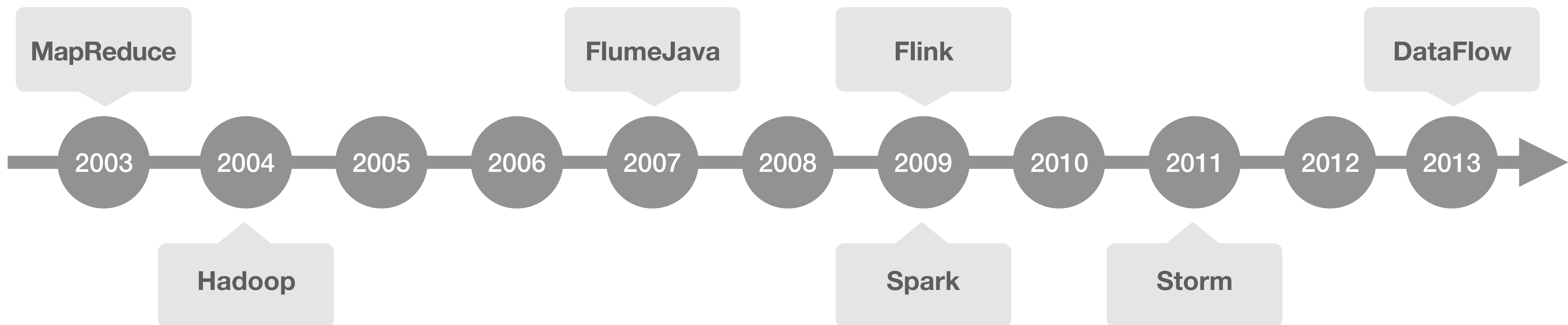




# 大数据的时间简史

为什么Flink正在成为下一代大数据计算引擎

巴真 @Flink China Meetup





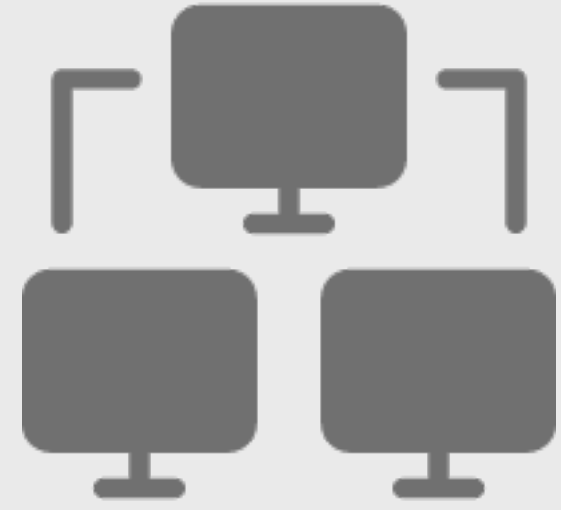
**Database**



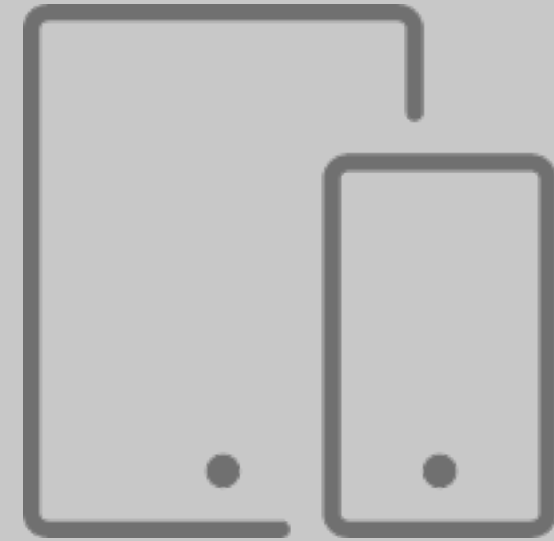
**DataWarehouse**



IT化



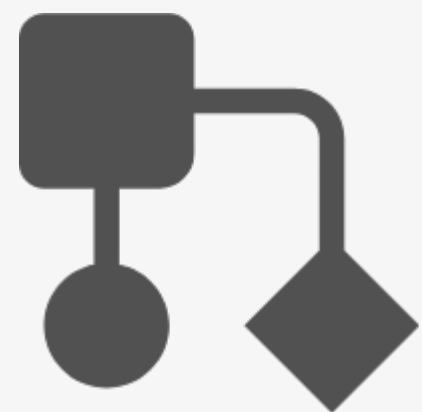
网络化



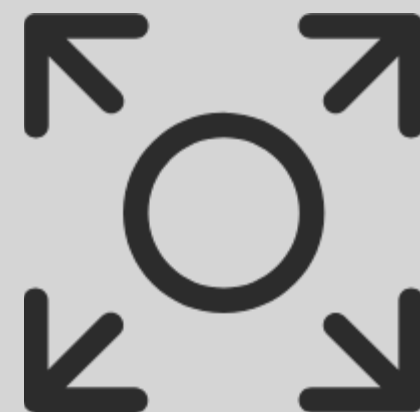
移动化



万物互联



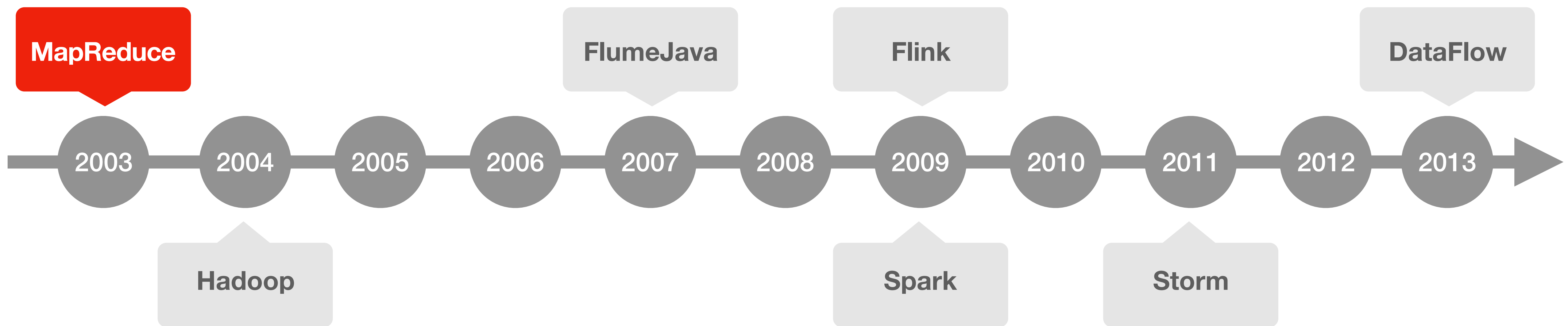
Process



Scalability



Fault-Tolerance



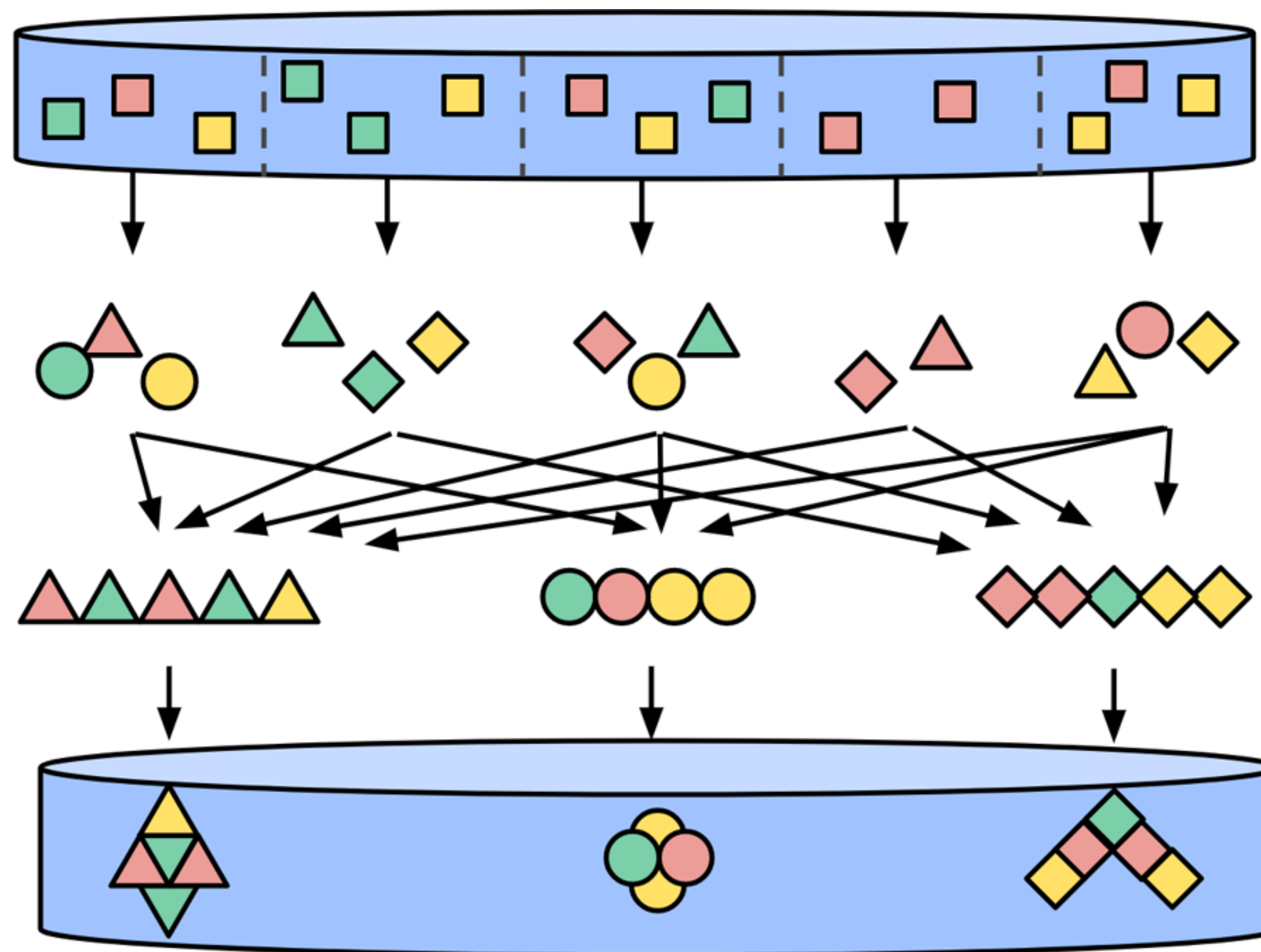


此URL需翻墙

## 《MapReduce: Simplified Data Processing on Large Clusters》

---

Google



## MapReduce模型

Data Processing is hard

数据路由/分而治之

Scalability is hard

Task FailOver/分布式存储

