



SEACOM

the leading open source architects



redhat.

An unconventional approach to the BigData projects: self-services data with Kafka, Elastic and Dremio

Stefano Pampaloni, CEO @Seacom srl



#RedHatOSD

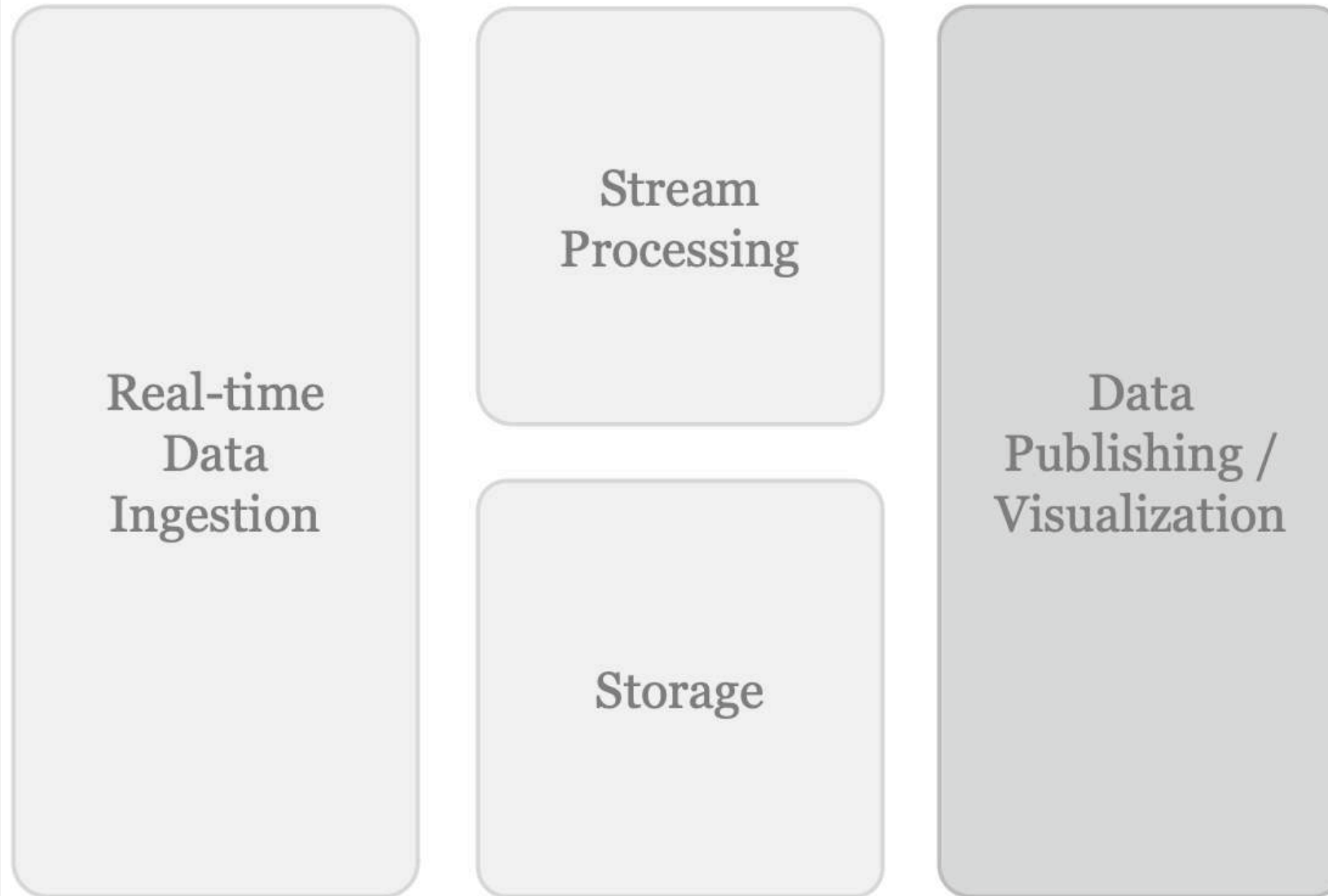
Typical High Level Architecture

Real-time
Data
Ingestion

Typical High Level Architecture



Typical High Level Architecture



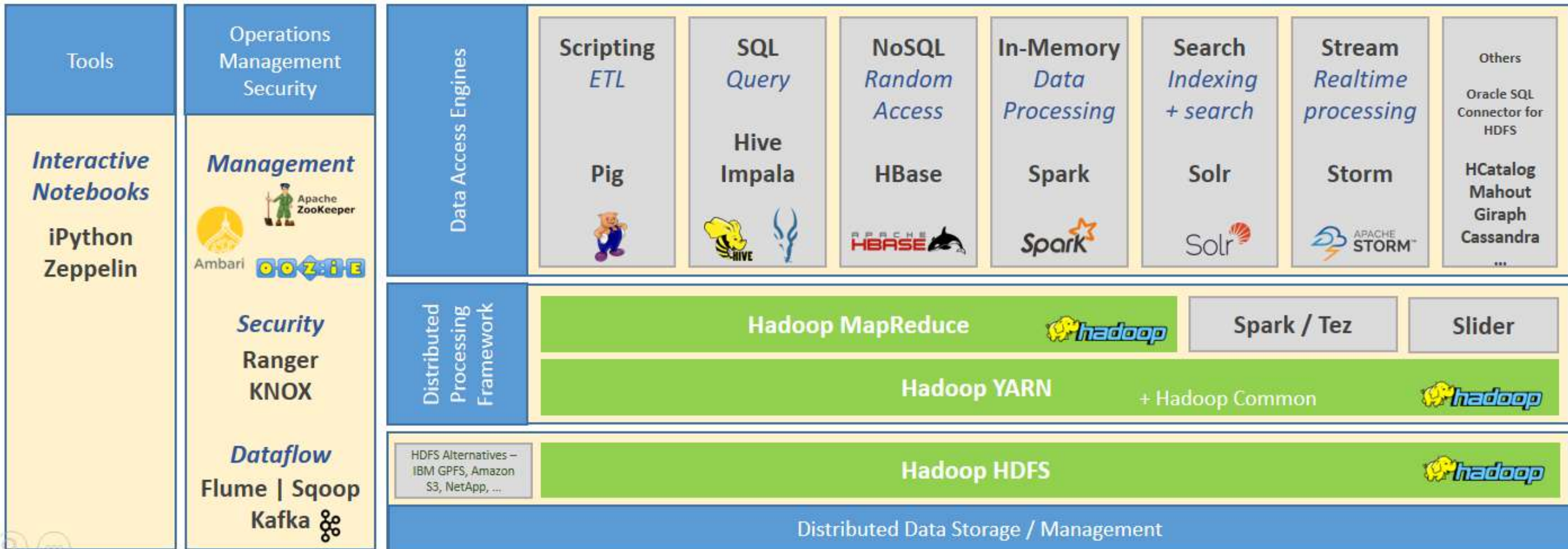
How many clusters do you count?

Apache Kafka

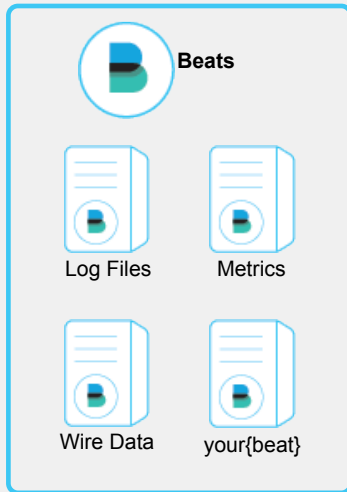
Storm, Flink,
Spark
Streaming,
Ignite, Akka
Streams, Apex,
...

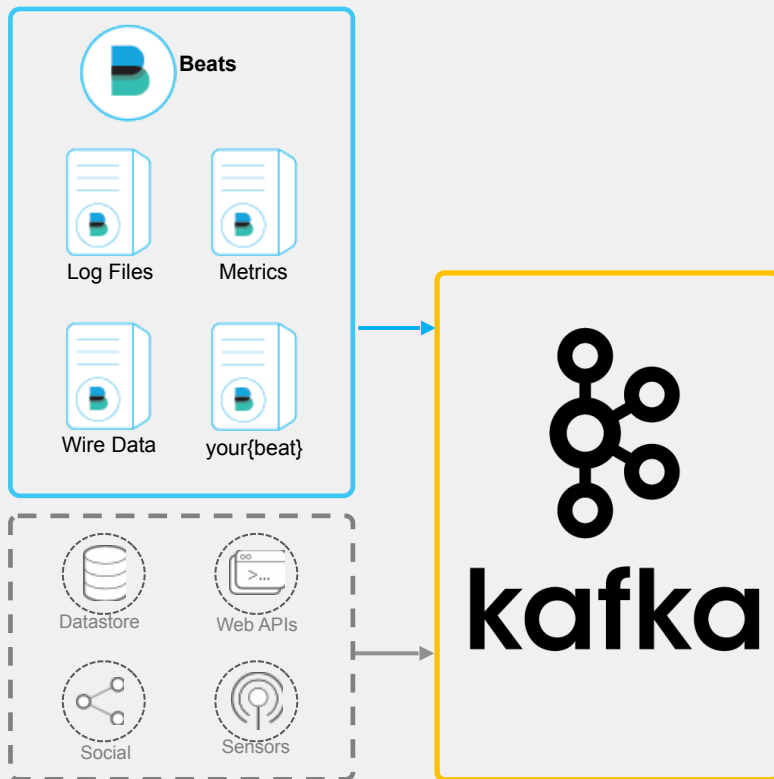
HDFS, NFS,
Ceph,
GlusterFS,
Lustre,
...

NoSQL
(Cassandra,
HBase,
Couchbase,
MongoDB, ...)
or
Elasticsearch,
Solr,
...





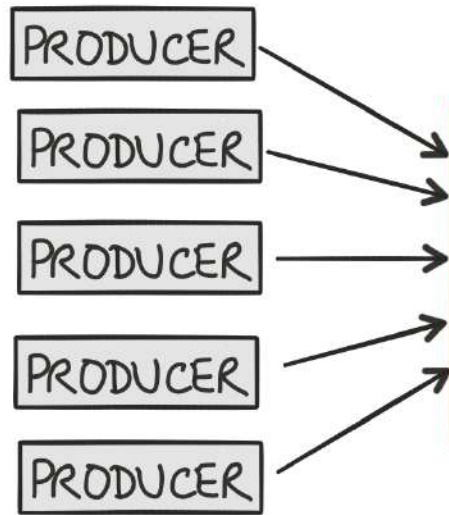




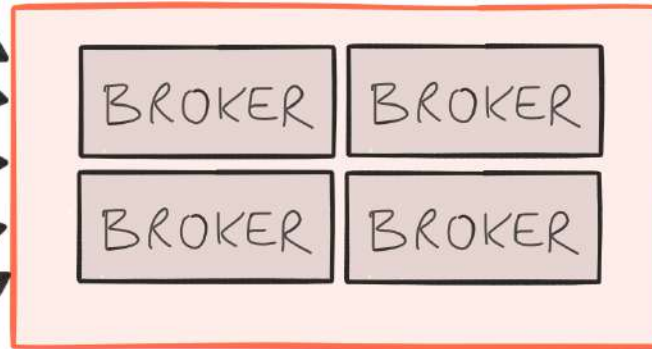
What's Apache Kafka?

