Grid Dynamics

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]

Grid Dynamics

SPB

Saint-Petersburg

SPB

Krakov


BIG
DATA

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

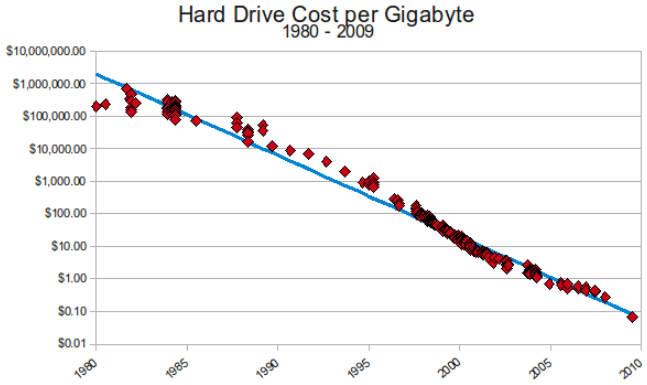
The Streaming Data Platform
as a heavy duty enterprise data marketplace
When to use it in your project.
The case of MMA

The Streaming Data Platform
as a heavy duty enterprise data marketplace
When to use it in your project.
The case of MMA

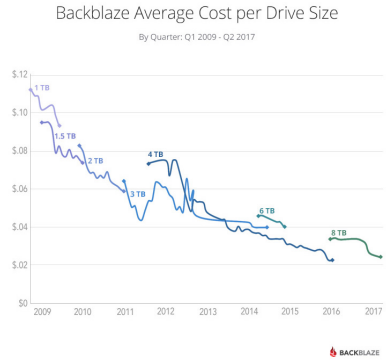
MMA - Multichannel Marketing Automation

Small Introduction to

Streaming Platforms

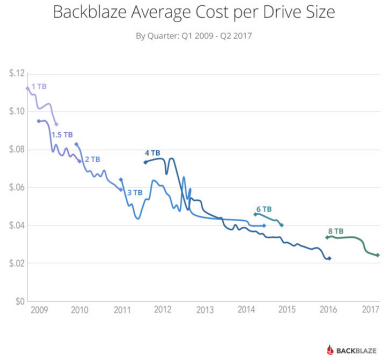


For the past 35+ years or so, hard drives prices have dropped, from around $500,000 per gigabyte in 1981 to less than $0.02 per gigabyte today



[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!?

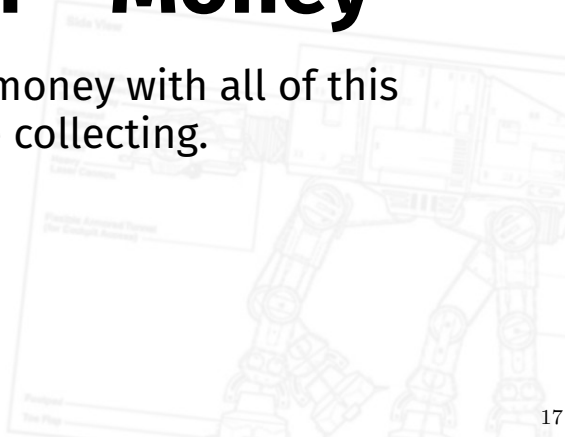


[https://www.backblaze.com/blog/hard-drive-cost-per-gigabyte/]

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

Do you know this company?

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.

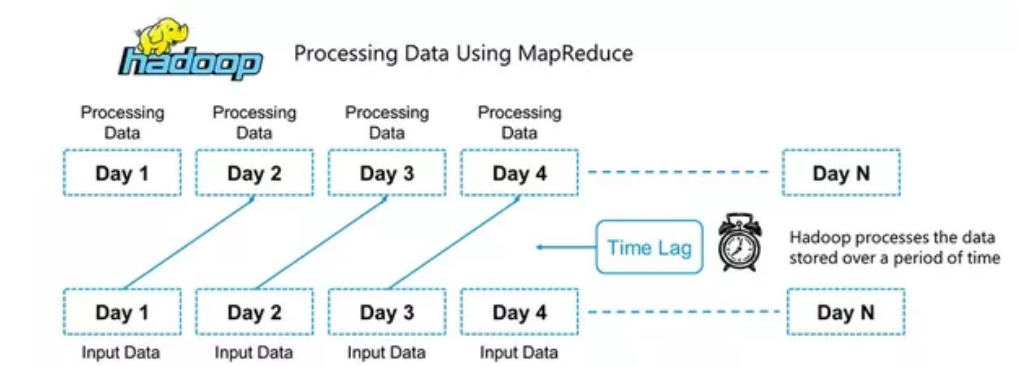
- BigData
- Lot of analytics

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

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.

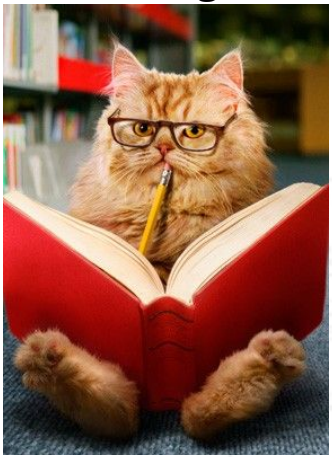
My previous project

Fraud Detection System

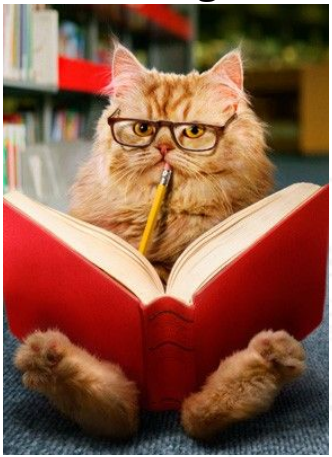


- Collect - 24hrs
- Process - 6hrs
- Block Fraudsters - 10 minutes

We’re loosing money!