Fault-tolerance is hard





## 《MapReduce: A major step backwards》

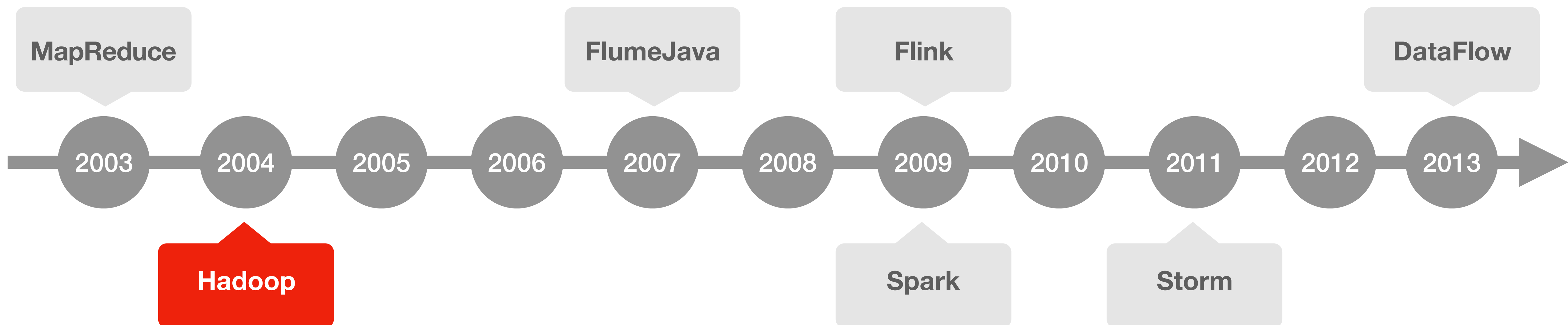
---

David DeWitt

**所谓大数据: 互联网行业快糙猛地解决大规模数据处理问题**  
为什么Google video做不过Youtube? / 如何快速拿到明星代言?

**所谓退步: 是互联网工程的妥协/Tradeoff**

Database -> NoSQL/NewSQL    Queue -> Kafka/Kinesis/RocketMQ





## Hadoop Official Site

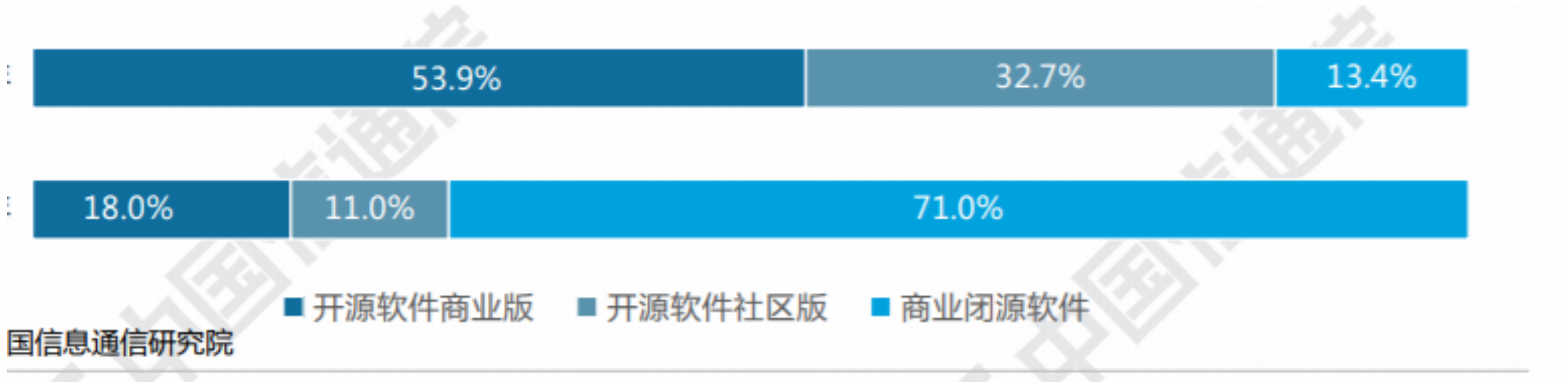
---

<http://hadoop.apache.org/>

### 互联网再一次山寨了山寨品

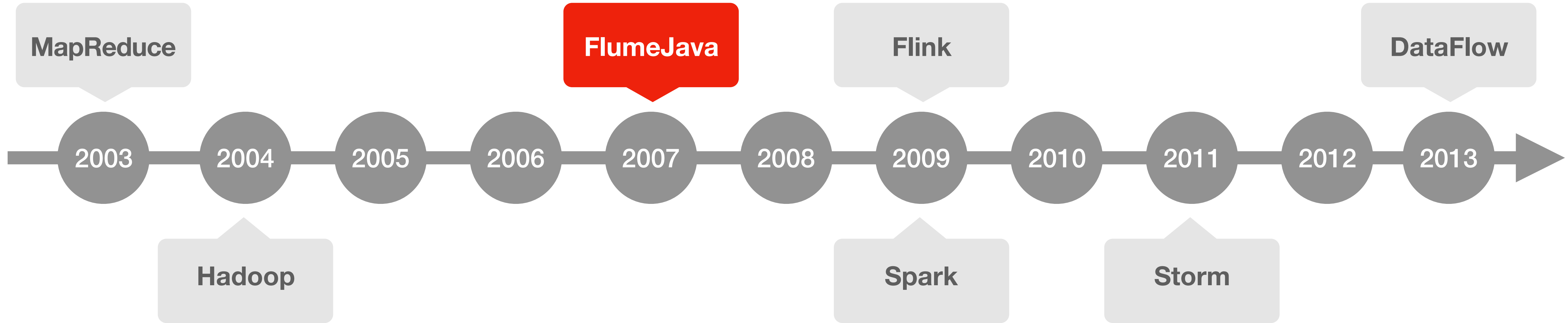
Hadoop折扣实现了大量Google的Paper，包括调度、存储。  
技术生态圈再次被开源占据，各个领域开源系统日益成为工业标准。

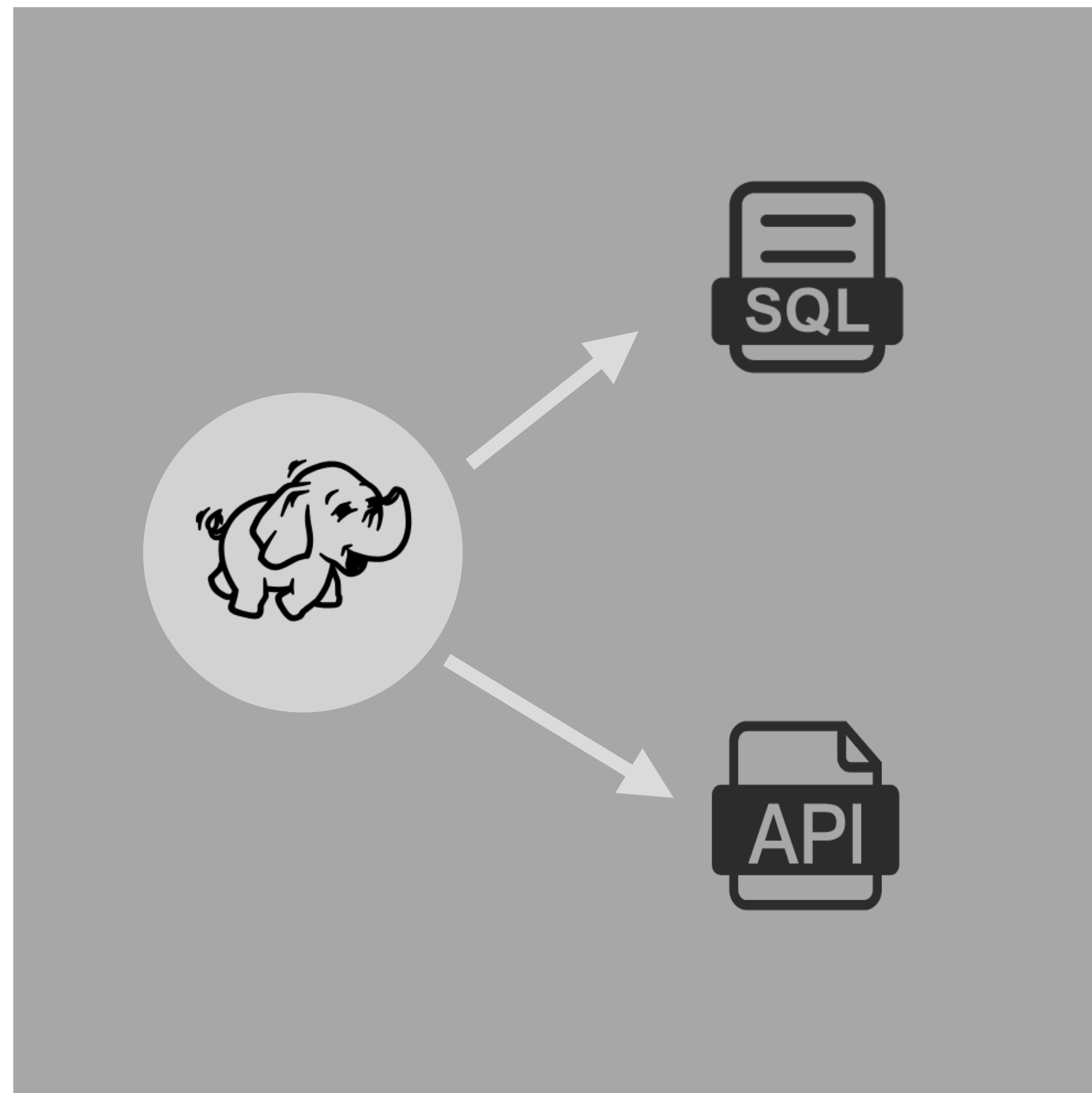
HasS巨大的市场机会，全球HaaS市场到2020是50B，约3300亿，中国HaaS市场到2020是578亿  
开源商业化版本对于企业化客户接受度高，超过50%。