- Pub/Sub Logic - Messaging, Done Right
- STORE - Hadoop, Made Fast
- ETL & Data Integration as a Platform
- Stream processing

WRITE

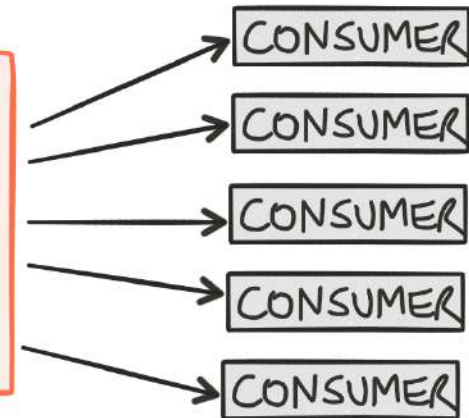


KAFKA
CLUSTER

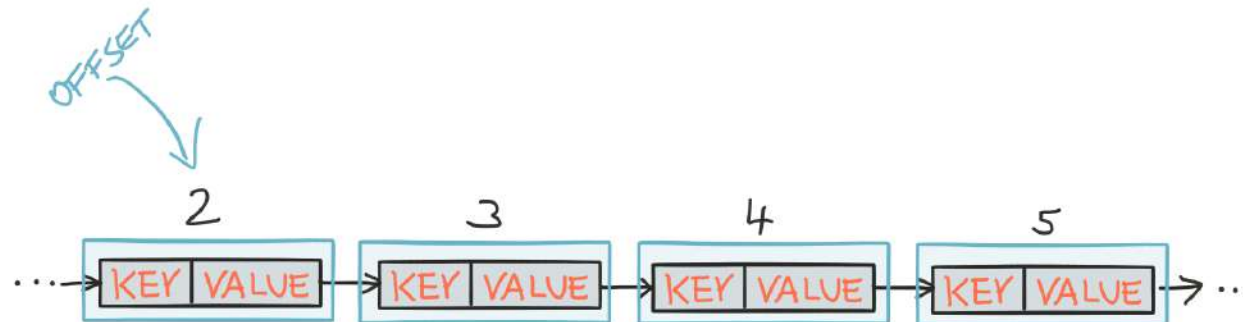
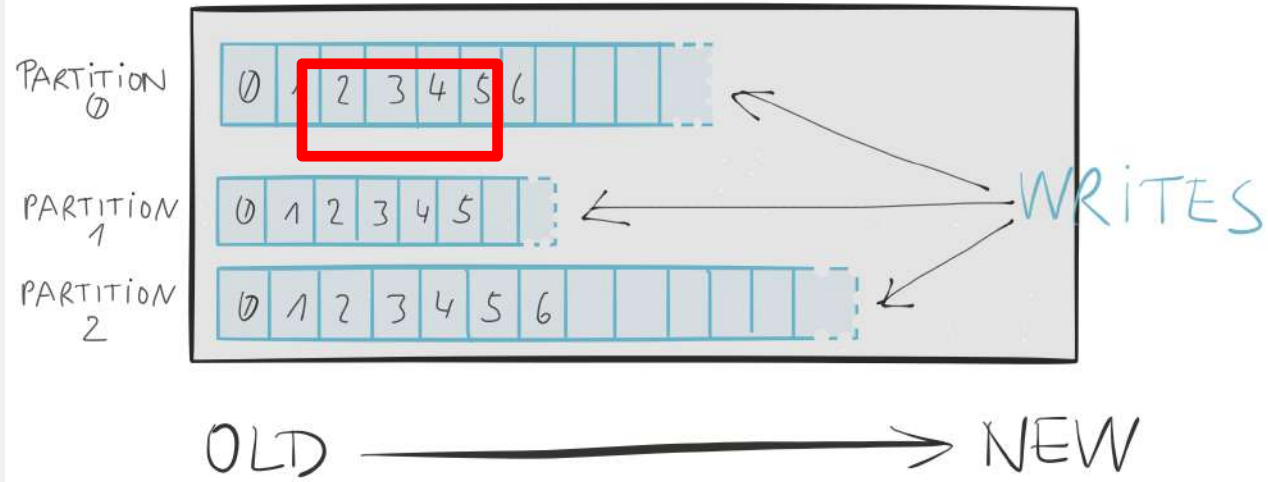


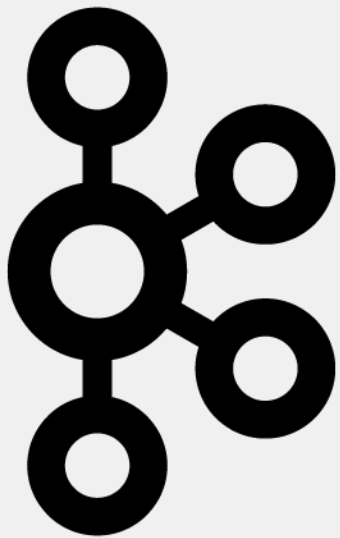
STORE + SERVE

READ



TOPIC (e.g. USER CLICKS)