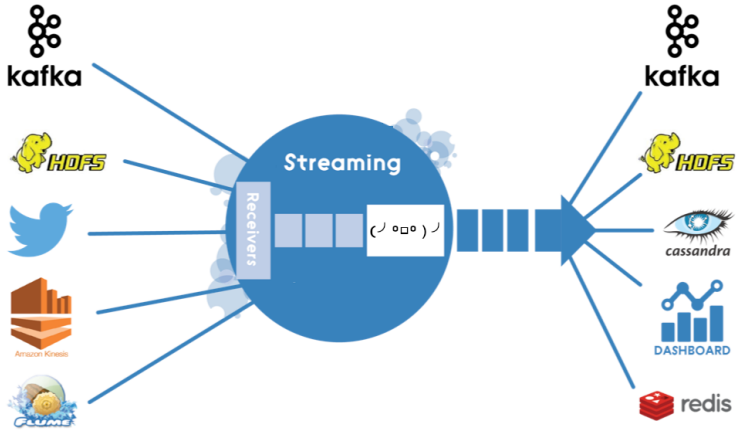
We’re loosing money!

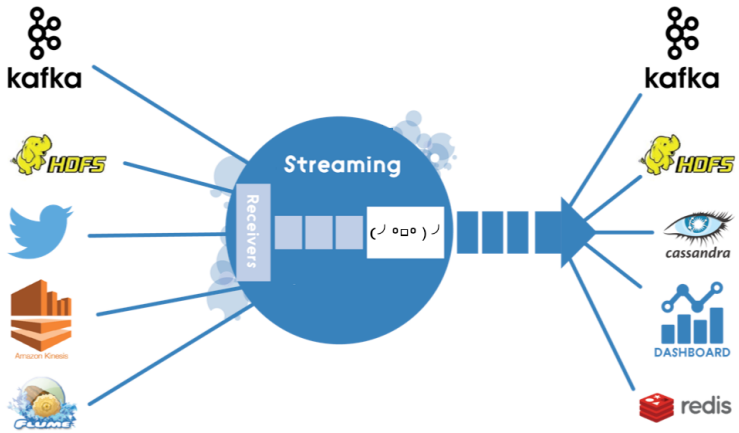


30hr lag →70 000$ per day

What could be improved here?

Reaction Time!

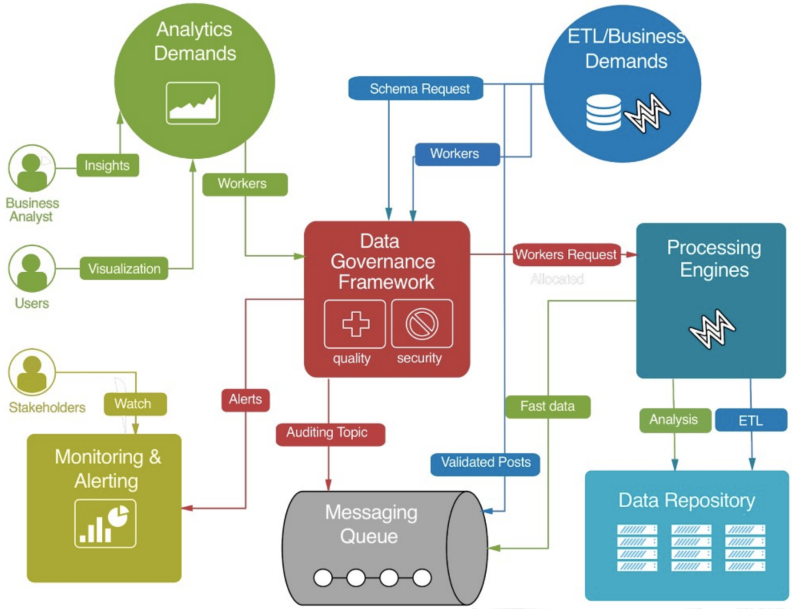

kafka
HDFS
Amazon Kinesis
FLUME
Streaming
Receivers
(ノ °□°)ノ
kafka
HDFS
cassandra
DASHBOARD
redis

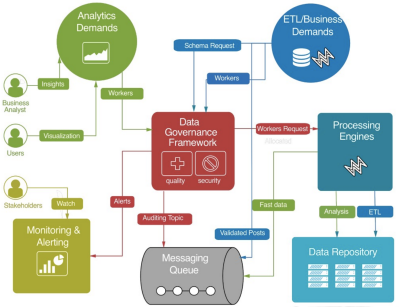


30hr →10mins
And Fraudsters were Disappointed

Was not so lucky

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


Analytics Demands
ETL/Business Demands
Schema Request
Workers
Insights
Business Analyst
Workers
Visualization
Users
Data Governance Framework
quality
security
Workers Request
Allocated
Processing Engines
Stakeholders
Watch
Alerts
Auditing Topic
Fast data
Analysis
ETL
Monitoring & Alerting
Validated Posts
Messaging Queue
Data Repository



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

References:

- 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/