# 非开源不用!

企业用户采购大数据系统只认可开源大数据  
50%+企业首选使用开源软件企业版本服务。



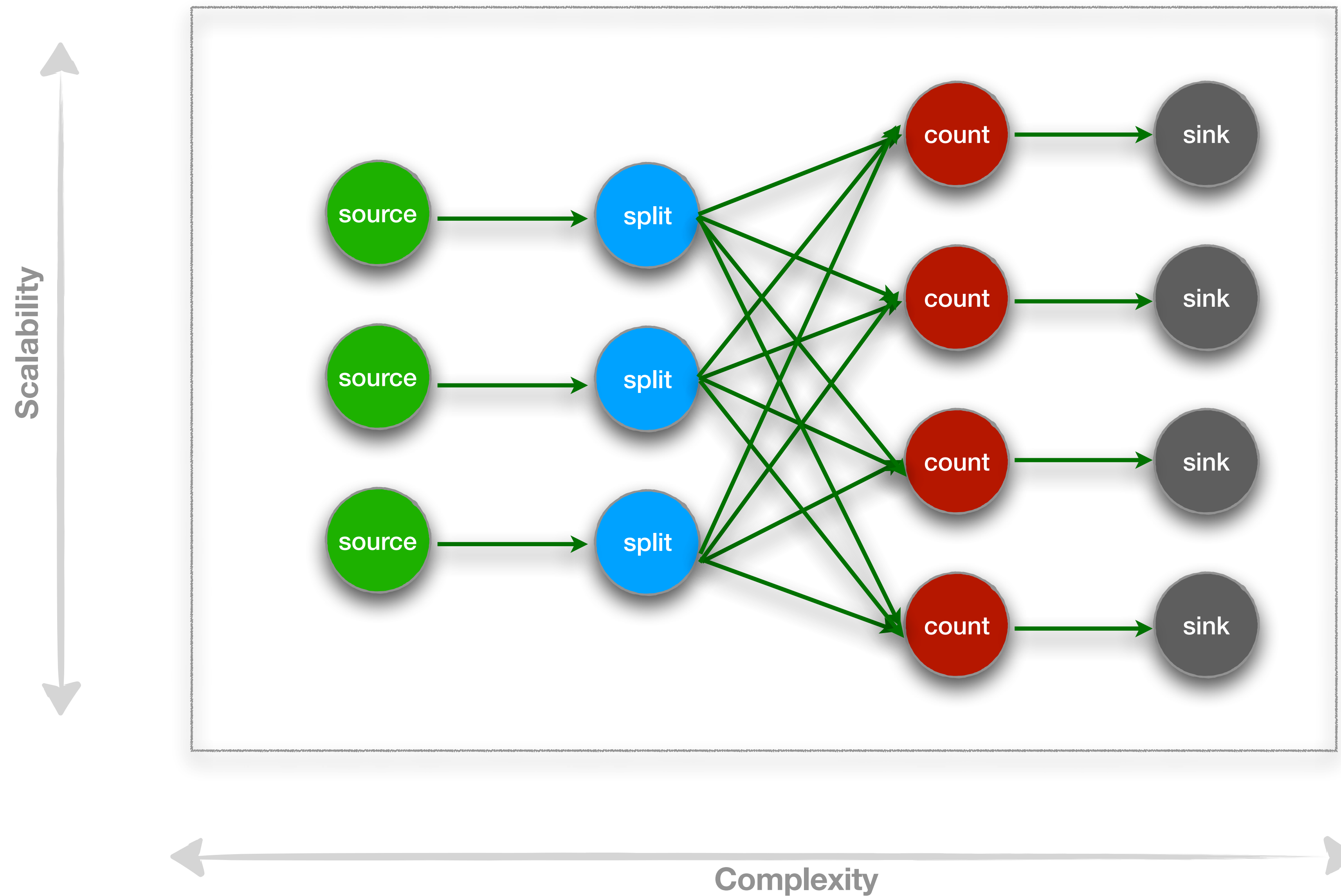


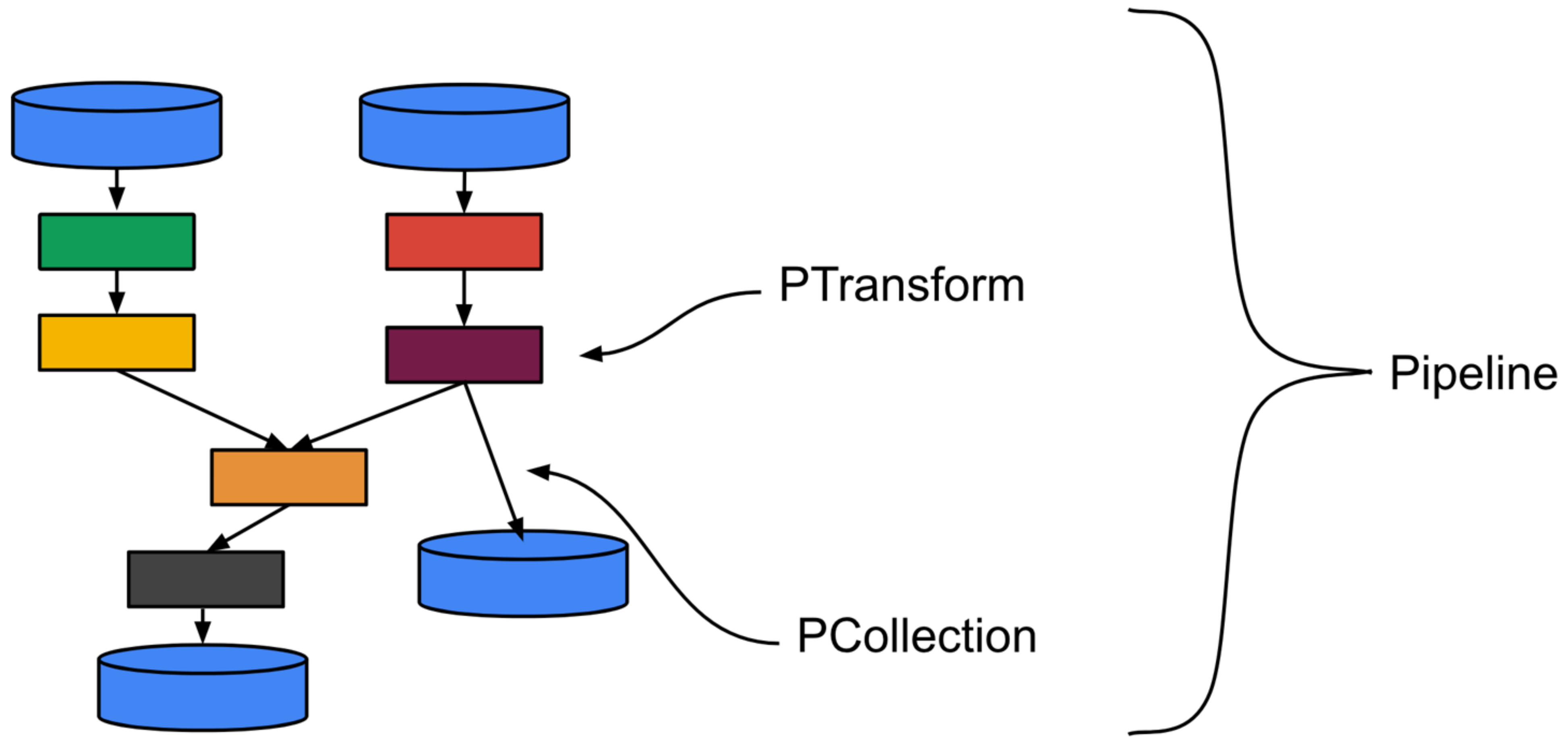
## MapReduce抽象层次太低，导致

1. 业务实现困难，复杂业务动辄上万行代码
2. 系统优化困难，黑盒导致框架无法优化

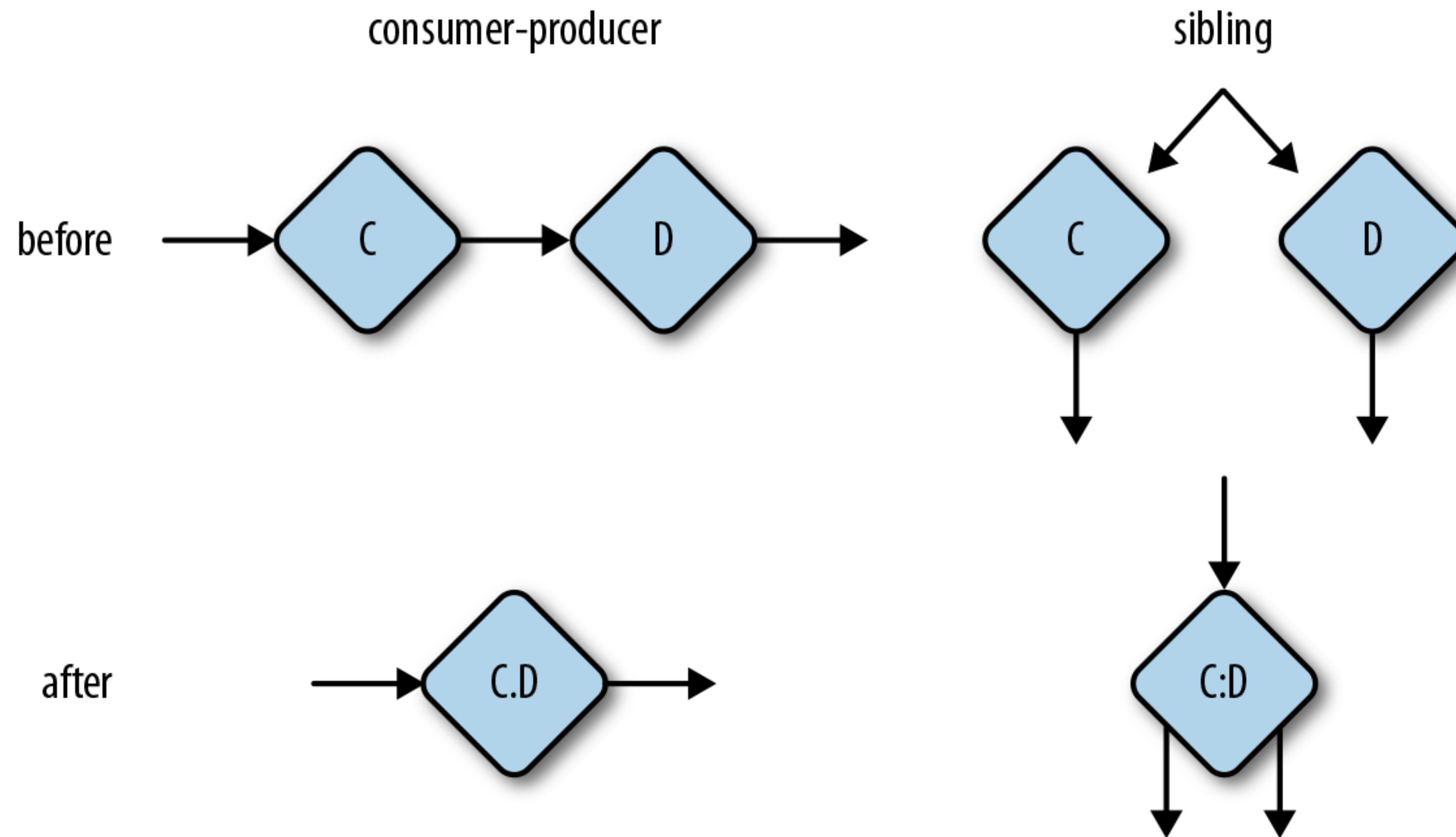


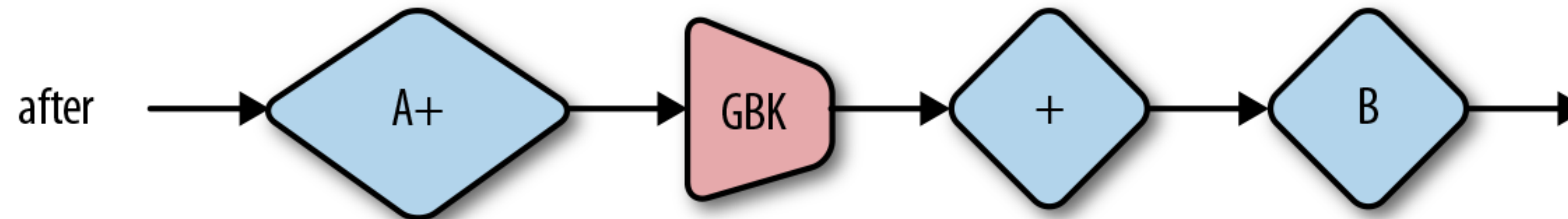
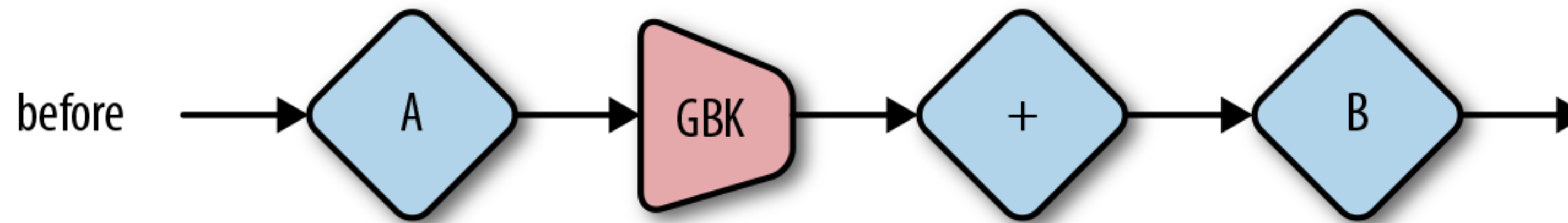
# WordCount Job