+



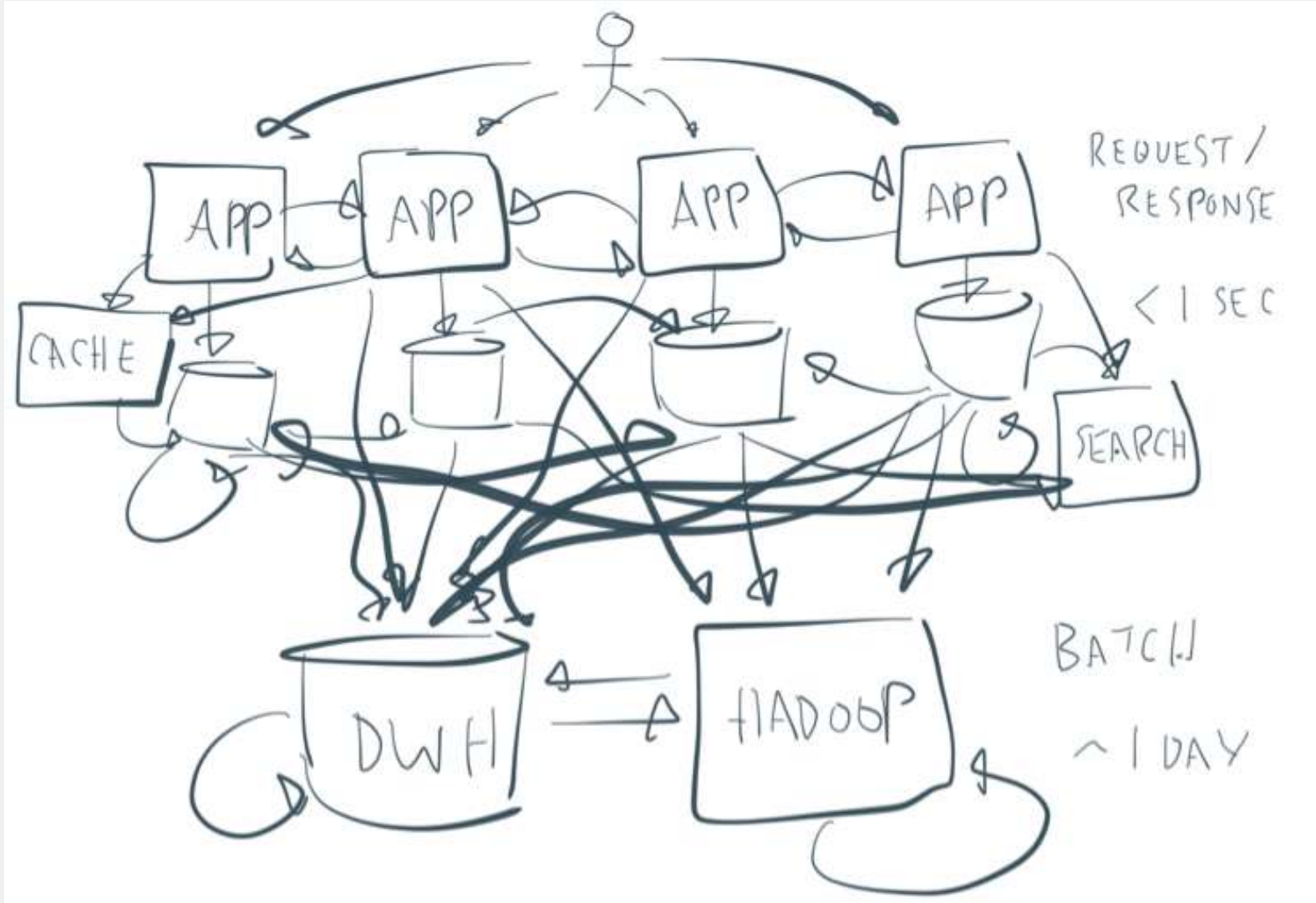
=

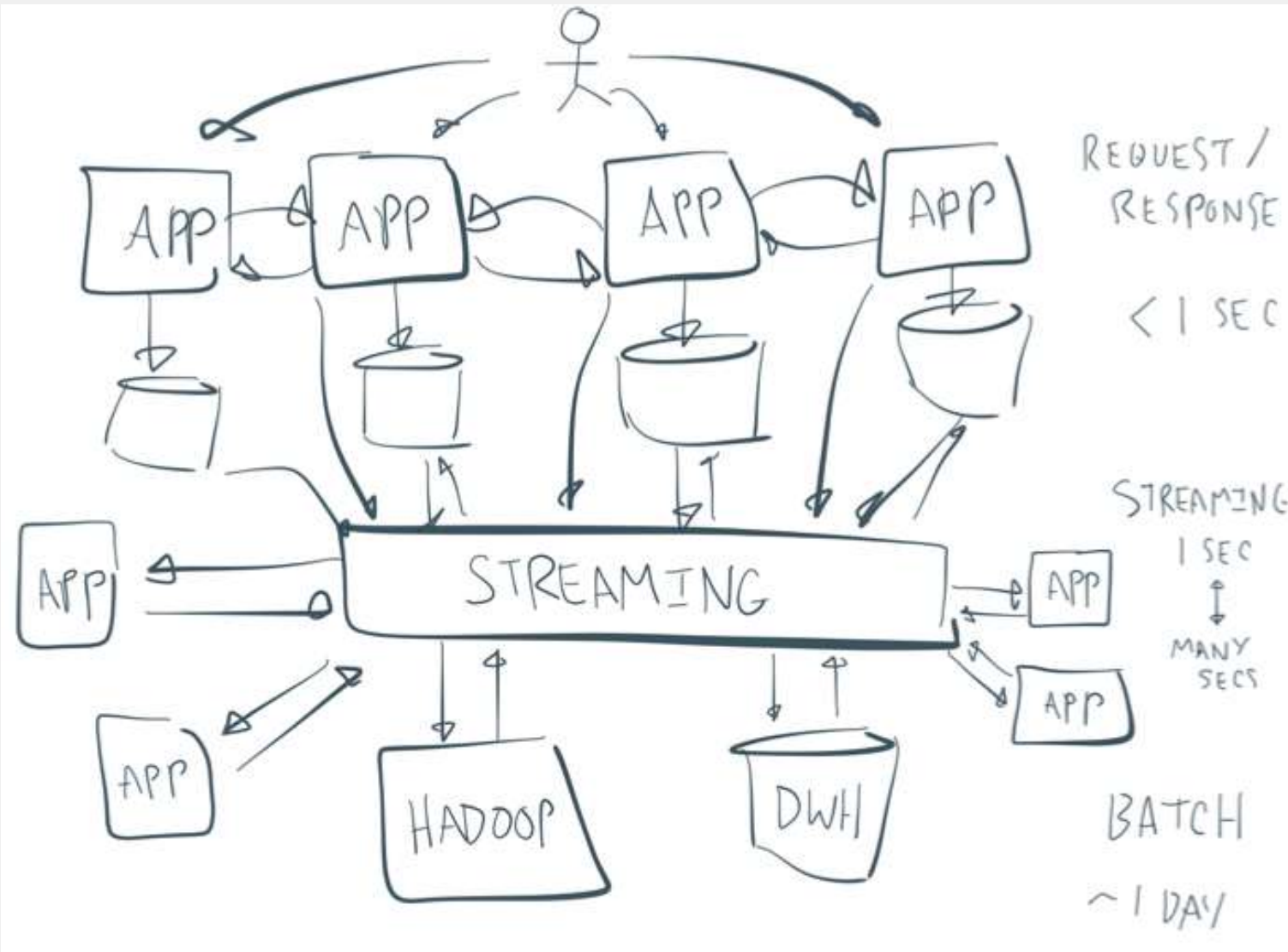
Stream
Processing



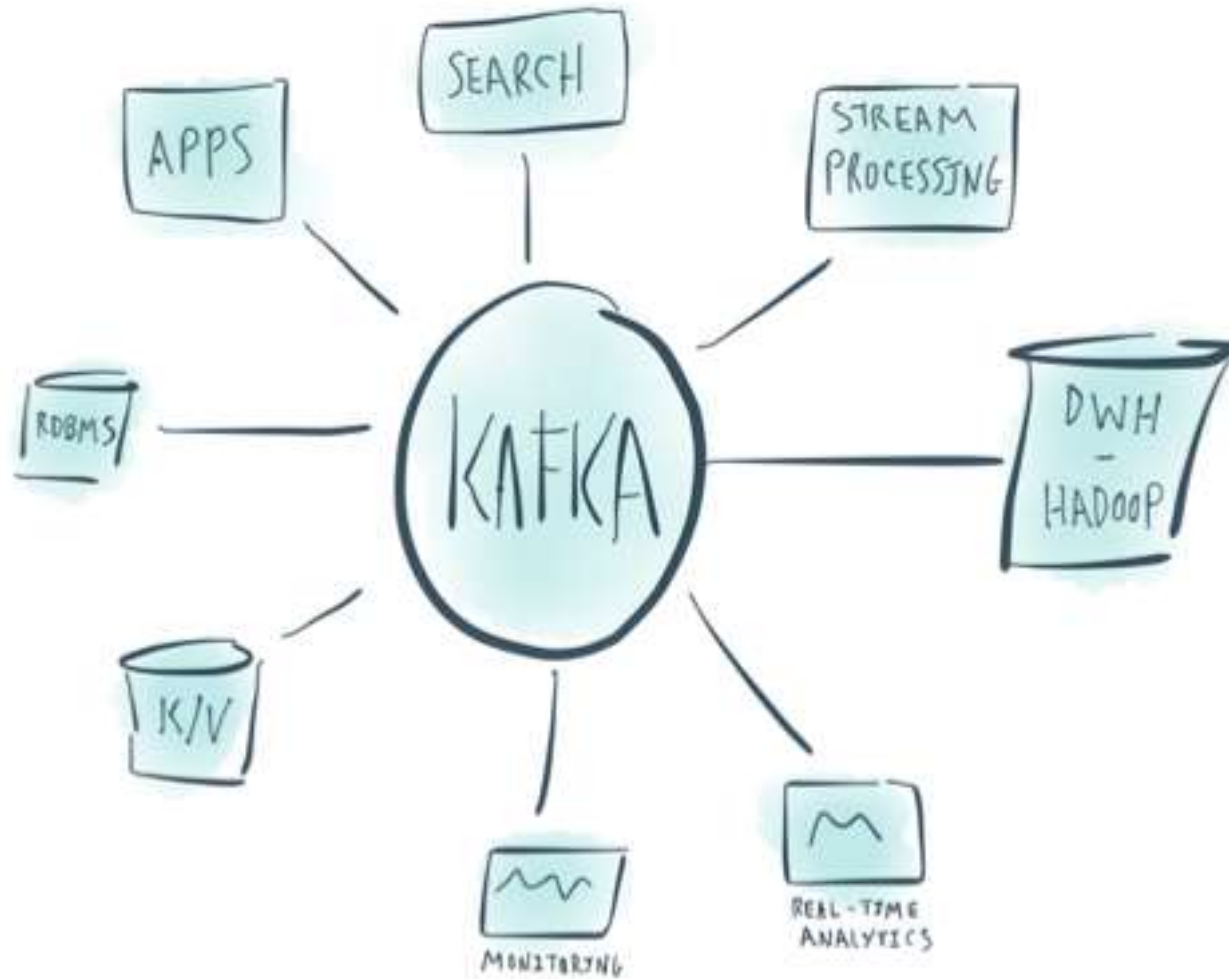


Apache Kafka: A distributed **streaming** platform

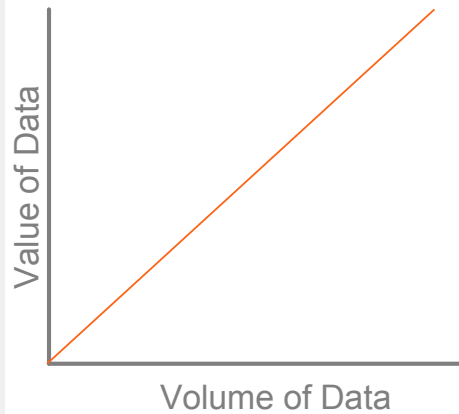




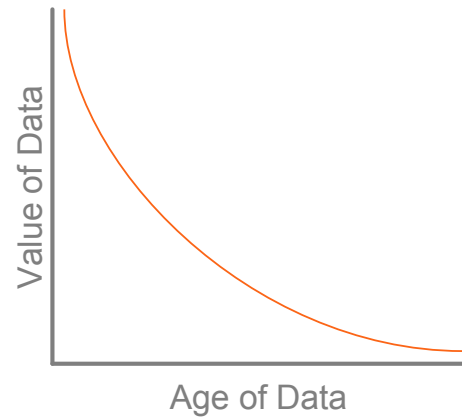
STREAMING PLATFORM



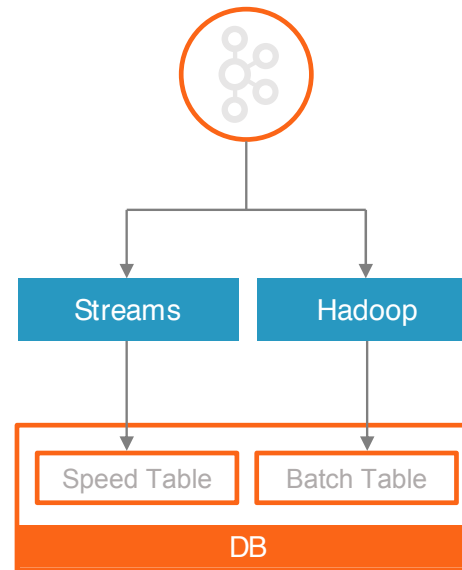
Big data was
