

Real-Time Analytics: Going beyond stream processing with Apache Pinot

David G. Simmons
Head of Developer Advocacy

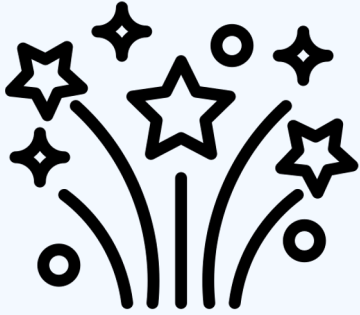


What is Real-Time Analytics?

Real-time analytics is the discipline that applies logic and mathematics to data to **provide insights for making better decisions quickly**.

Gartner[®]

Events -> Insight -> Action



Events

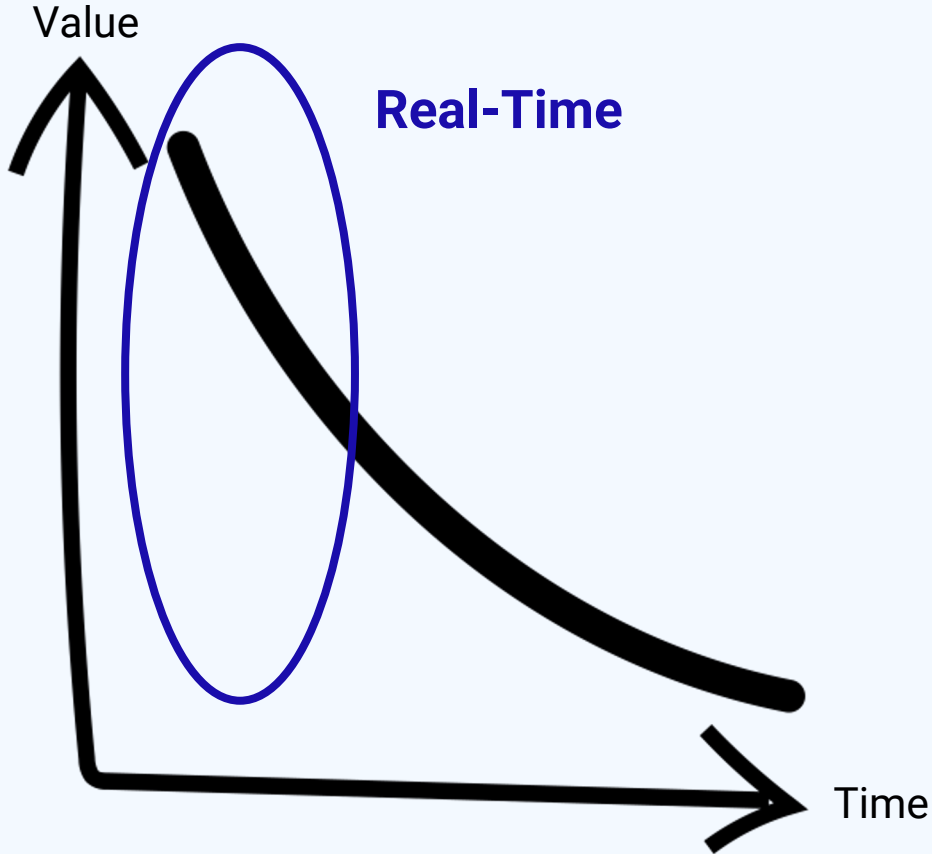


Insight



Action

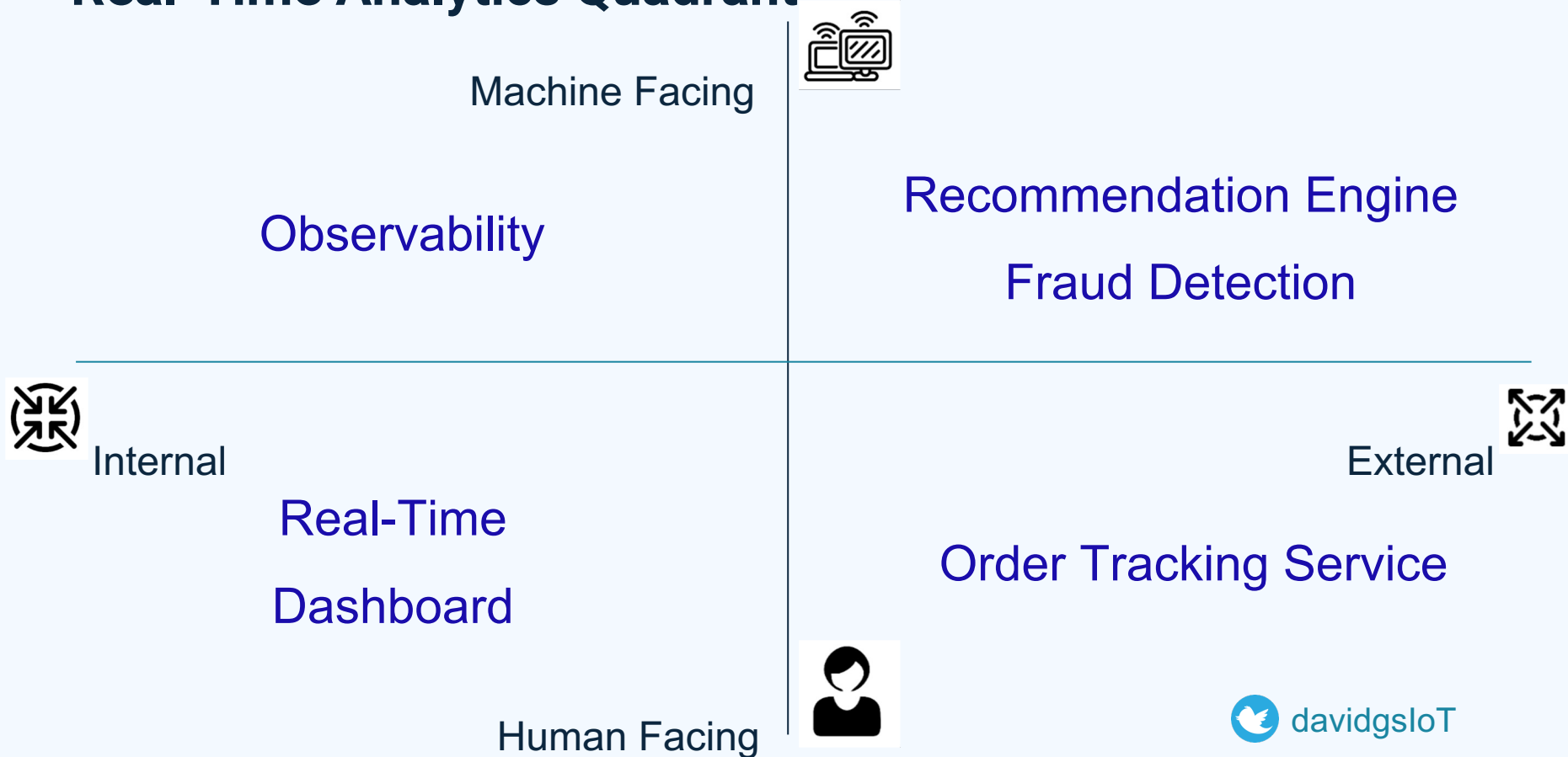
The value of data over time



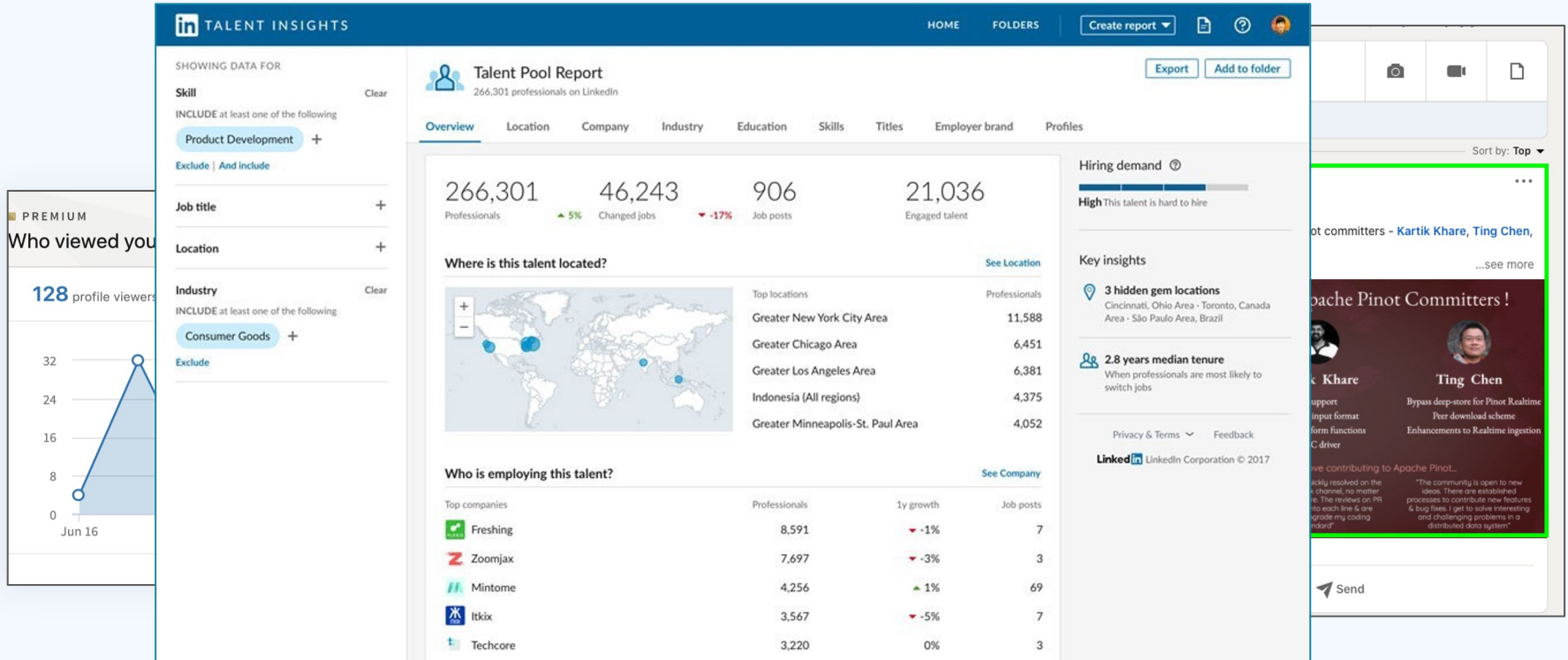
Who's interested in this data?