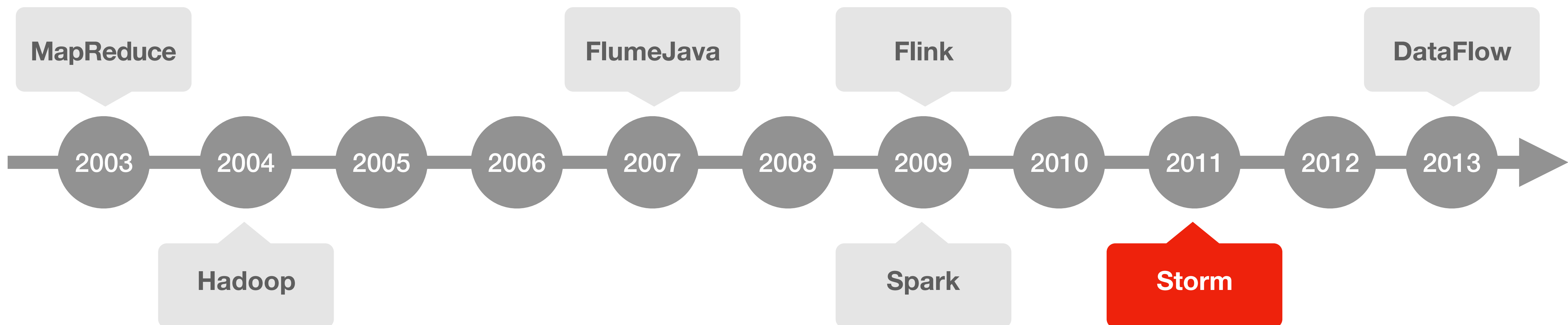


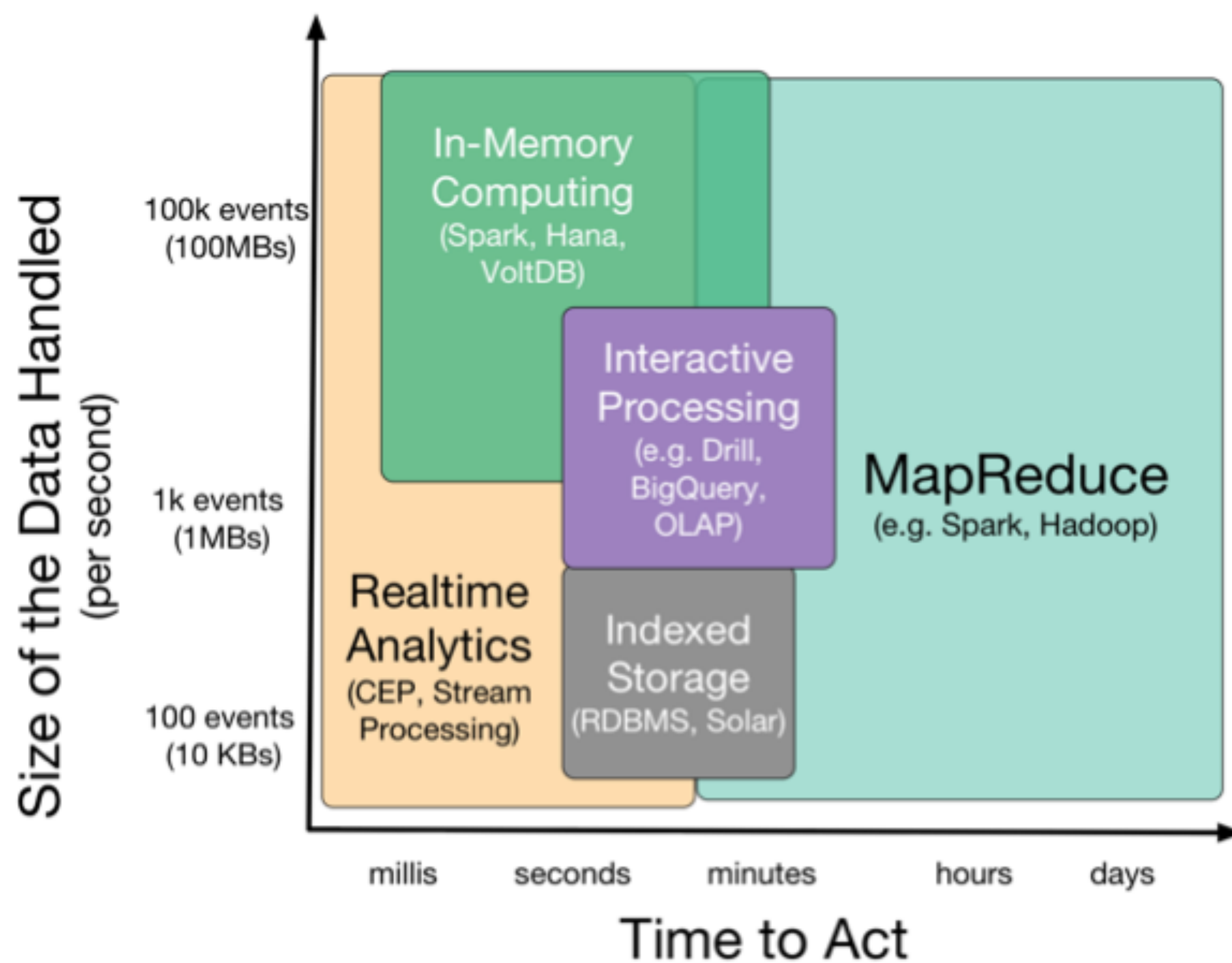
此URL需翻墙

## 《Flume Java: Easy, Efficient Data-Parallel Pipelines》

### FlumeJava有效解决MR抽象层次太低的问题

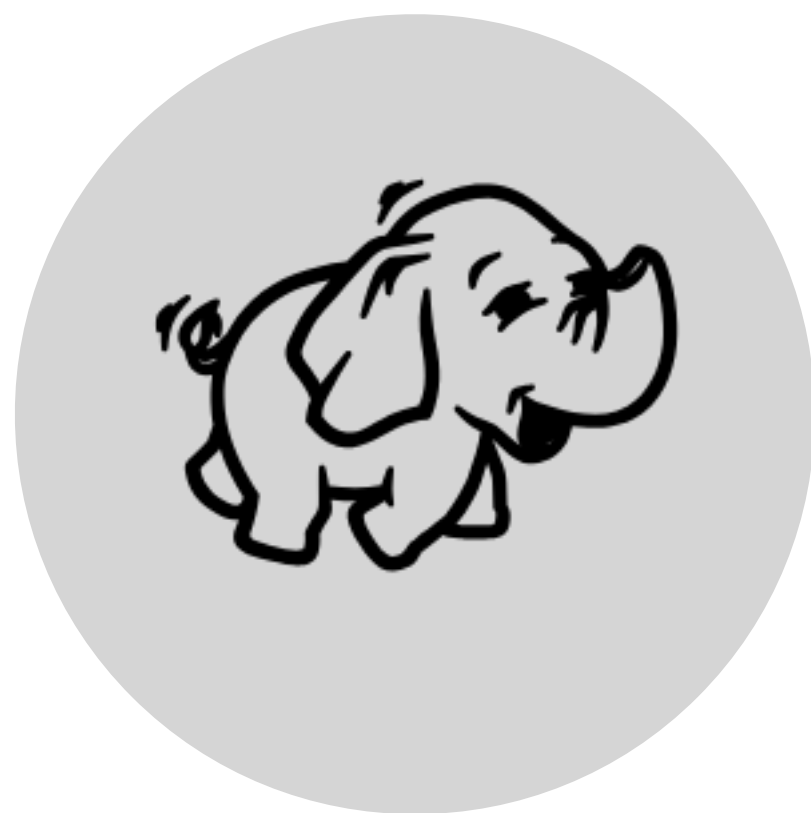
1. 业务可以使用较少量的代码实现曾经MR成千上万行代码
2. 系统可以根据用户的业务CODE实现优化





## Batch被山寨了，Stream还远吗？

1. 数据的价值随着时间增长迅速降低。
2. 数据流实时产生，实时处理



**Apache Hadoop**

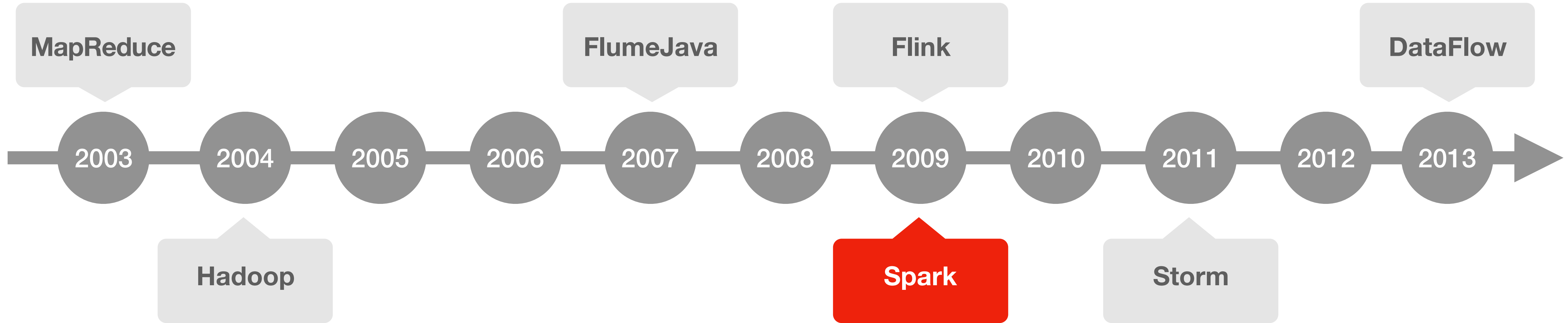


**Apache Storm**



# Flink: 下一代流式处理系统

——为什么Flink要比Storm/Spark更加优秀？







《An Architecture for Fast and General Data Processing on Large Clusters》

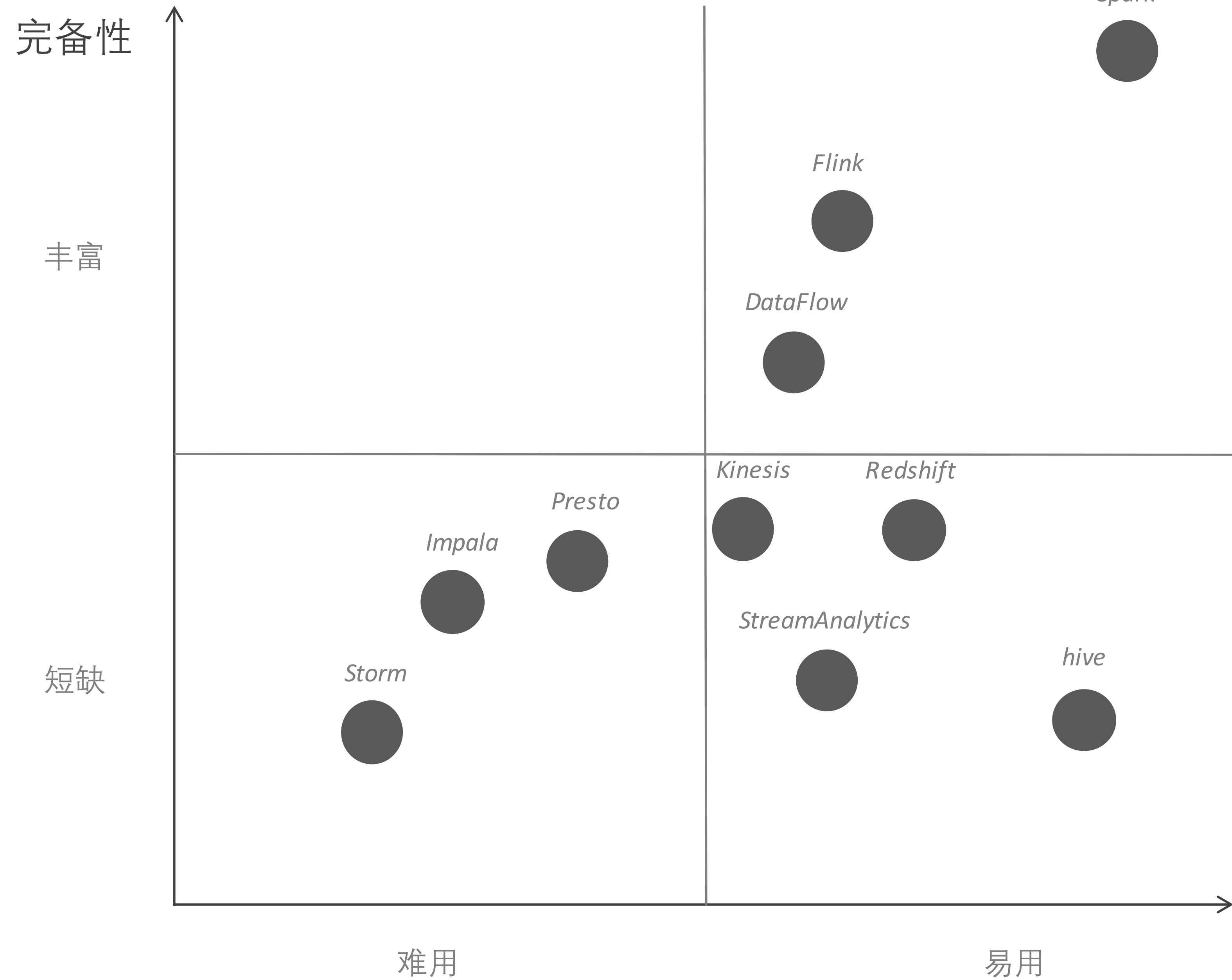
## 涵盖大数据入口处理业务的一站式大数据计算引擎

1. 1/3的社区调查用户认为一套引擎解决大部分问题是Spark重要优势
2. Spark是Hadoop/Storm的下一代大数据处理引擎



---

计算类型多种多样，但入口业务仅此数种



### Spark优势明显

功能完备: Batch/Stream/ML  
简单易用: API丰富/文档较丰富/生态对接  
社区活跃: 运营出彩、大数据最活跃社区

### 云产品系列

产品化: 产品化成熟，易于上手  
稳定性: 稳定性好，商业托管  
社区小: 产品封闭，不易形成社区

### 开源软件

社区活跃: 产品开放，易于形成社区  
稳定性差: Bug多，不稳定  
产品粗糙: 产品化初级，上手较难



Spark期望一套软件覆盖主要计算模型，但实际覆盖不完整

21%用户认为Spark Streaming在功能(集中在窗口)和时延(亚秒)等比不上Flink，增量流式业务考虑使用Flink

17%用户认为Spark ML部分落后，包括提供更多算法、对接TF，部分业务迁移到TF框架运行

25%的用户认为当前缺乏好用、内置的上层平台，包括开发界面、工作流调度，用户使用Spark同样需要重新搭建平台系统

稳定性/调优/排错 仍未解决

31% 用户吐槽Spark集群不稳定，经常性OOM导致业务产出不稳定

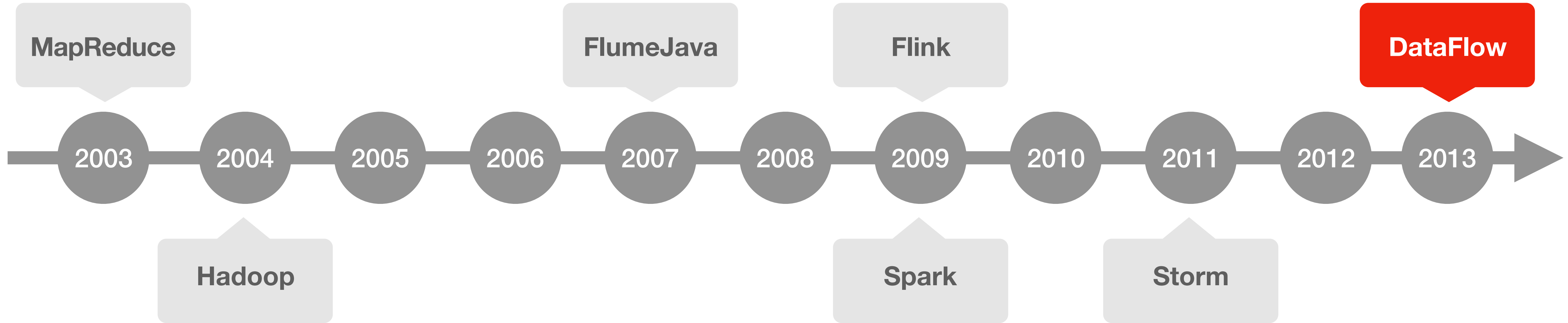
另外，几乎同样客户群体(说明都是深入生产使用Spark用户)都认为Spark作业排错、调优困难，易用性不够