The more the better



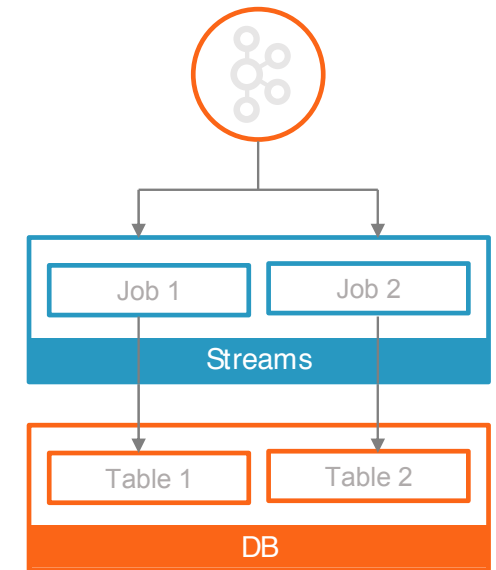
Stream data is
The faster the better



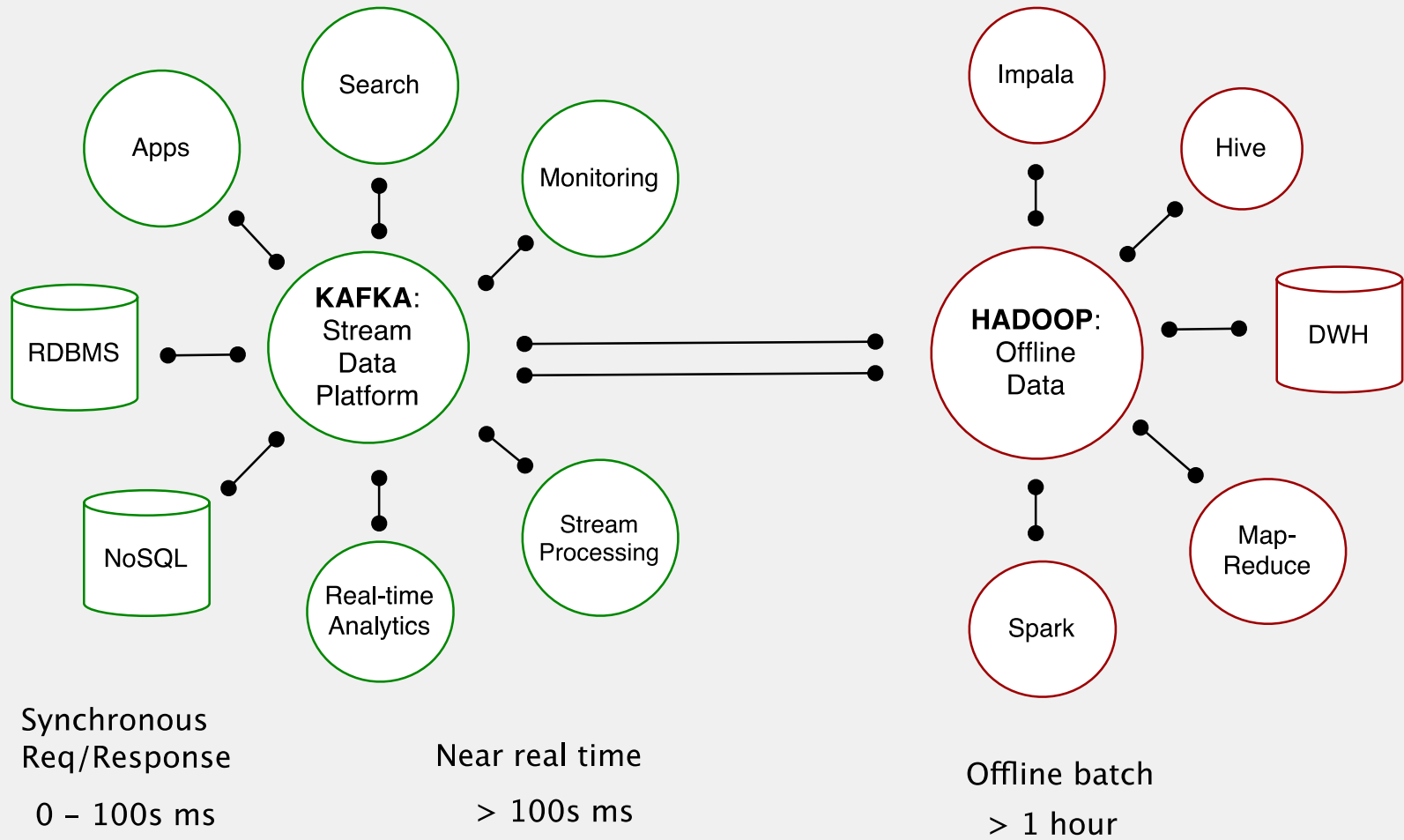
Stream data can be
big or fast (Lambda)

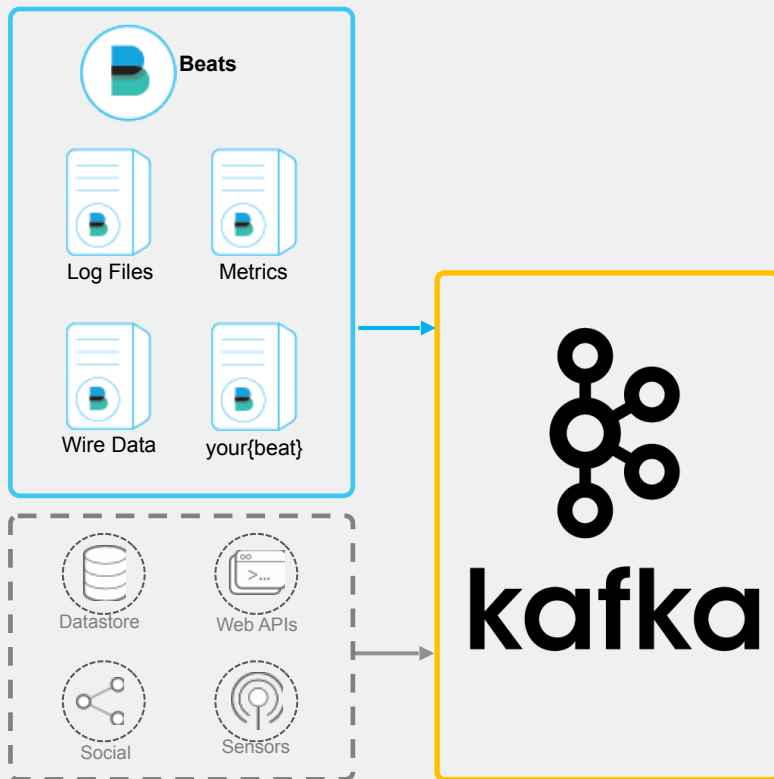


Stream data will be
big AND fast (Kappa)



Apache Kafka is the enabling technology of this transition







How many clusters do you count?