- Analysts
- Management
- Users

Real-Time Analytics Quadrant

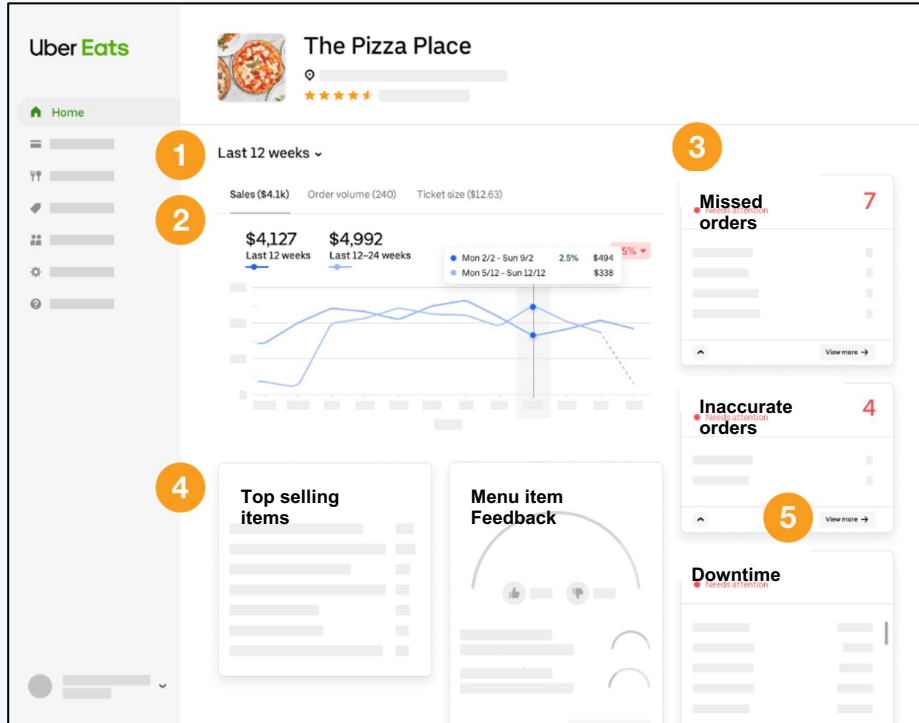


Examples of Real-Time Analytics



Examples of Real-Time Analytics

Uber Eats



Total users	500,000+
QPS	100s
Latency SLA	< 100 ms p99th
Freshness	Seconds - Minutes

Examples of Real-Time Analytics

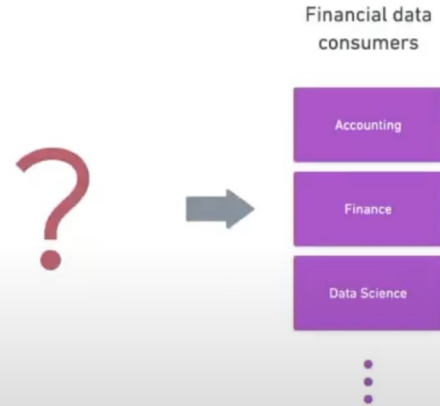
Stripe's Real-Time Analytics Landscape

Unique Challenges and Opportunities

- High precision/accuracy requirements
 - Aggregations must be exact
 - Many small-unit transactions plus small currency units vary
- Transactional-level granularity
- Reports must be reproducible
- Frequent manual action required
- Strict Compliance & Security Requirements
 - Regulators and auditors are an end-customer

stripe

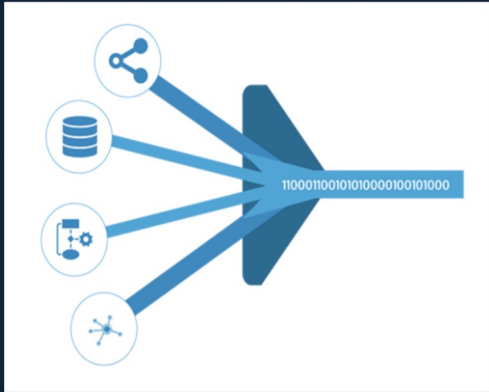
... solving?



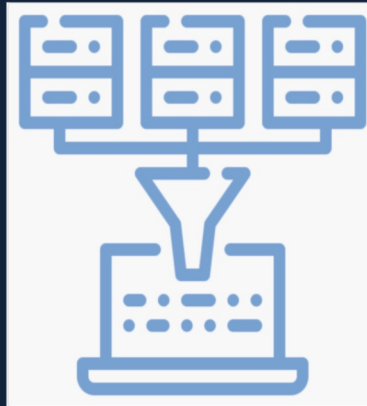
Source:
Peter Bakkum, Engineering Manager @Stripe Financial