中文资料/社区严重缺乏，未能形成有效组织

30%用户吐槽当前文档、案例过少，特别在调优、排错方面，用户往往不知所措

用户同时认为相关中文资料相比更少，时效性也落后英文社区太多

中美语言差异导致中国市场更加空白

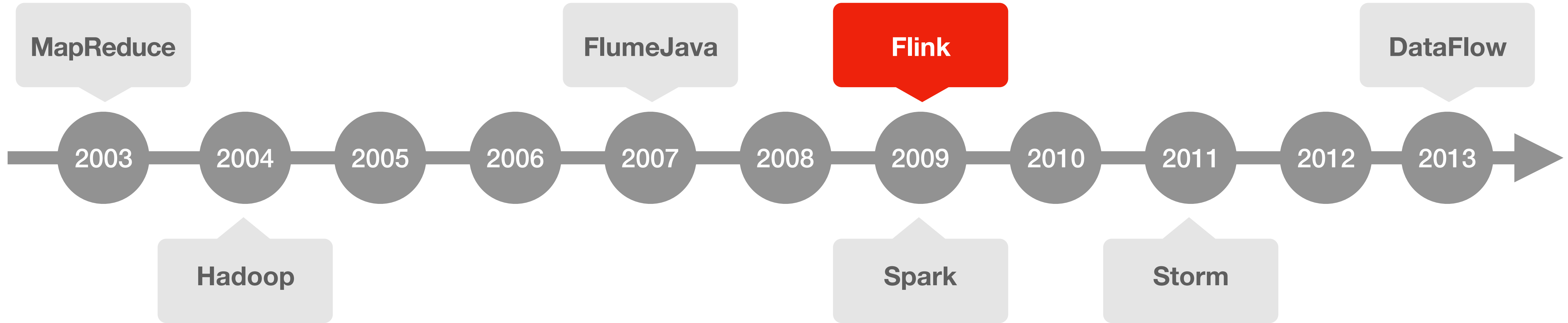




The Dataflow Model: A Practical Approach to Balancing Correctness, Latency, and Cost in Massive-Scale, Unbounded, Out-of-Order Data Processing