Apache Kafka

Storm, Flink,
Spark
Streaming,
Ignite, Akka
Streams, Apex,
...

HDFS, NFS,
Ceph,
GlusterFS,
Lustre,
...

NoSQL
(Cassandra,
HBase,
Couchbase,
MongoDB, ...)
or
Elasticsearch,
Solr,
...

Simplicity is the ultimate sophistication

Apache Kafka

Distributed Streaming Platform

Publish & Subscribe

to streams of data like a messaging system

Store

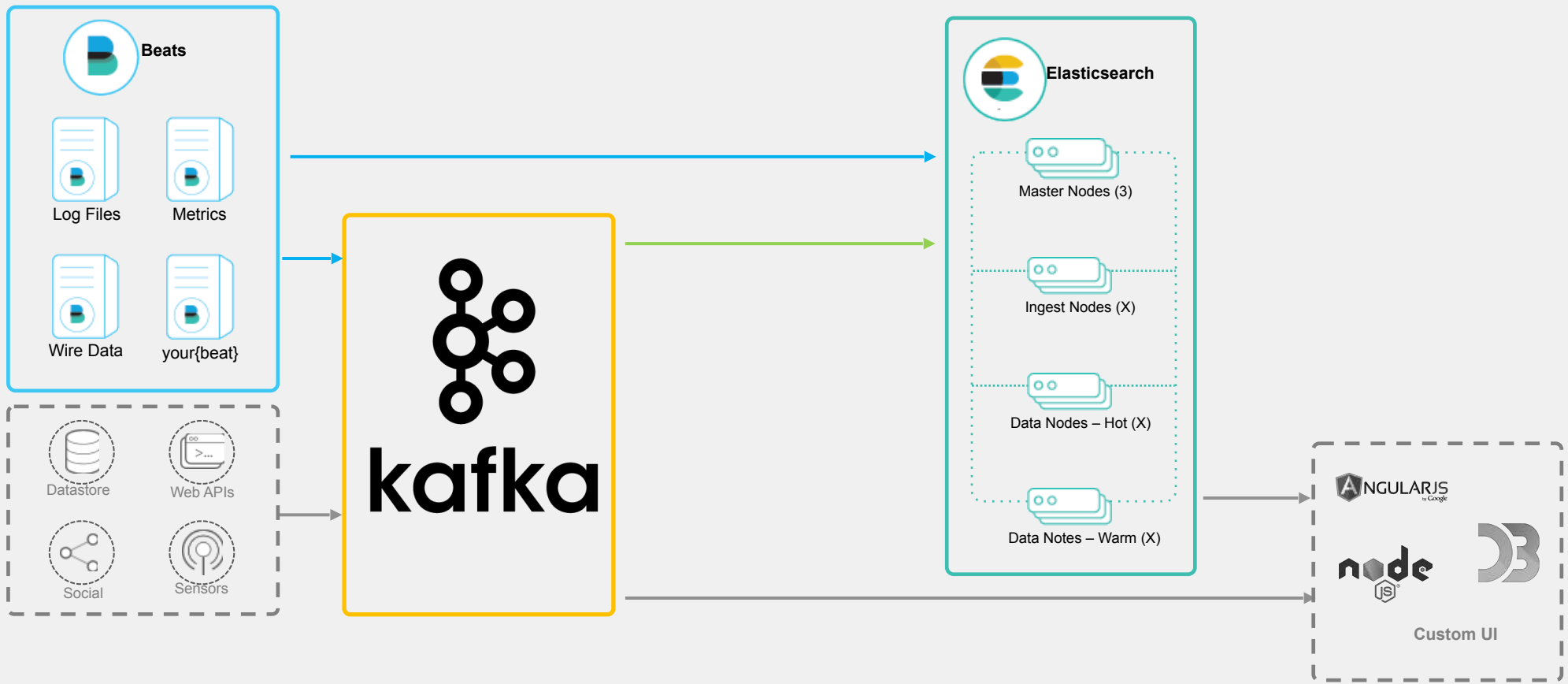
streams of data safely in a distributed replicated cluster

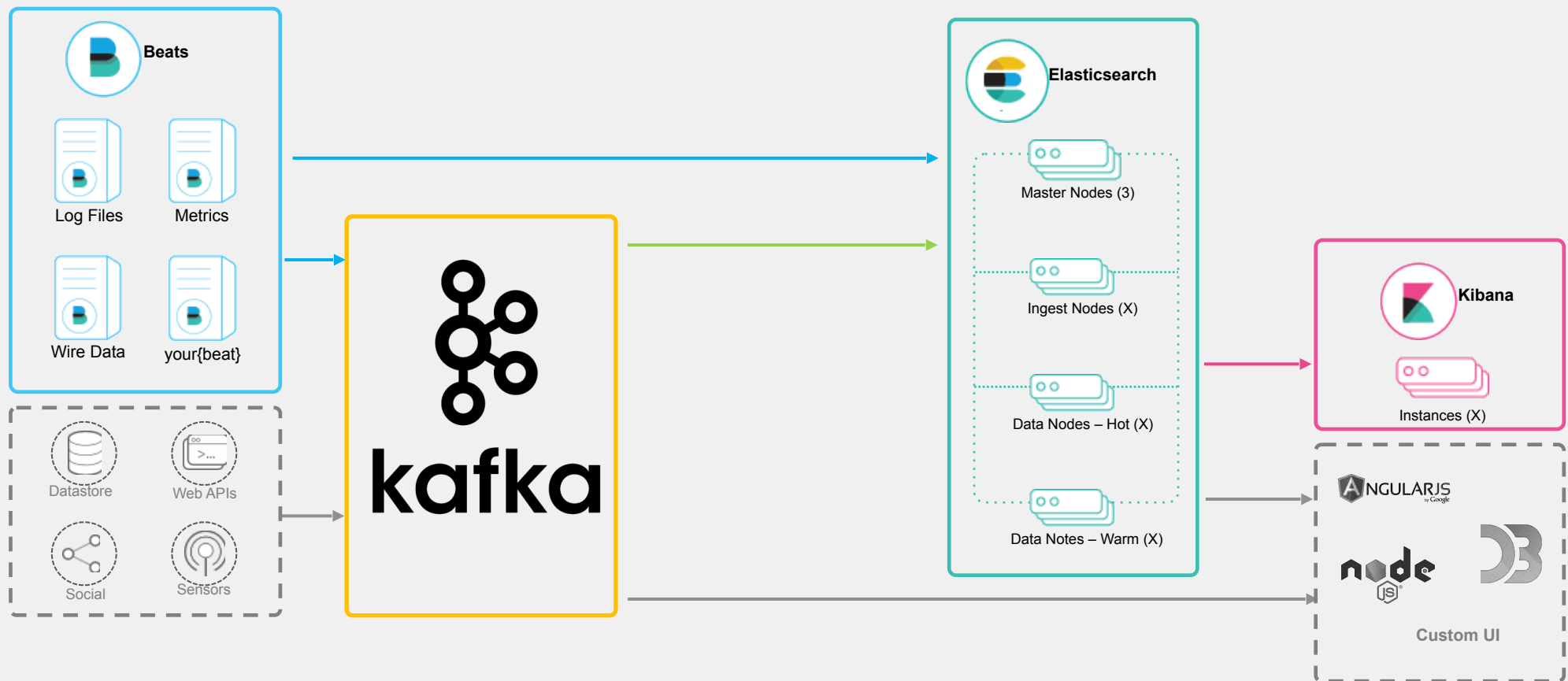
Process

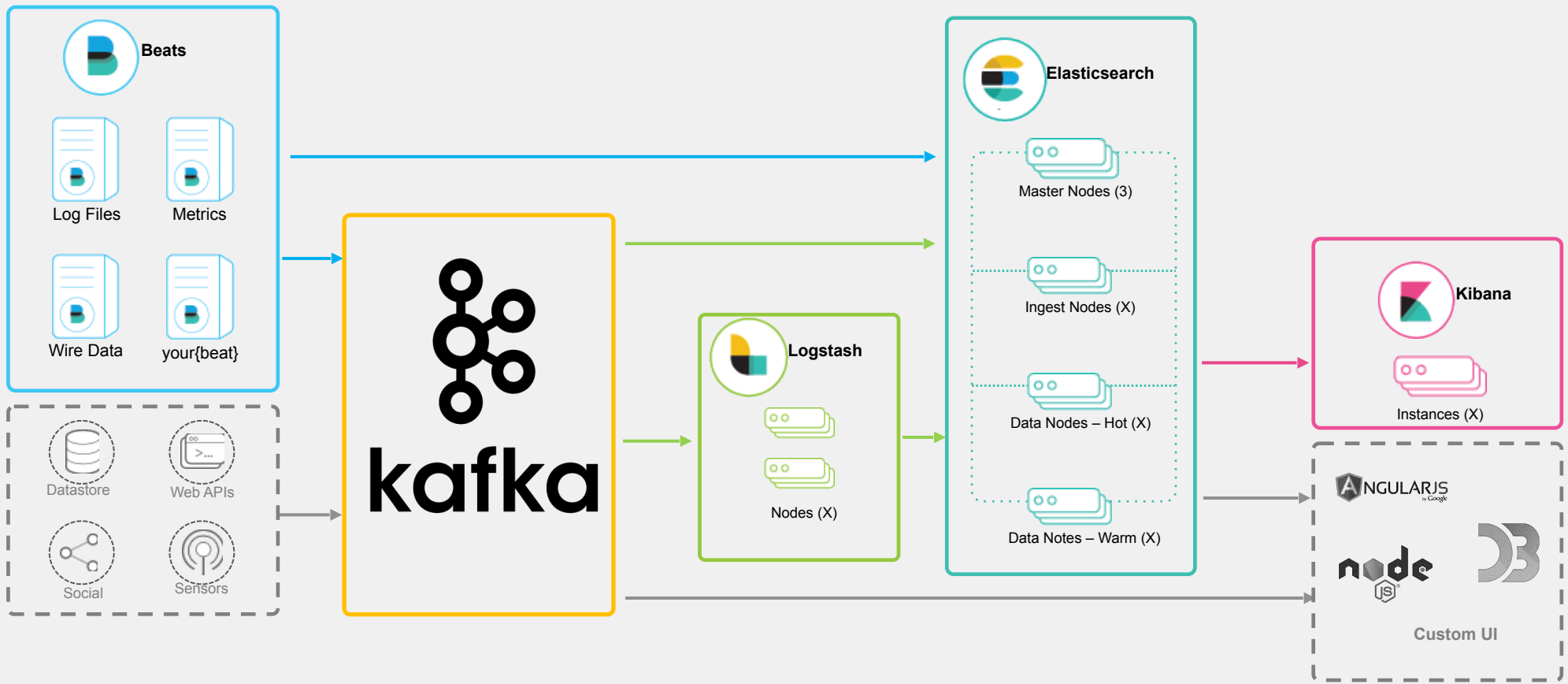
streams of data efficiently and in real-time