Properties of Real-Time Analytics Systems

Speed of Ingestion



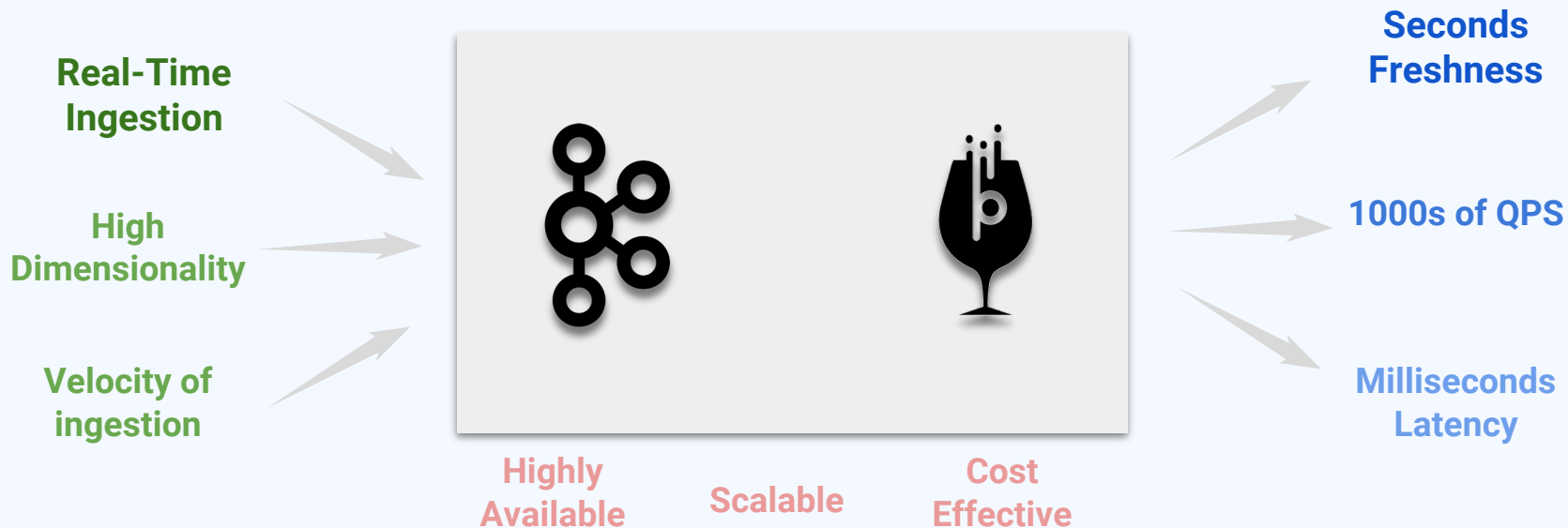
Speed of Queries



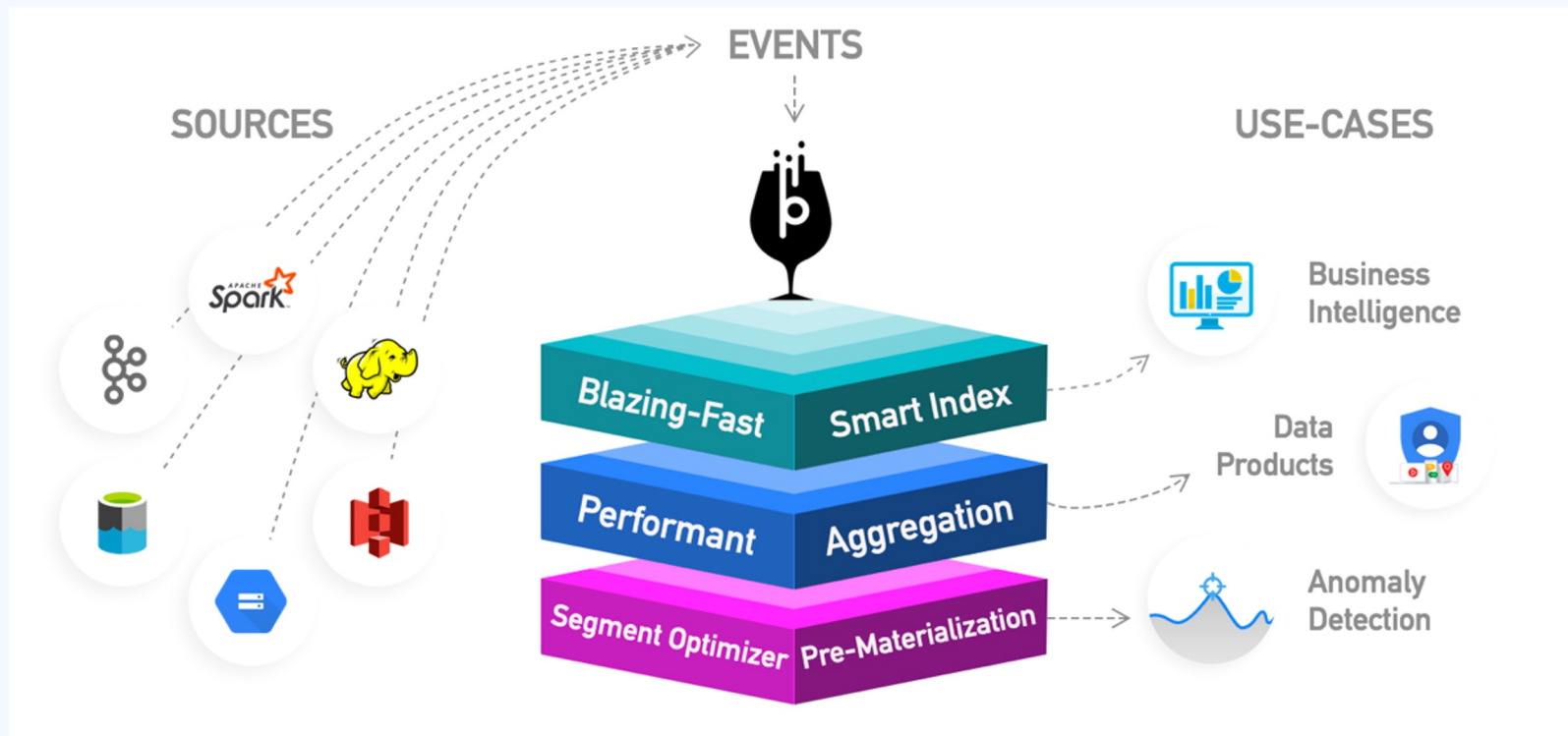
Scalability



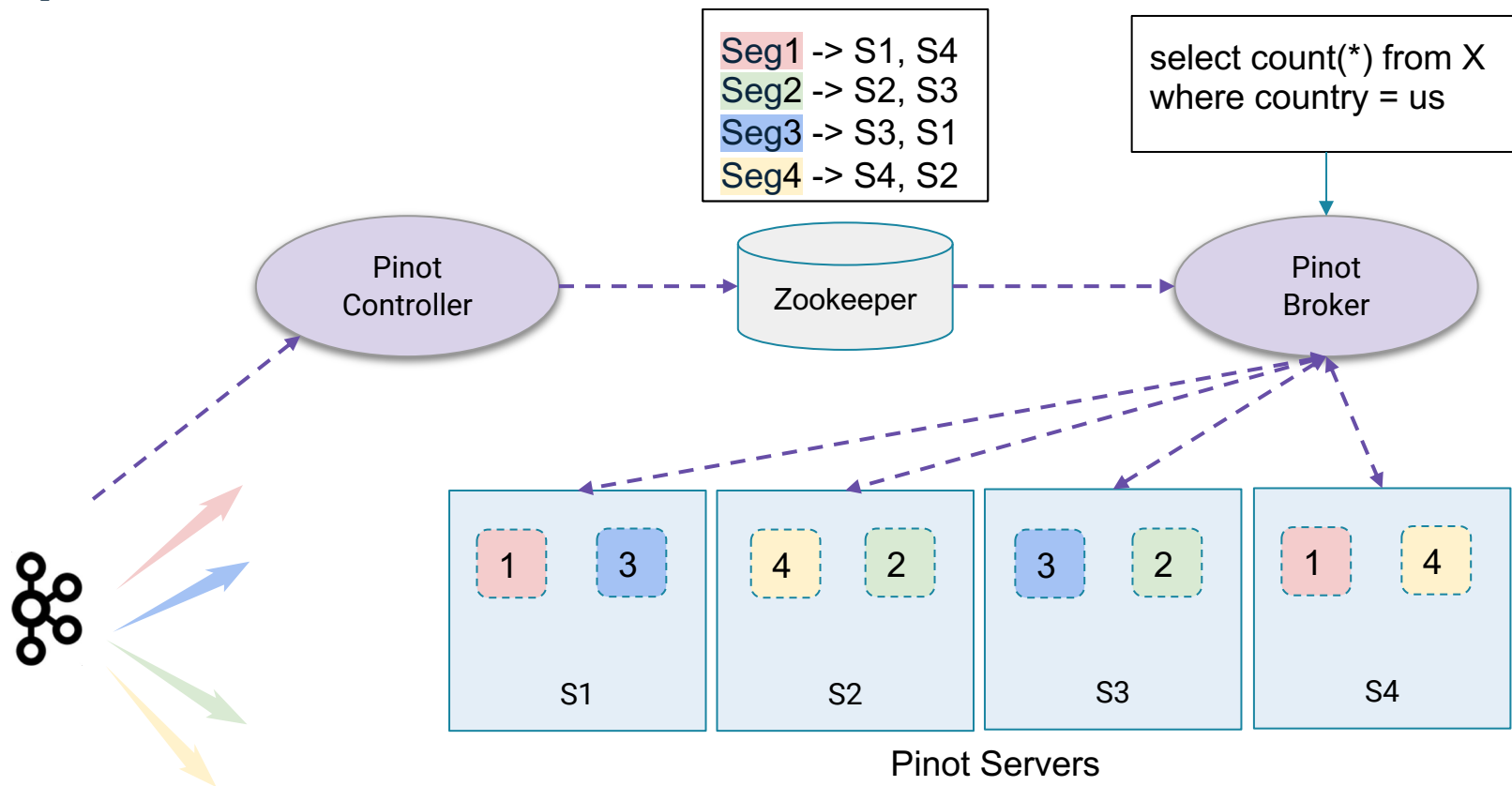
Building a User-facing Real-Time Analytics System



What is Apache Pinot?



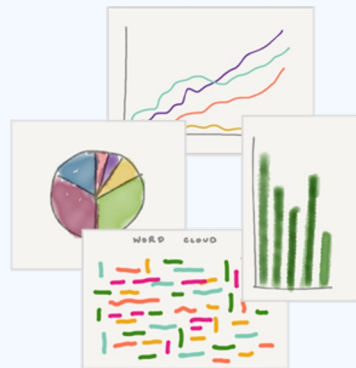
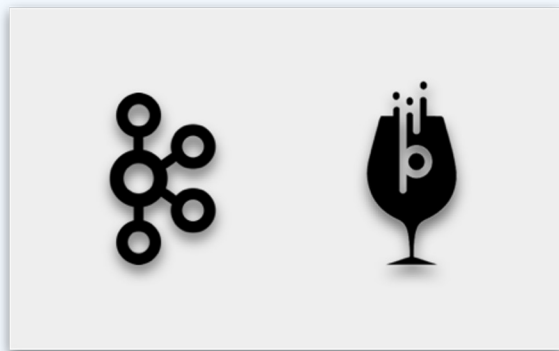
Apache Pinot Architecture



<insert any user-facing real-time analytics use case here>



Events



Realtime
Analytics

**Simply point Pinot to
Kafka,
and query!**

Powered by Apache Pinot

Community

100+

Companies

2400+

Slack Users

3.9k

Github Stars

Performance

1M+

Events/sec

200k+

Peak QPS

ms

Query Latency



pinot.apache.org



Takeaways

- Real-time analytics lets us create applications that give users actionable insights
- Properties of these systems: Fresh data, fast querying, at scale
- Kafka + Pinot is the perfect combination to achieve this

Thank you!

 davidgs@startree.ai

 @davidgs@tty0.social

 @davidgsloT

 in/davidgsimmons

 davidgs.com



dev.startree.ai



davidgsloT