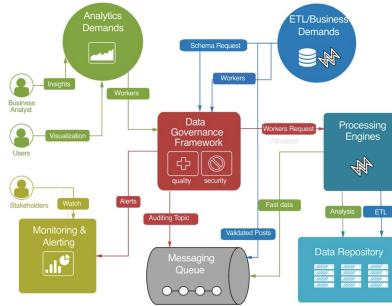
Was not so lucky

# GOTTA GO FAST!



- Reduce reaction time
- Minimize risk surface
- Compete for best offers in market
- ...





- Improve ETL
- Realtime Analytics
- Automated Governance
- Superior Visualization

## 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/>