Node.js





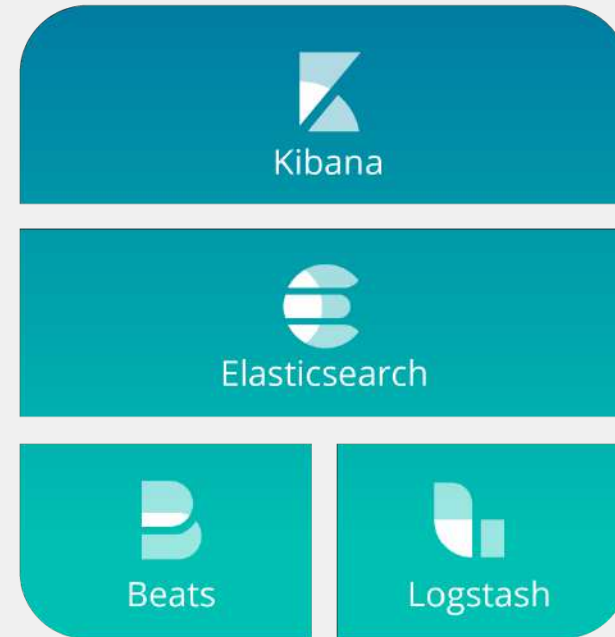




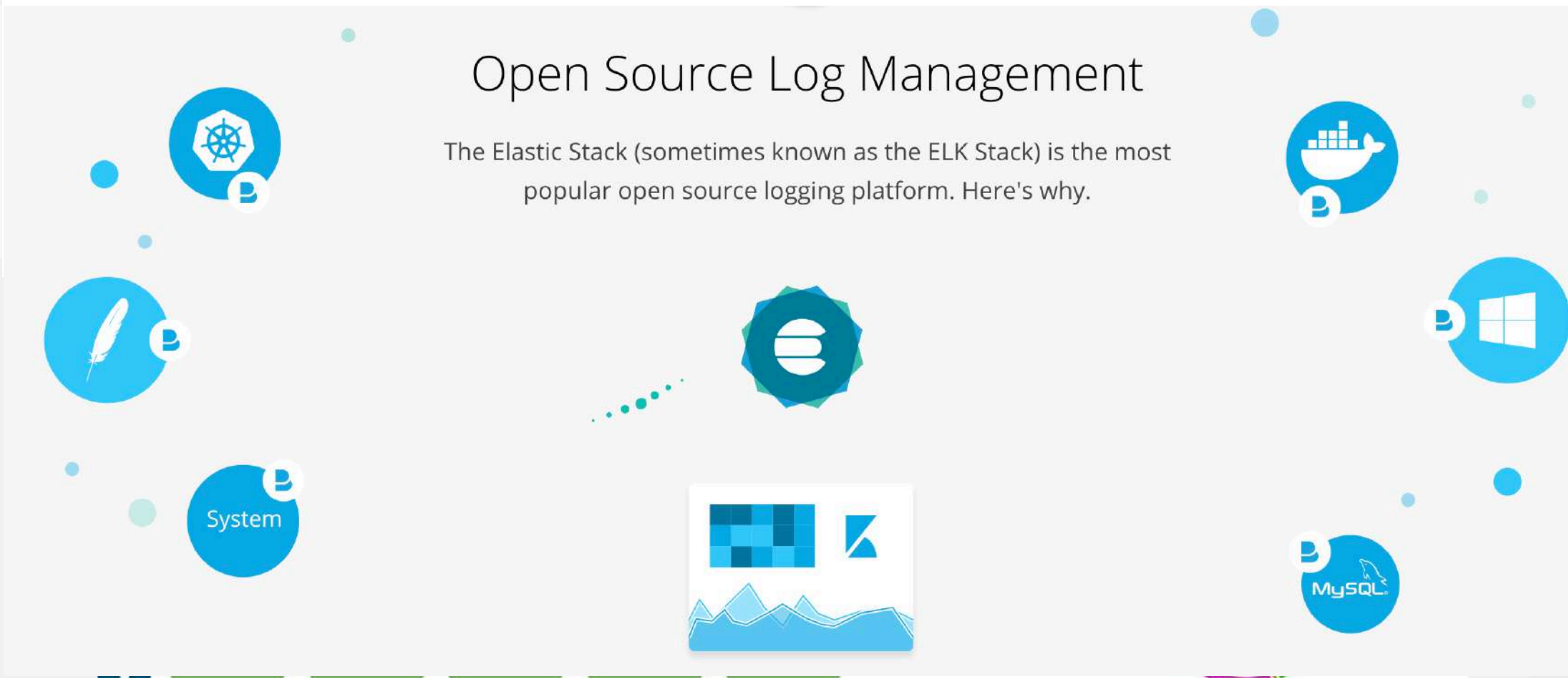


Elastic Stack

100% open source
No enterprise edition



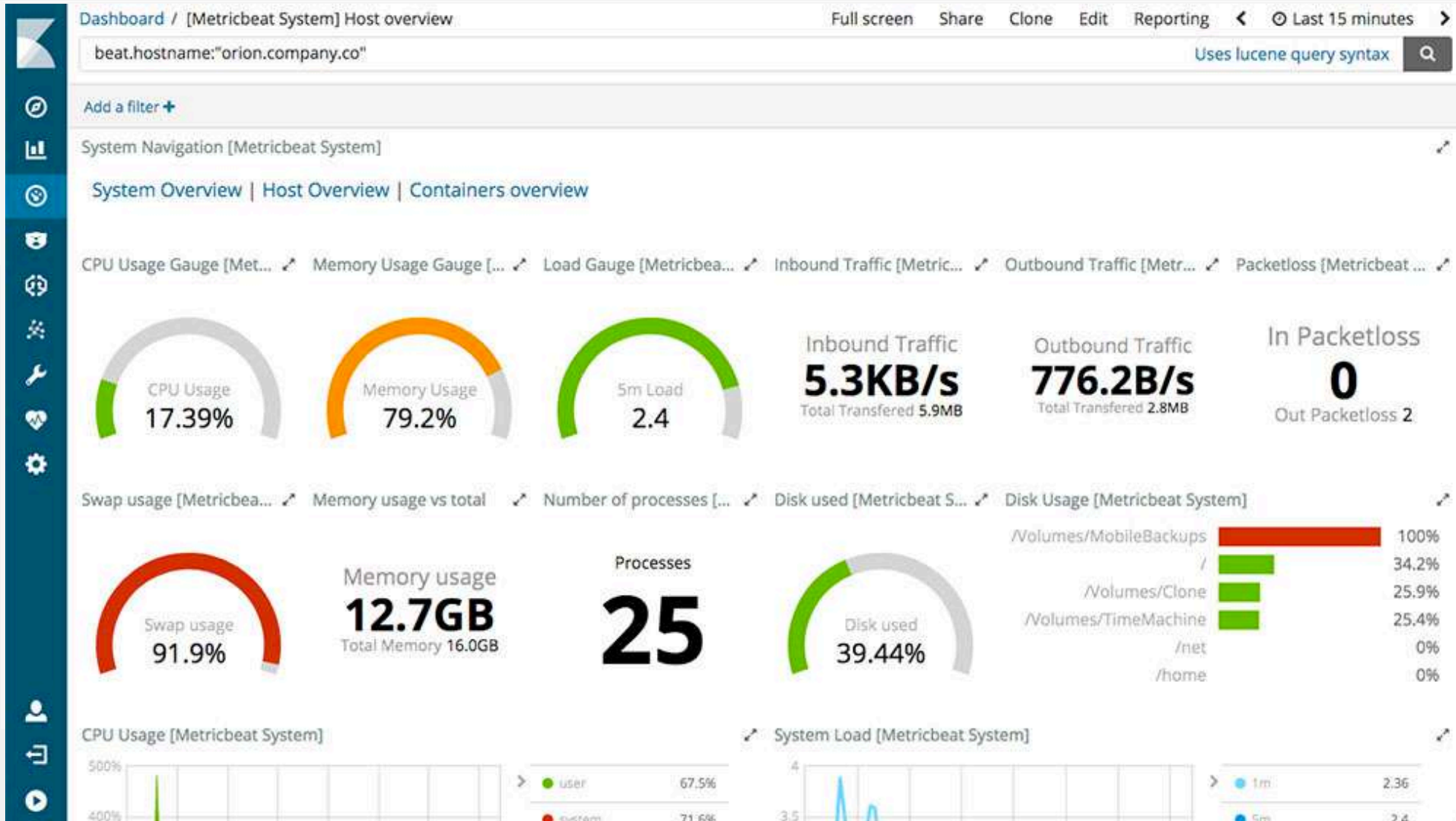
Da strumento a Soluzione



Da strumento a Soluzione



Logging








Da strumento a Soluzione



Metrics

CAI Studio @applepig

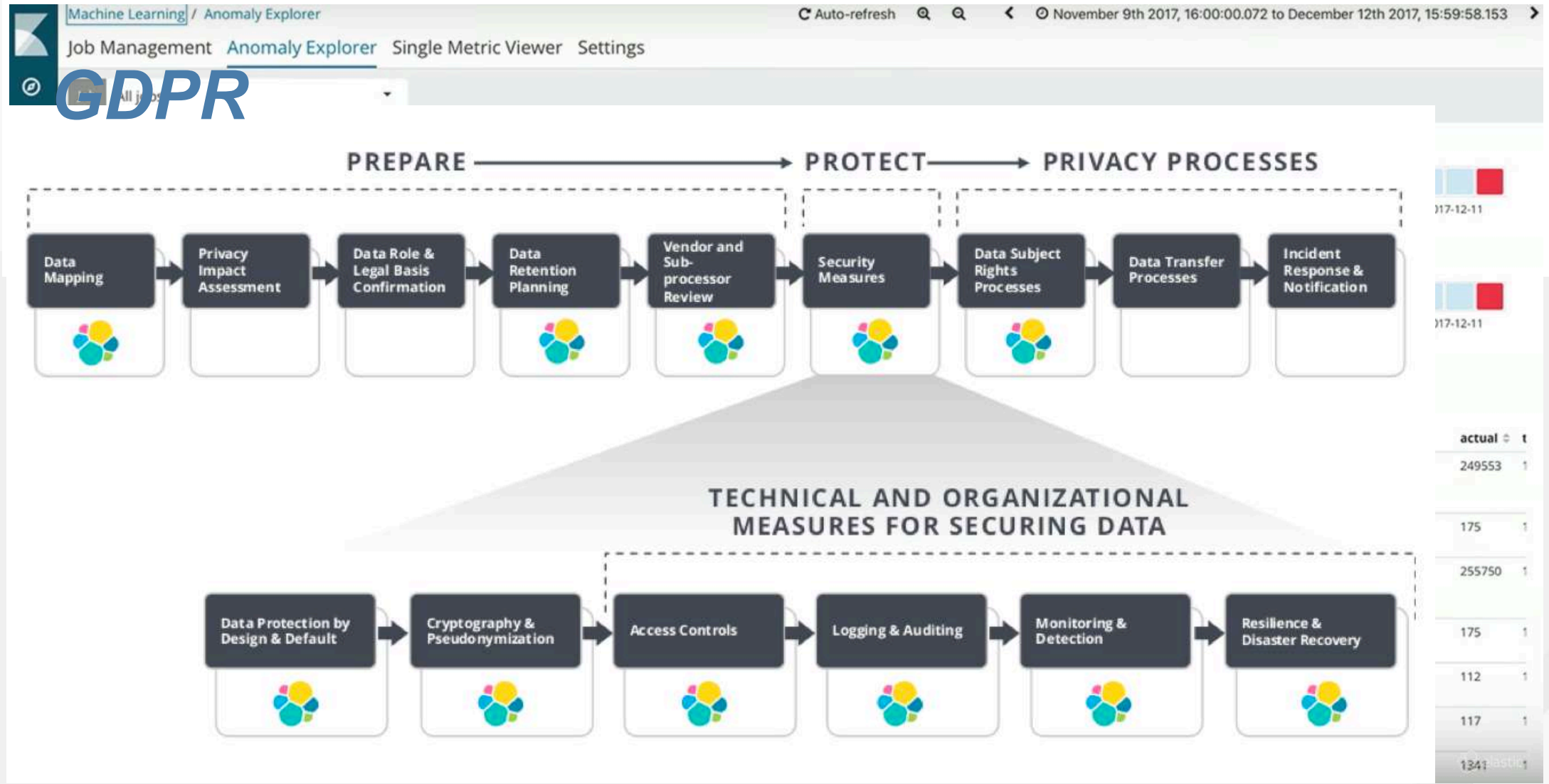
	Grey Lawson Style Sofa \$799.00 ★★★★★
	Eggshell Mid Century Modern Sofa \$795.00 ★★★
	Classic Contemporary Loveseat Sofa \$499 ★★★
	Charcoal Contemporary Loveseat Sofa \$499 ★★
	Tan Contemporary Mid Century Sofa \$599 ★