### 批流统一执行引擎的大数据计算引擎

1. 具备现代流式处理特征的批流统一大数据处理引擎
2. 同一套Runtime，同样一套QP



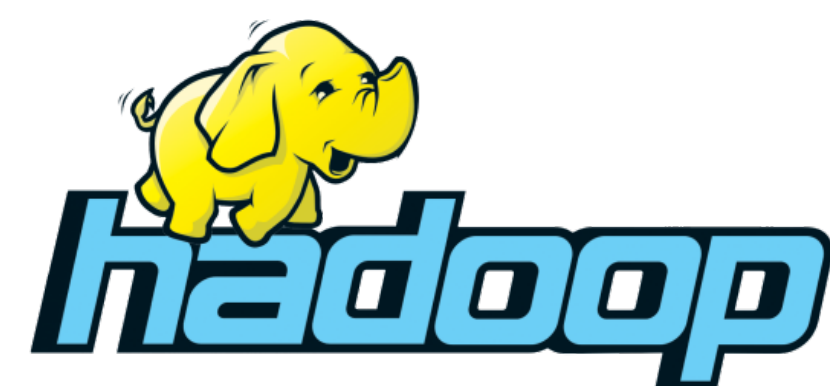


《Apache Flink™: Stream and Batch Processing in a Single Engine》

## 批流统一执行引擎的开源大数据计算引擎

1. 具备现代流式处理特征的批流统一大数据处理引擎
2. 同一套Runtime，同样一套QP





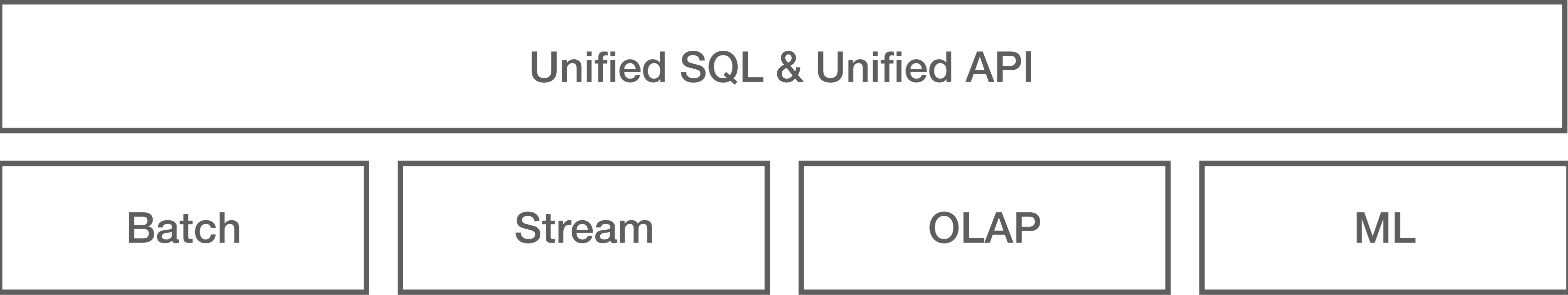
第一代

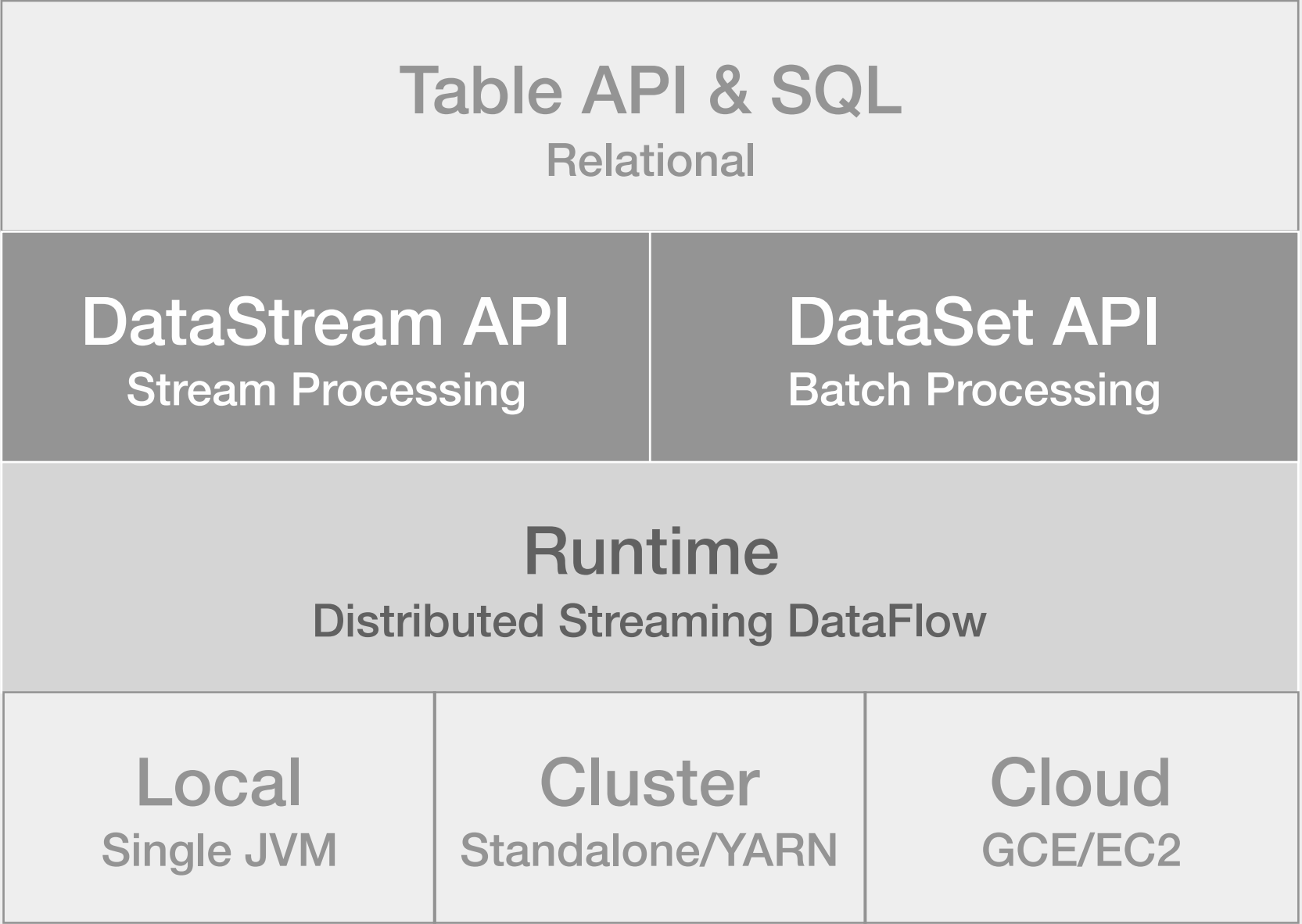


第二代

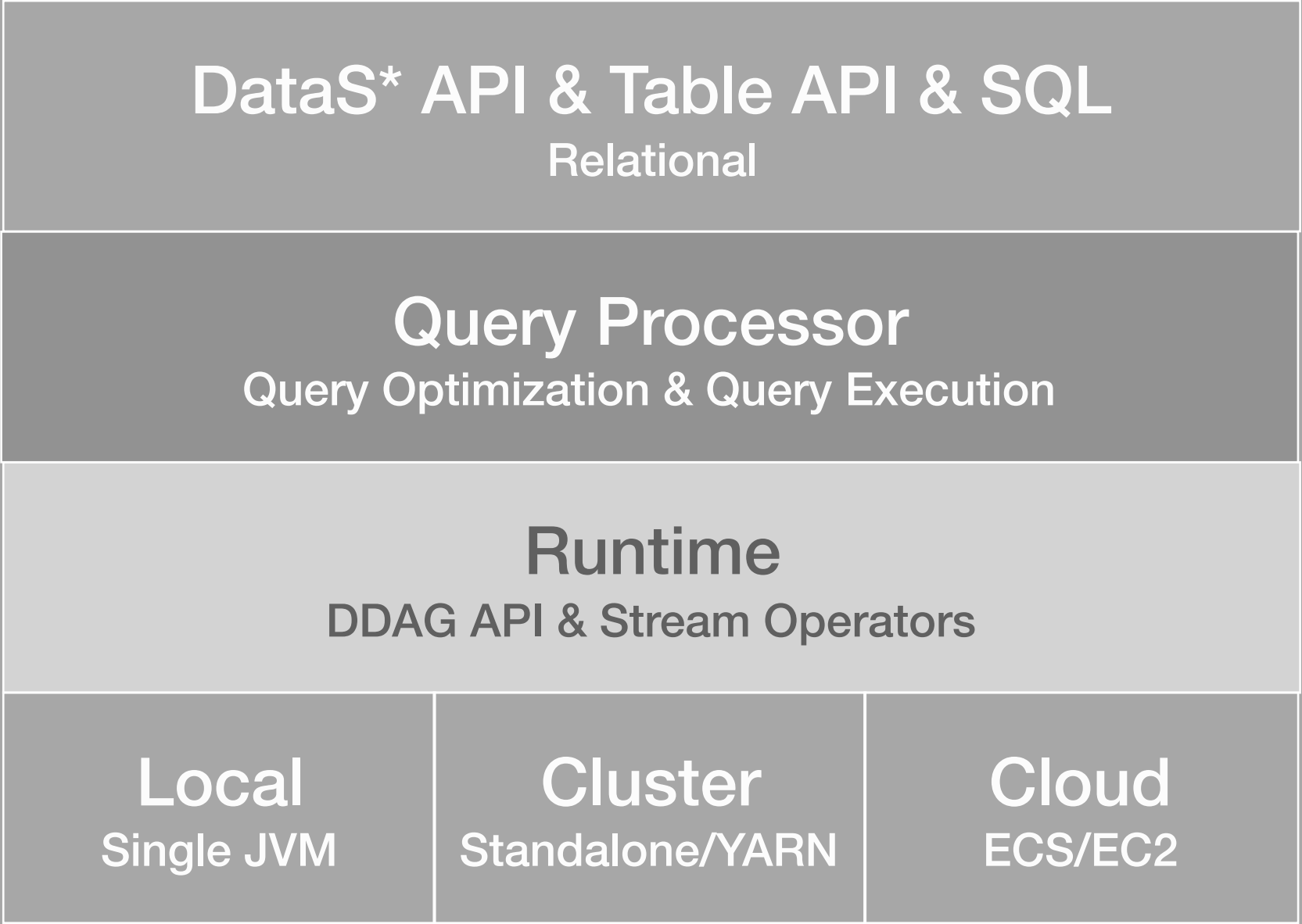


第三代

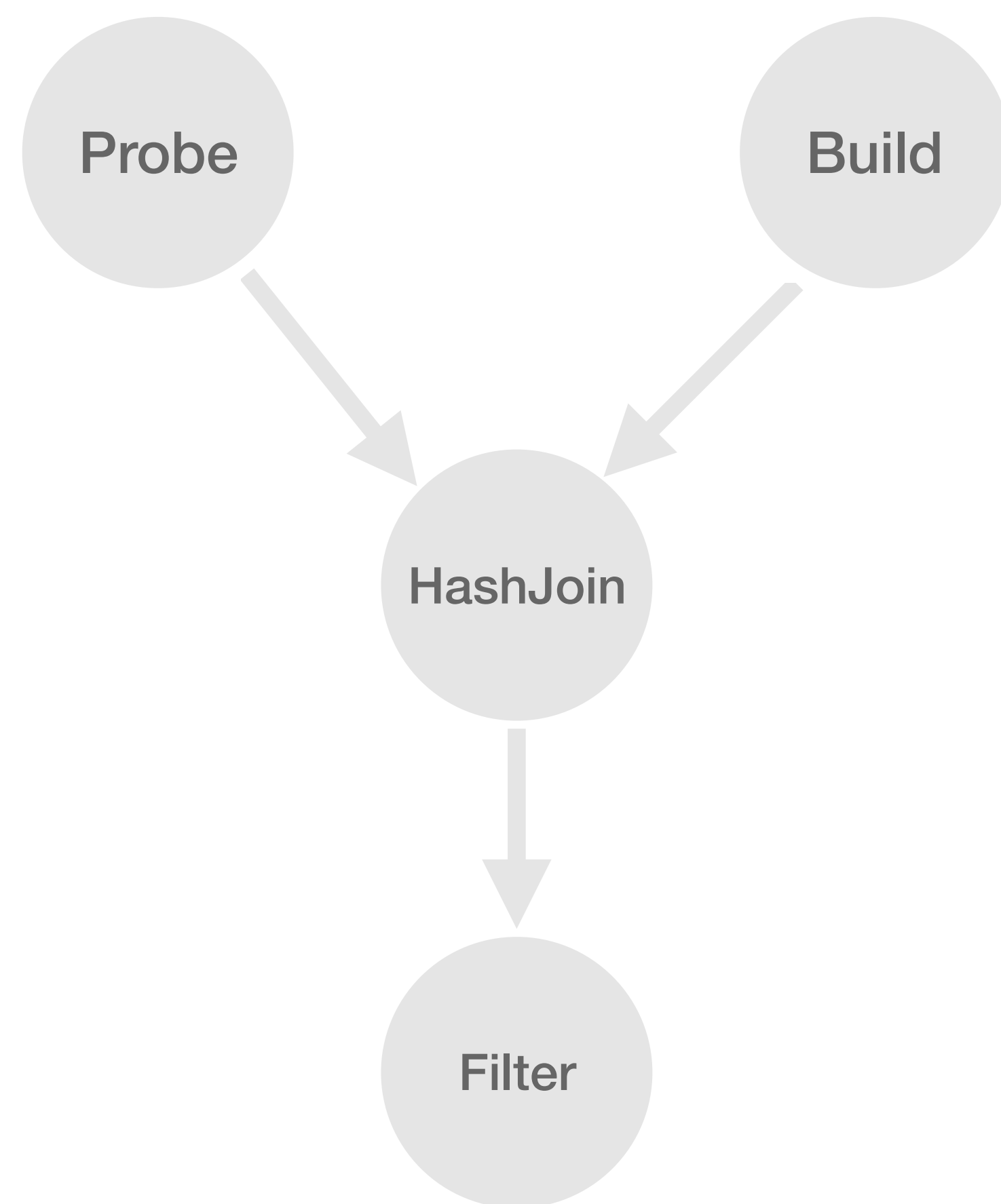