Rating 6

Da strumento a Soluzione



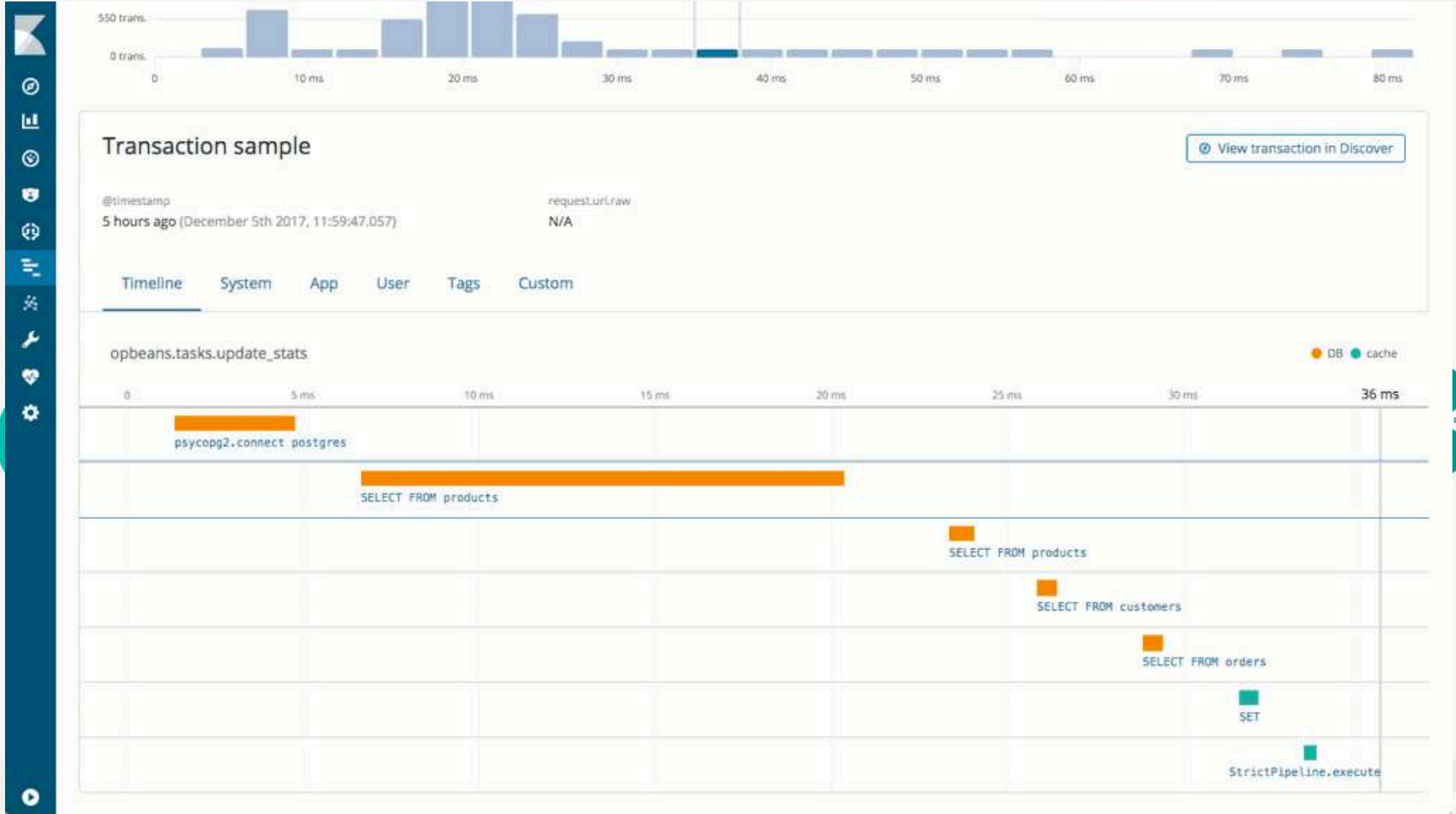
Site Search

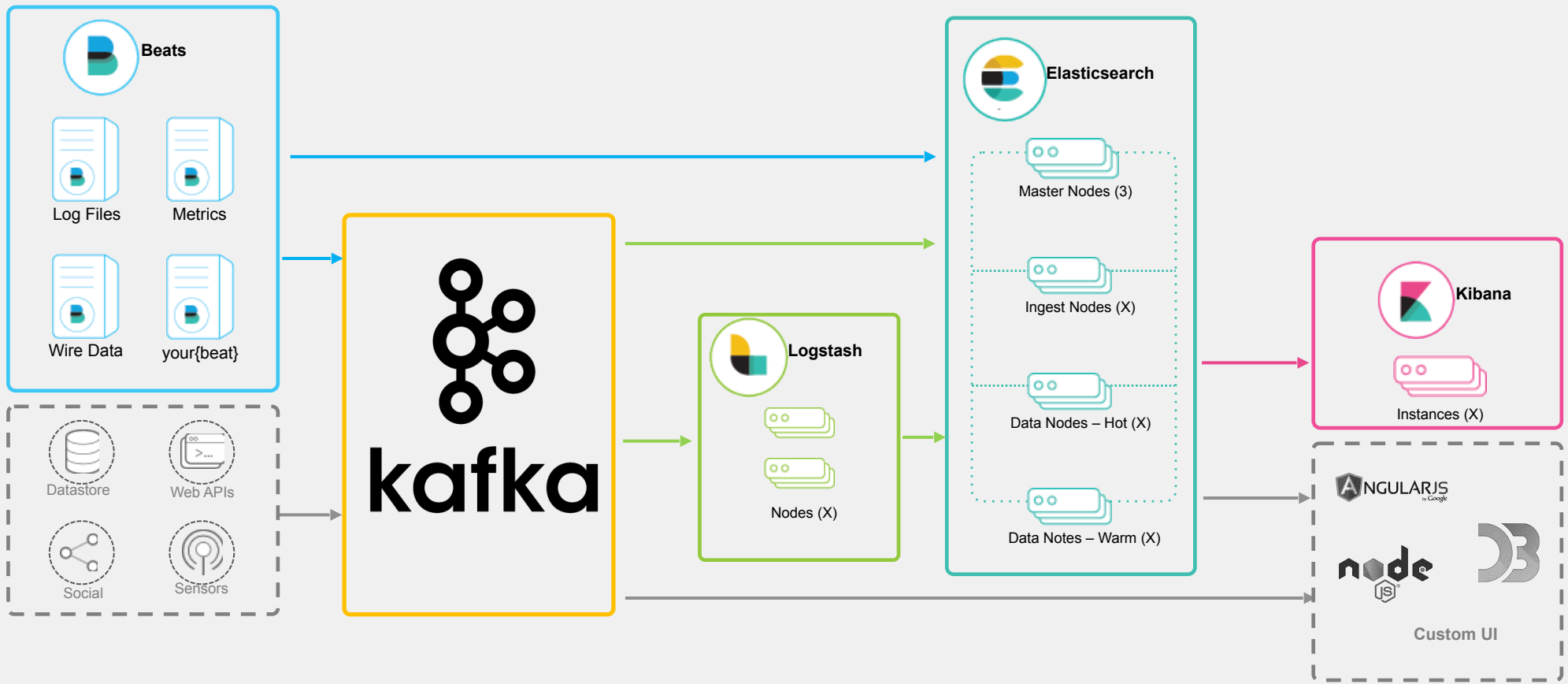


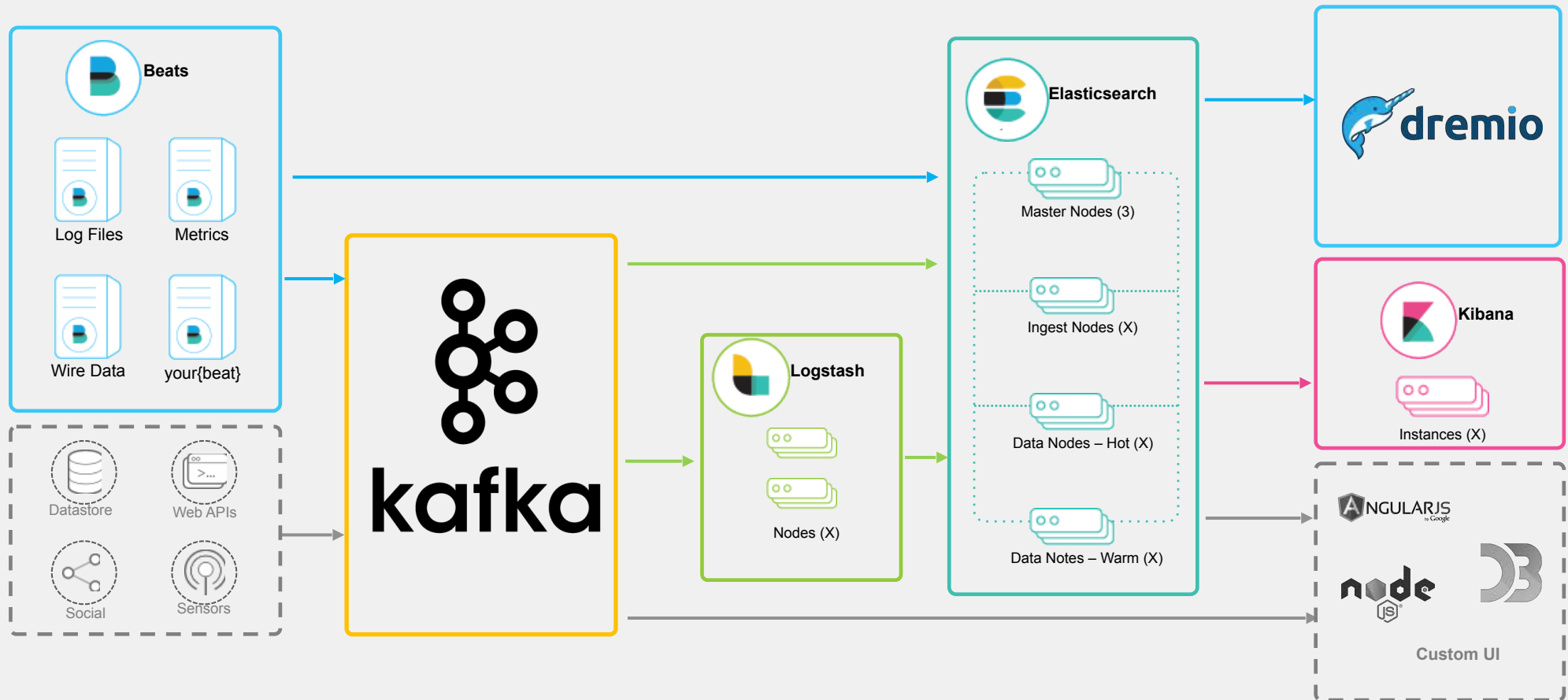
Da strumento a Soluzione

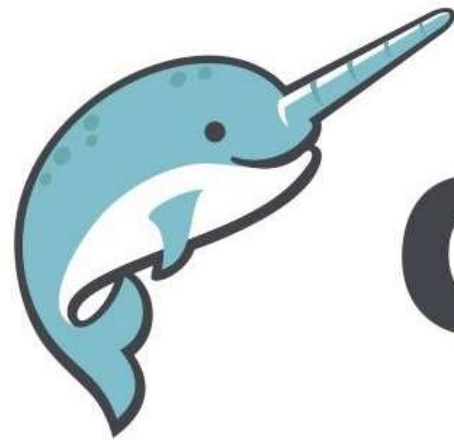


Security Analytics



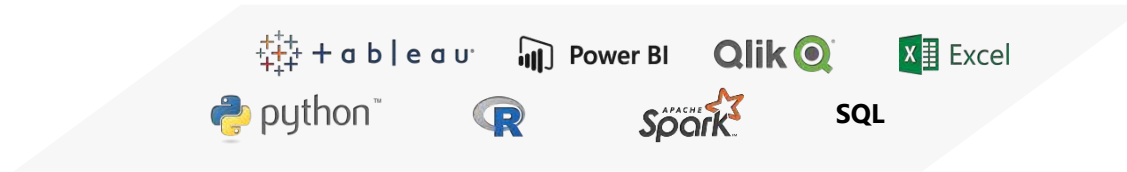






dremio

Your analysts and data scientists are hungry for data



dremio

Data is a massive engineering project today



Data Staging

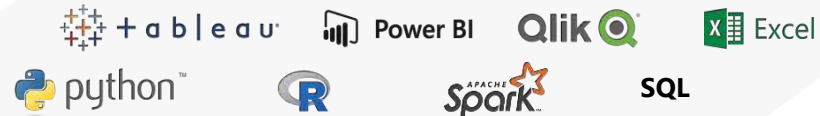


- Custom ETL
- Fragile transforms
- Slow moving

dremio

Data is a massive engineering project today

Data Warehouse



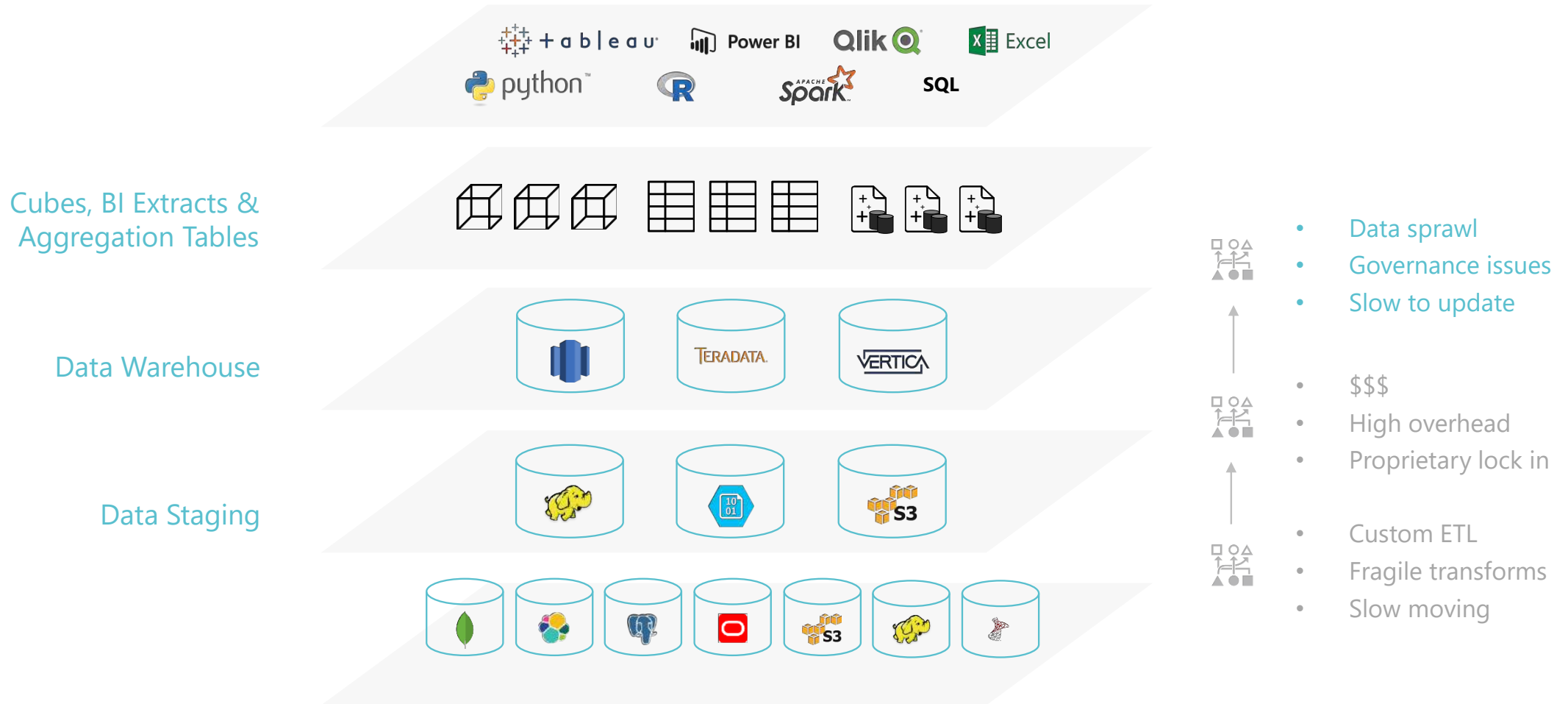
Data Staging



- \$\$\$
- High overhead
- Proprietary lock in
- Custom ETL
- Fragile transforms
- Slow moving

dremio

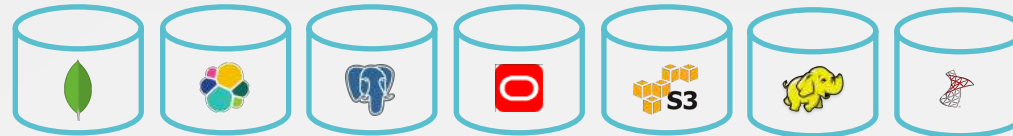
Data is a massive engineering project today



dremio



SQL



These popular first projects are easy wins

 + a b l e a w  Power BI **SQL**

 + a b l e a w  Dashboards

 + a b l e a w  Power BI  Excel



Data Lake

Accelerate SQL on your Data Lake



Hybrid Lake

Join across Data Lake + other systems

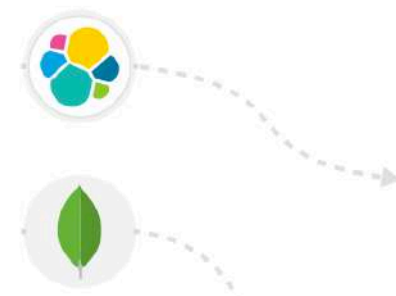


Self-Service Data

Across all your systems, on-prem + cloud

dremio

Sources(2)



Parents(2)

review

Jobs: 128»
Descendants: 10
Fields: 11

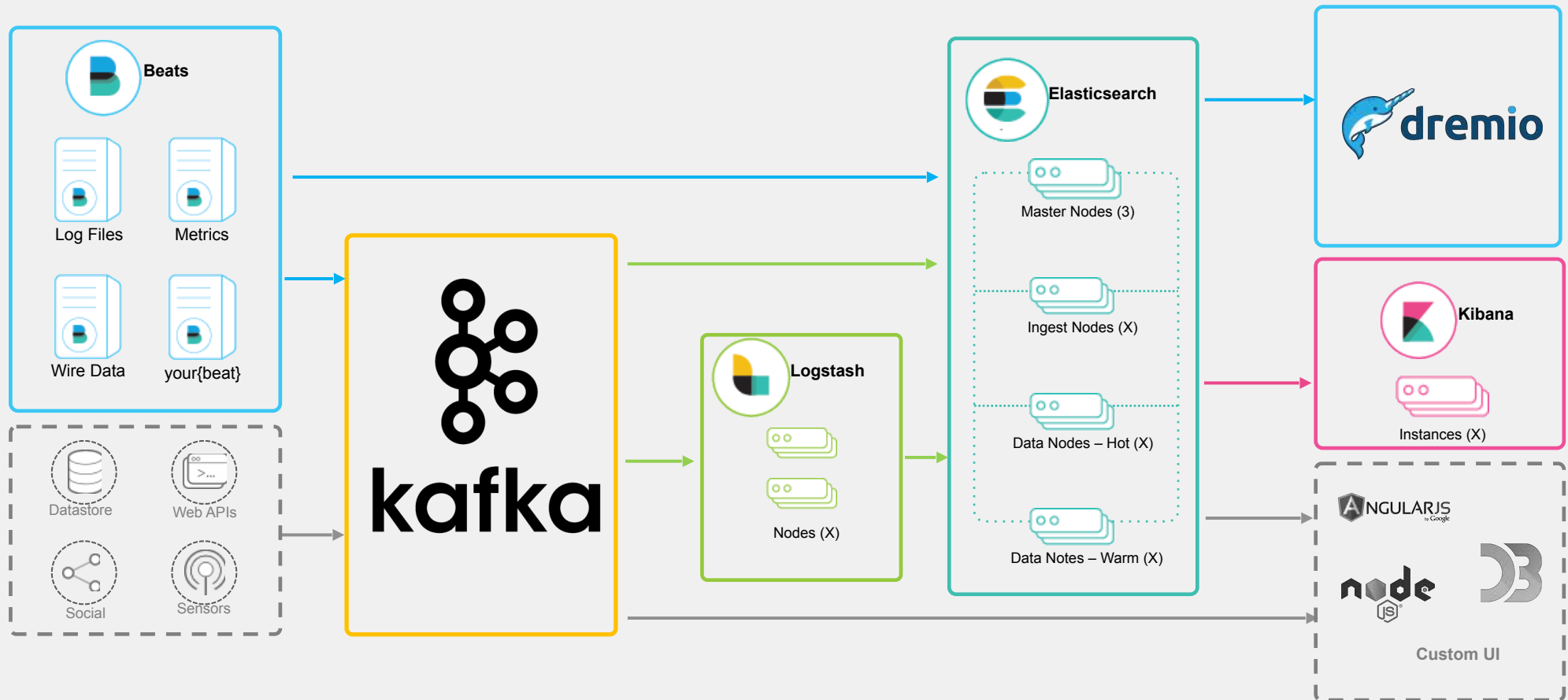
business

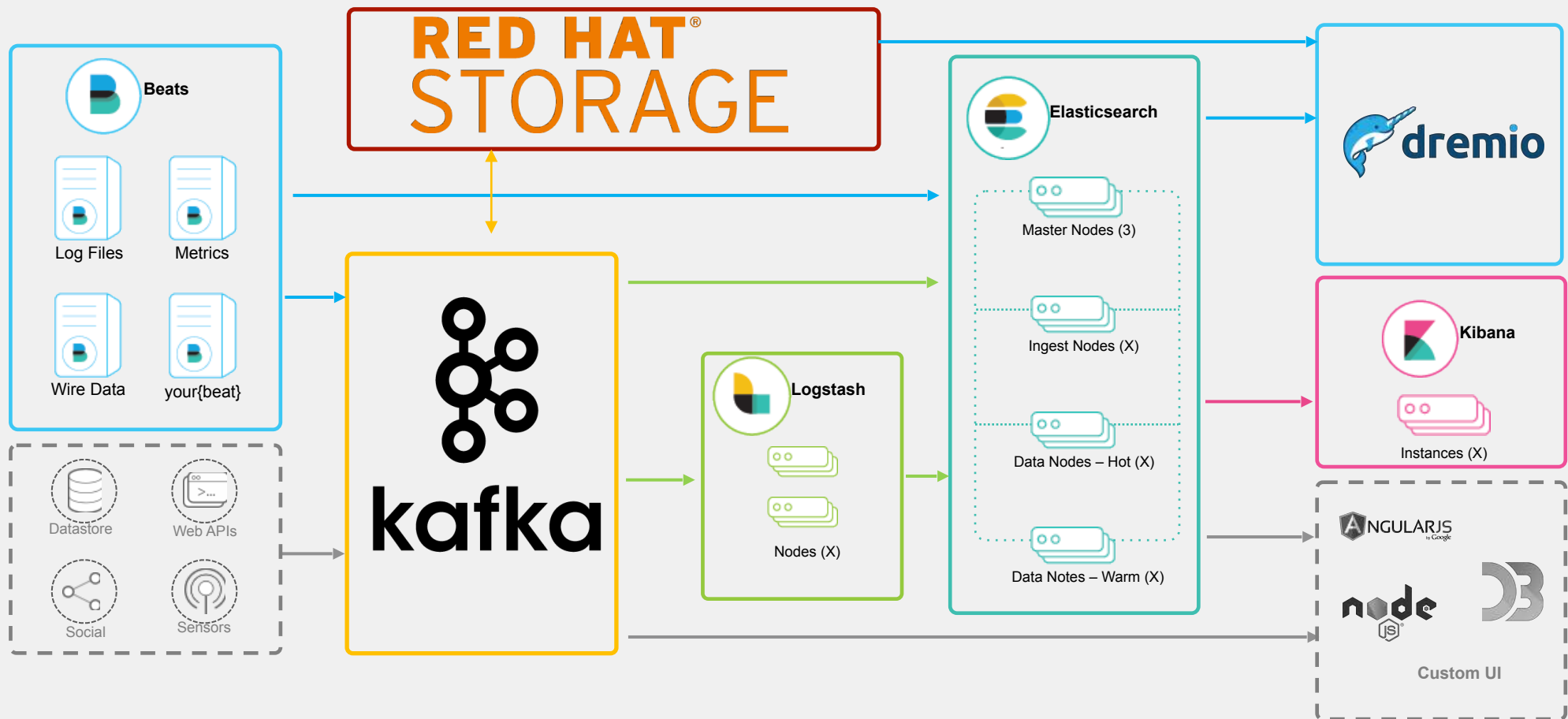
Jobs: 443»
Descendants: 22
Fields: 16

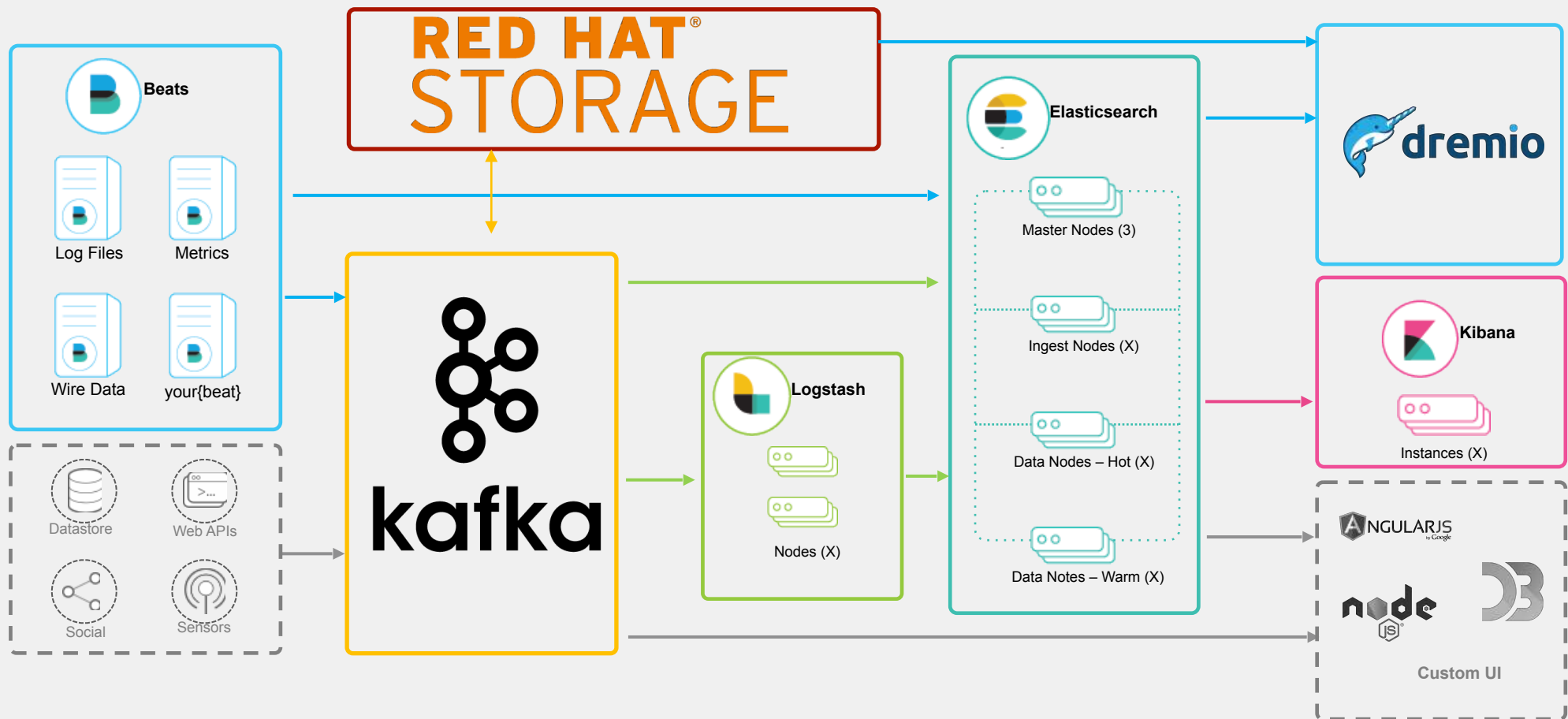
amazingreviews

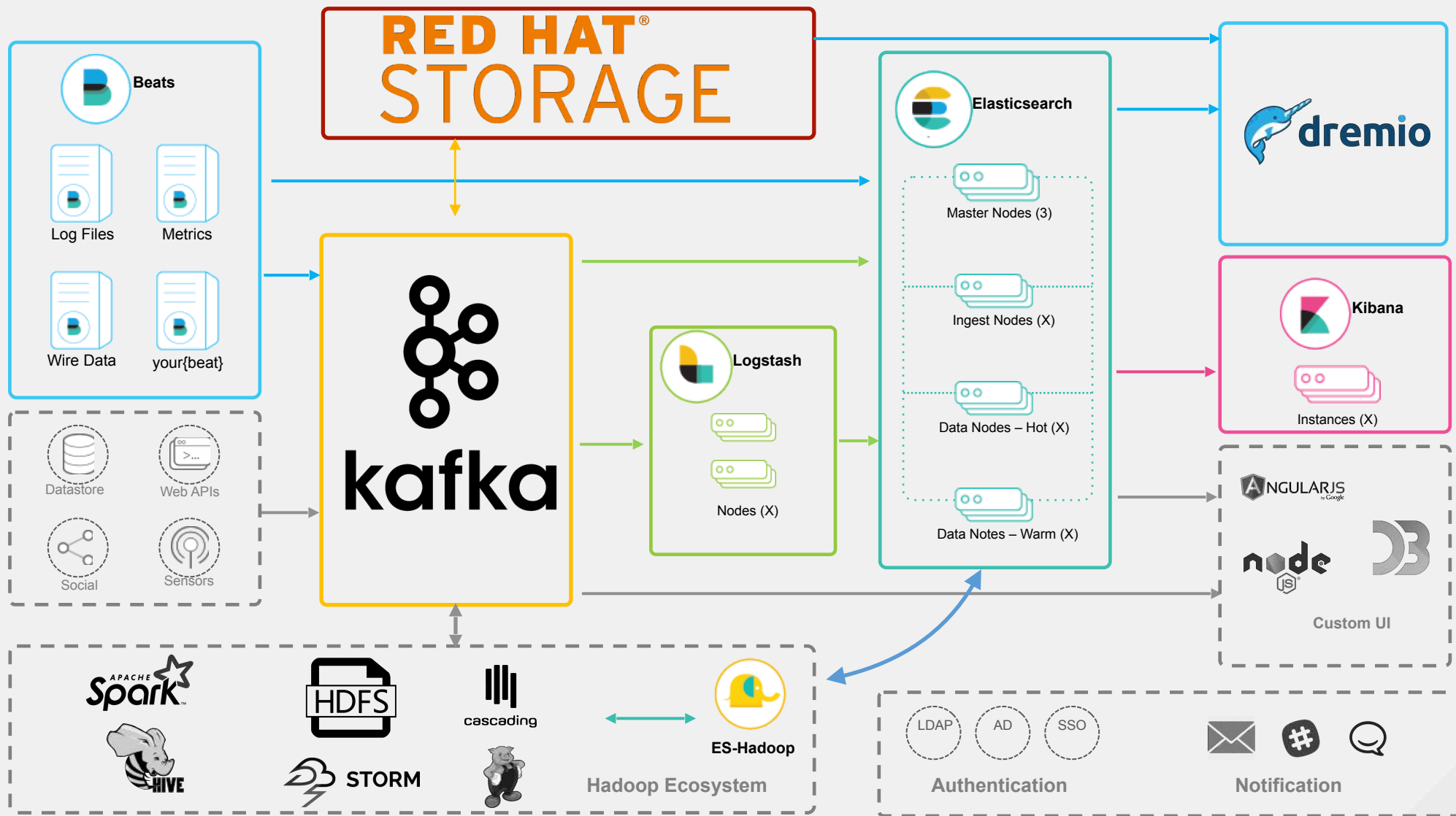
Jobs: 9»
Descendants: 0
Fields: 27

Abc _index
Abc _type
Abc _id
? votes
Abc user_id
Abc review_id

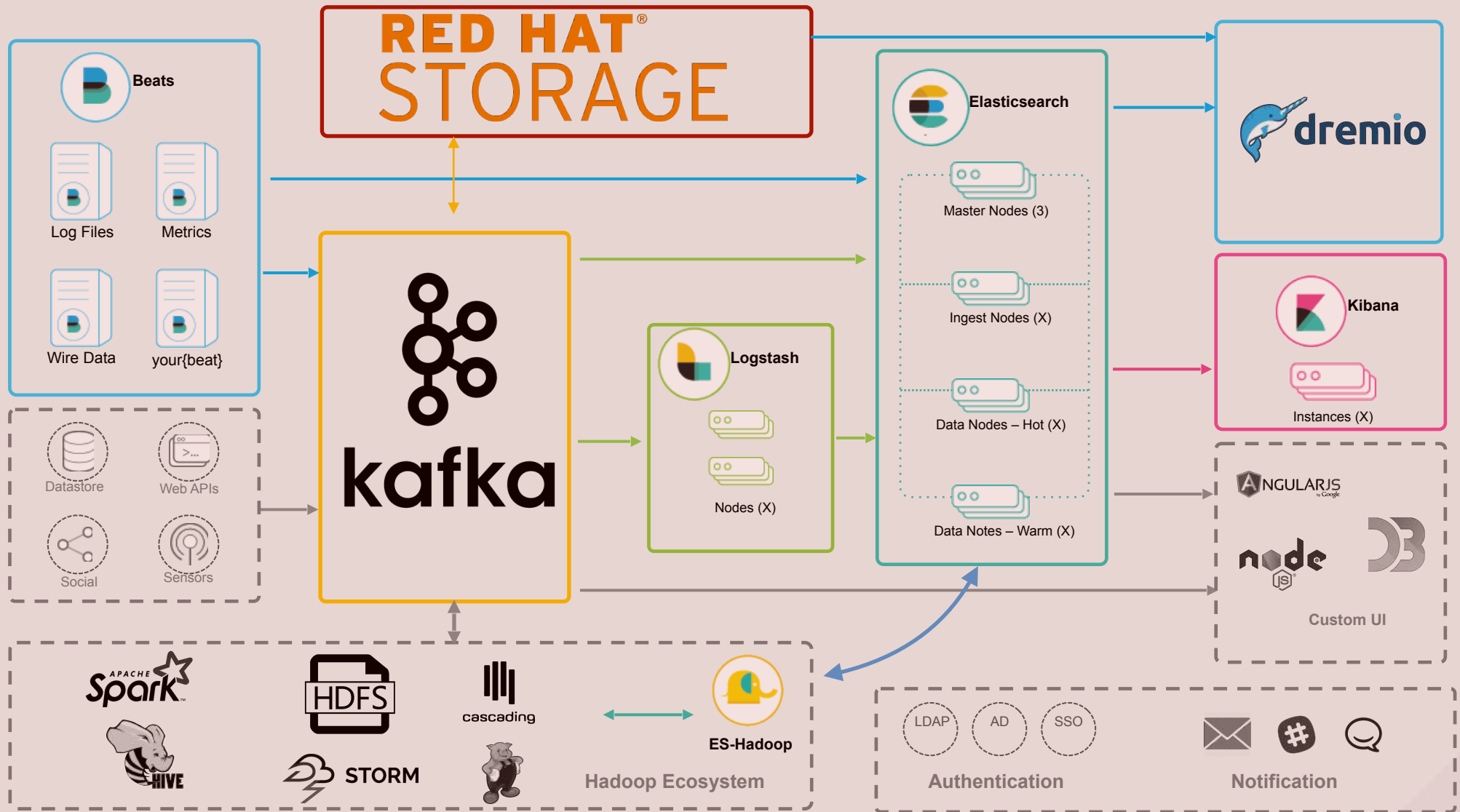








**RED HAT®
OPENSIFT**





SEACOM

the leading open source architects



redhat.

GRAZIE PER L'ATTENZIONE

Stefano Pampaloni, CEO @Seacom srl



#RedHatOSD