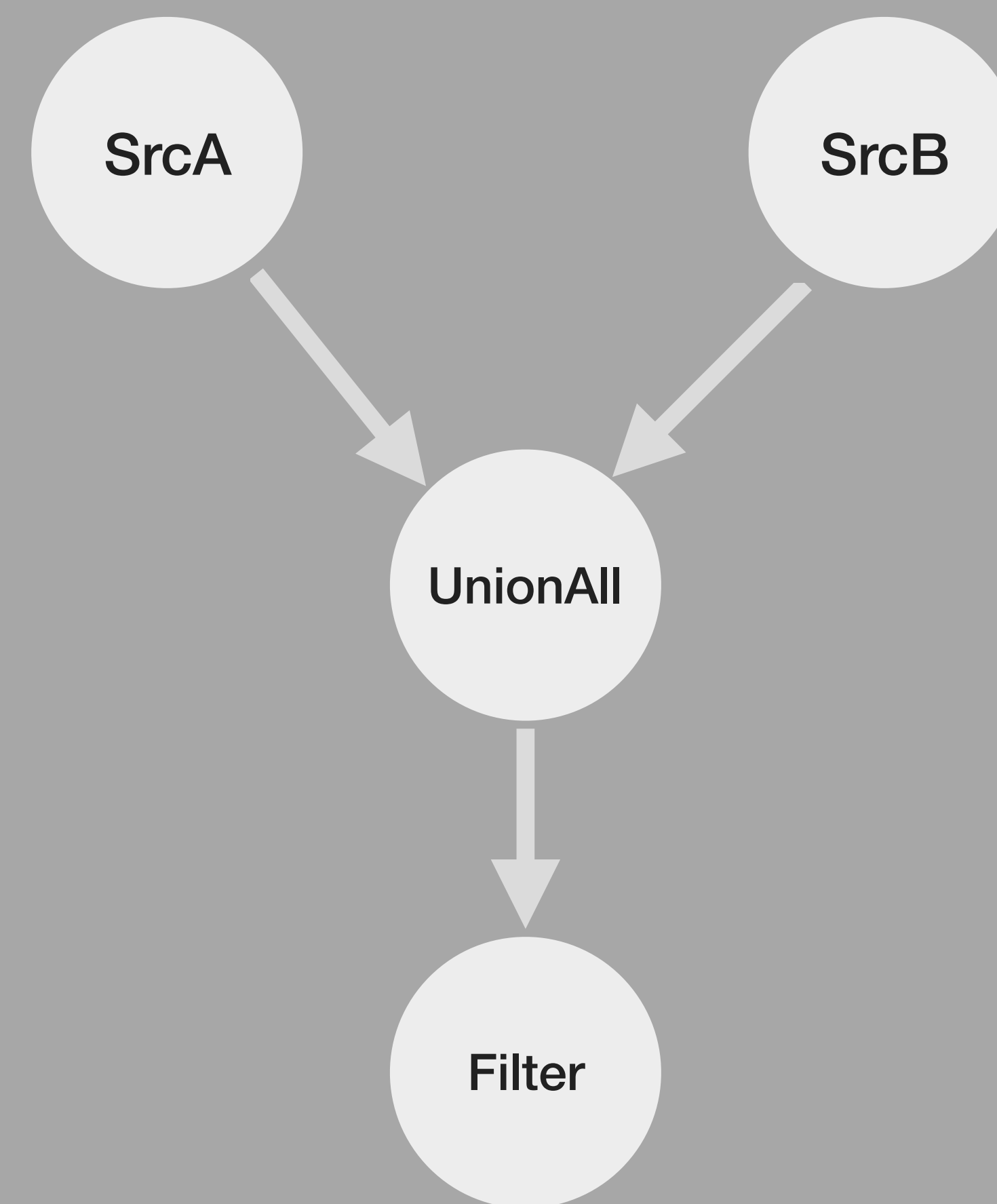
Old Design



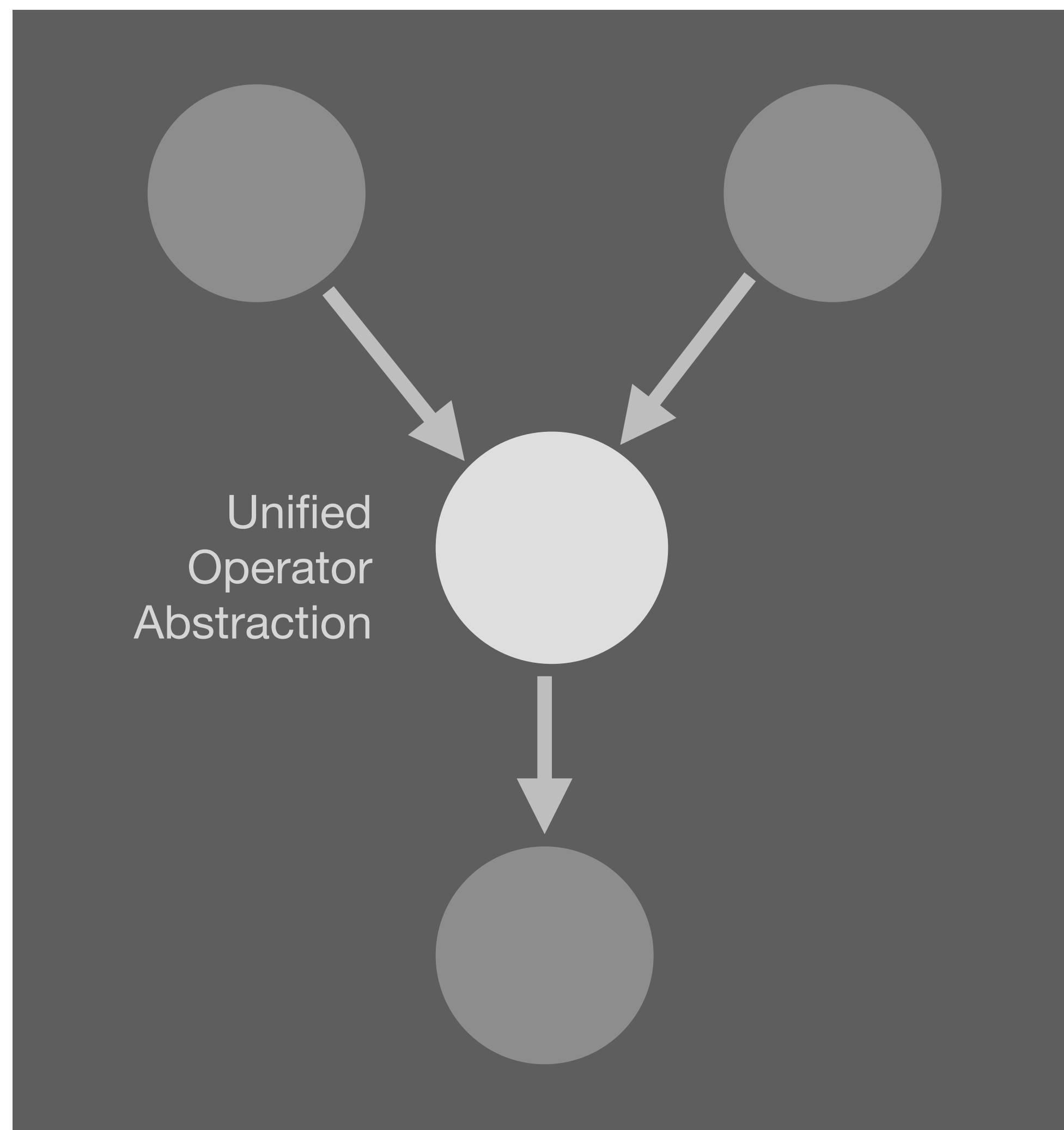
New Design



Batch Model



Stream Model



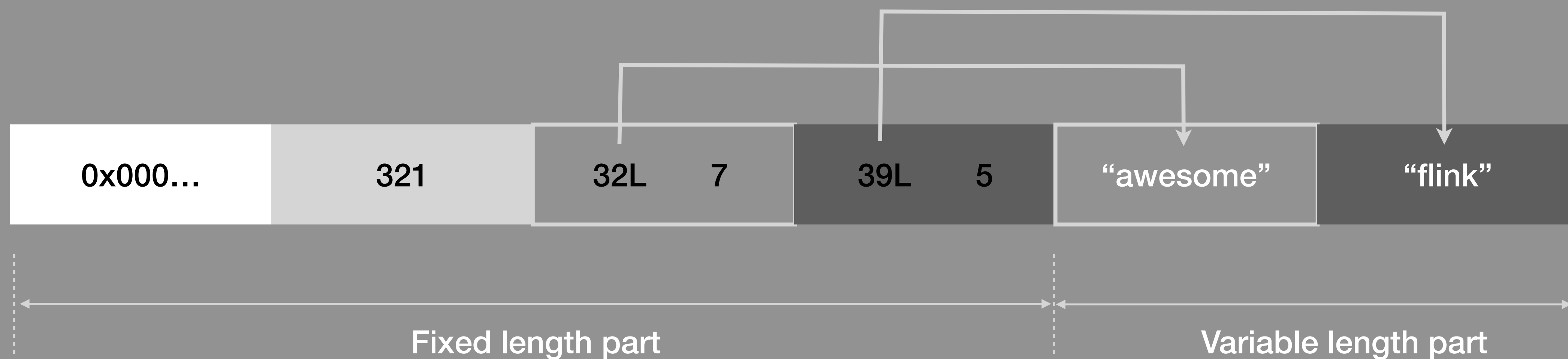
## *Unified Operator Abstraction*

- *Operators can choose inputs*
- *Operators can be chained easily*



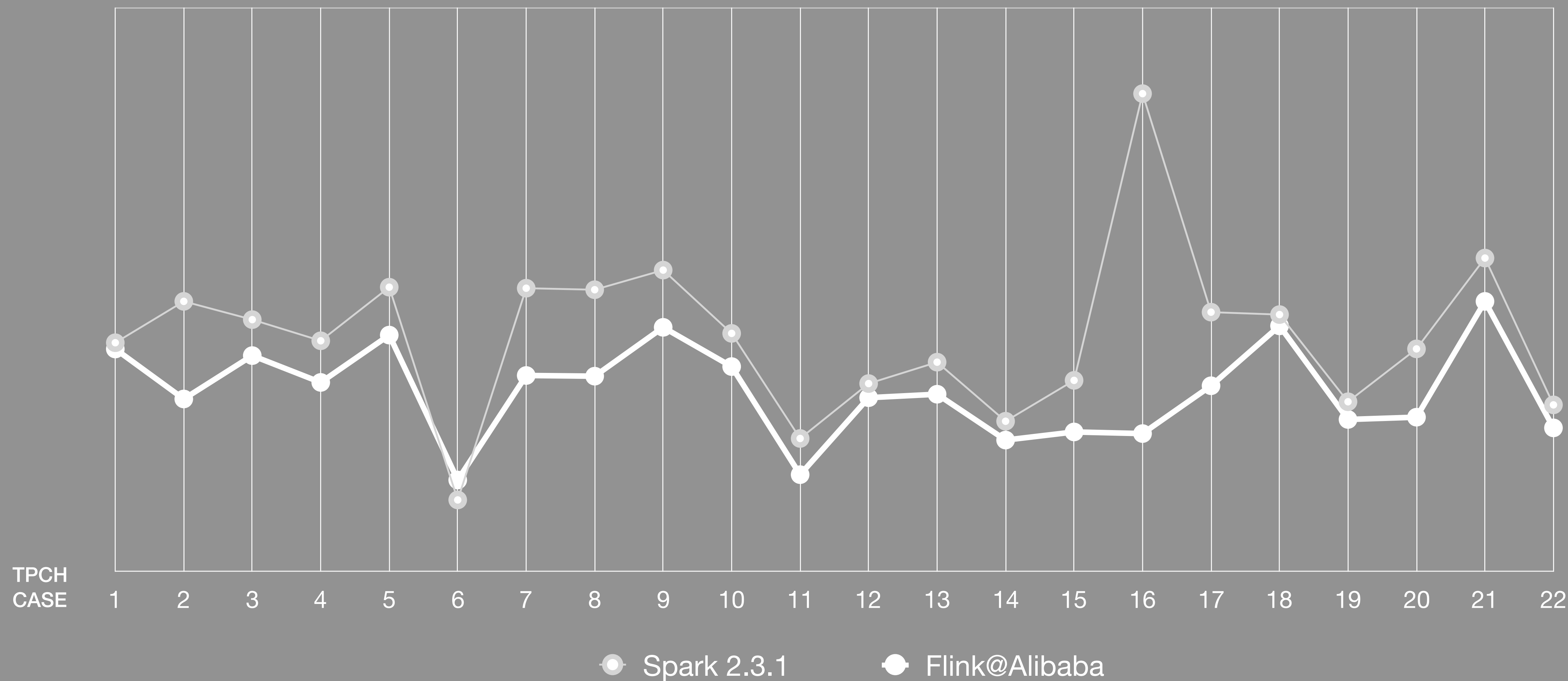
## Record Format

- Introduced new row format: BinaryRow
- Tight integration with memory management
- Avoid deserialization cost





# TPCH Performance *(the Lower, the Better)*





20+

BU

3K+

Job

40w

Core

7K+

机器

4.7亿

总体峰值

1.8亿

大屏峰值

1s

延迟

74%

利用率





Apache Flink



Realtime Compute  
Alibaba Cloud





未来基础设施

Alibaba/Amazon/Microsoft/Google 正在成为未来基础设施





**Flink Forward China**

**12月20-21日 @北京 国家会议中心**

亚洲区第一场Flink官方大会，预计3000+参会人员

**Flink中文社区大牛&公司悉数到场**

华为、腾讯、滴滴、饿了么、阿里等共同举办

**Call For Keynotes&Sponsors**

召集keynotes&Sponsors



# Thanks

