

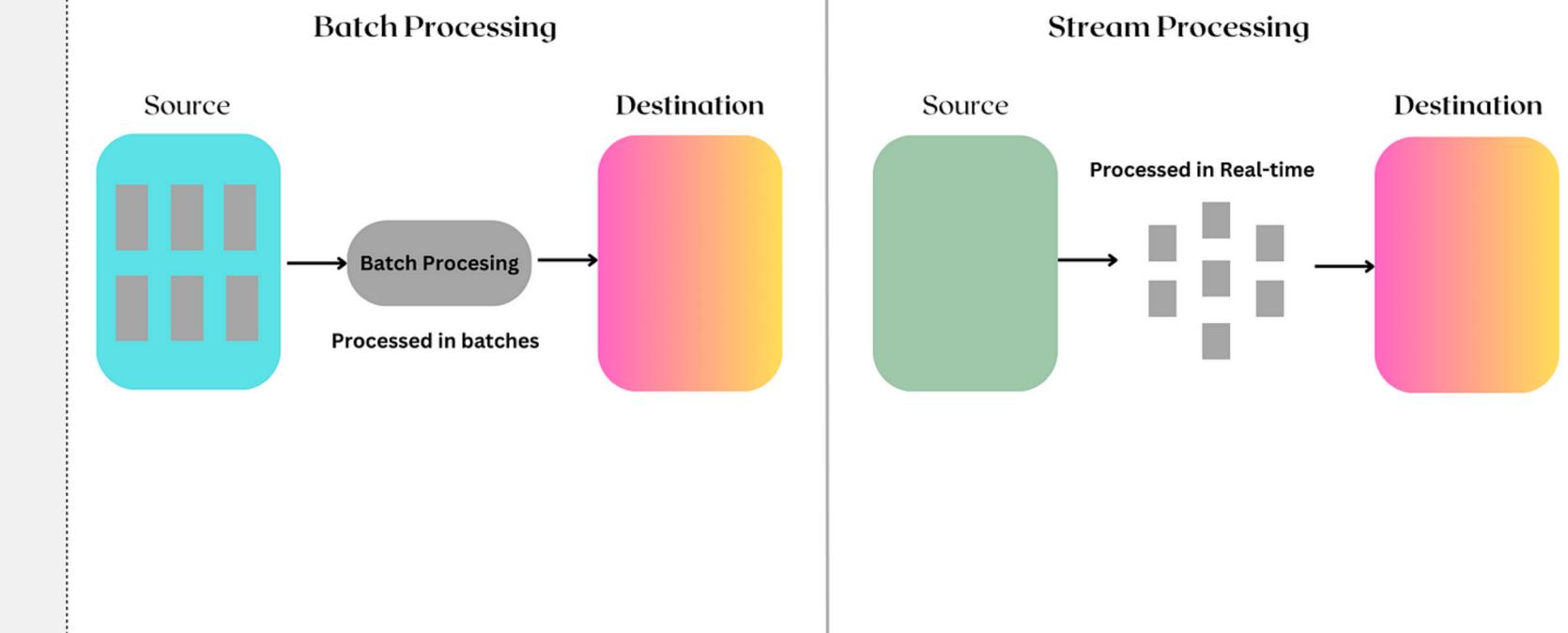
AWS Kinesis: A Stream Data Processing Powerhouse

Author: Hognogi Ana-Maria Cristina

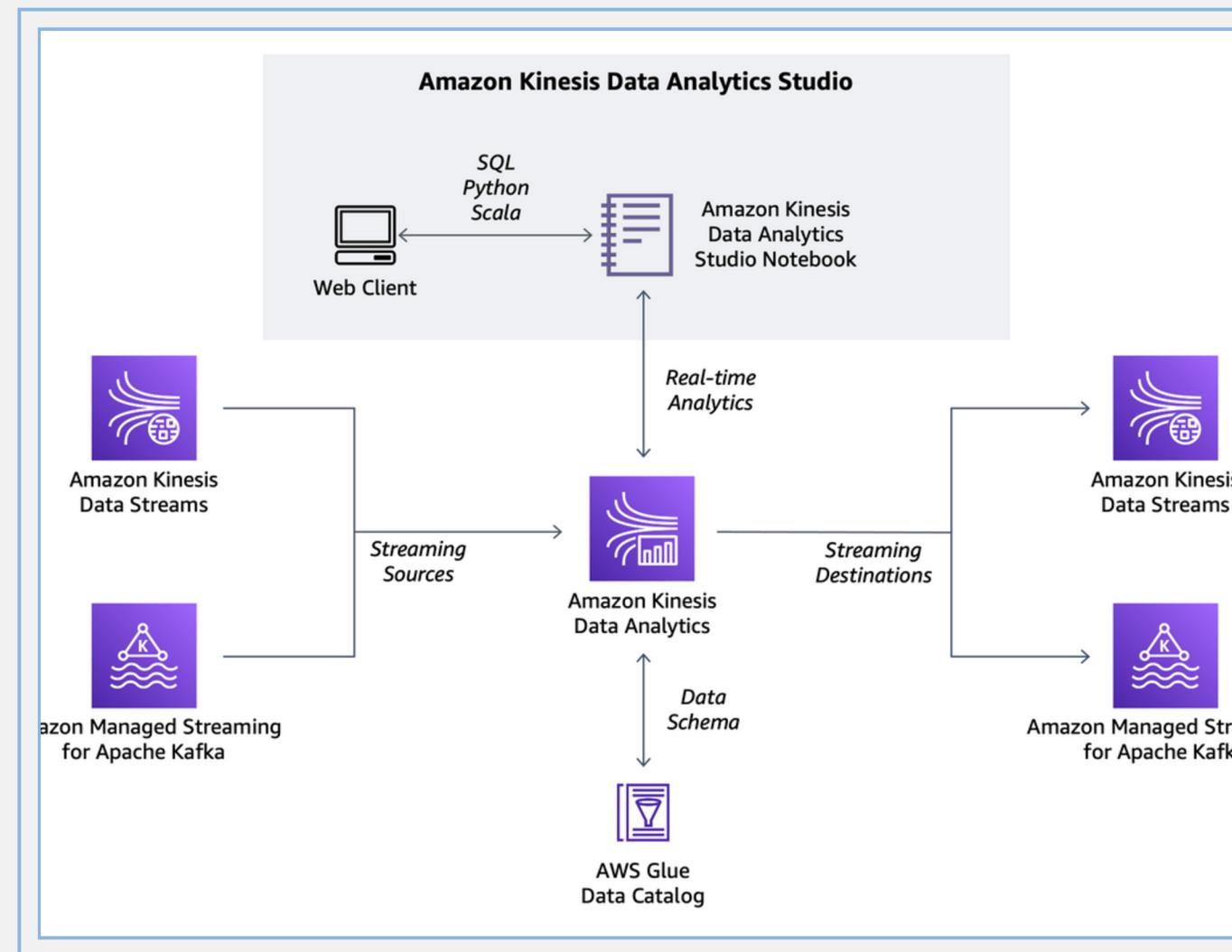
The Need for Real-Time Data Processing

With the rise of IoT, social media, and real-time applications, processing data as it arrives is crucial. AWS Kinesis enables businesses to handle these data flows in real-time without managing complex infrastructure.

Batch vs Stream Processing



What is AWS Kinesis?



Kinesis is a family of services designed to handle different aspects of real-time data workflows. Whether you're processing logs, video feeds, or applying SQL queries to streams, Kinesis has a tailored solution.

Kinesis Services

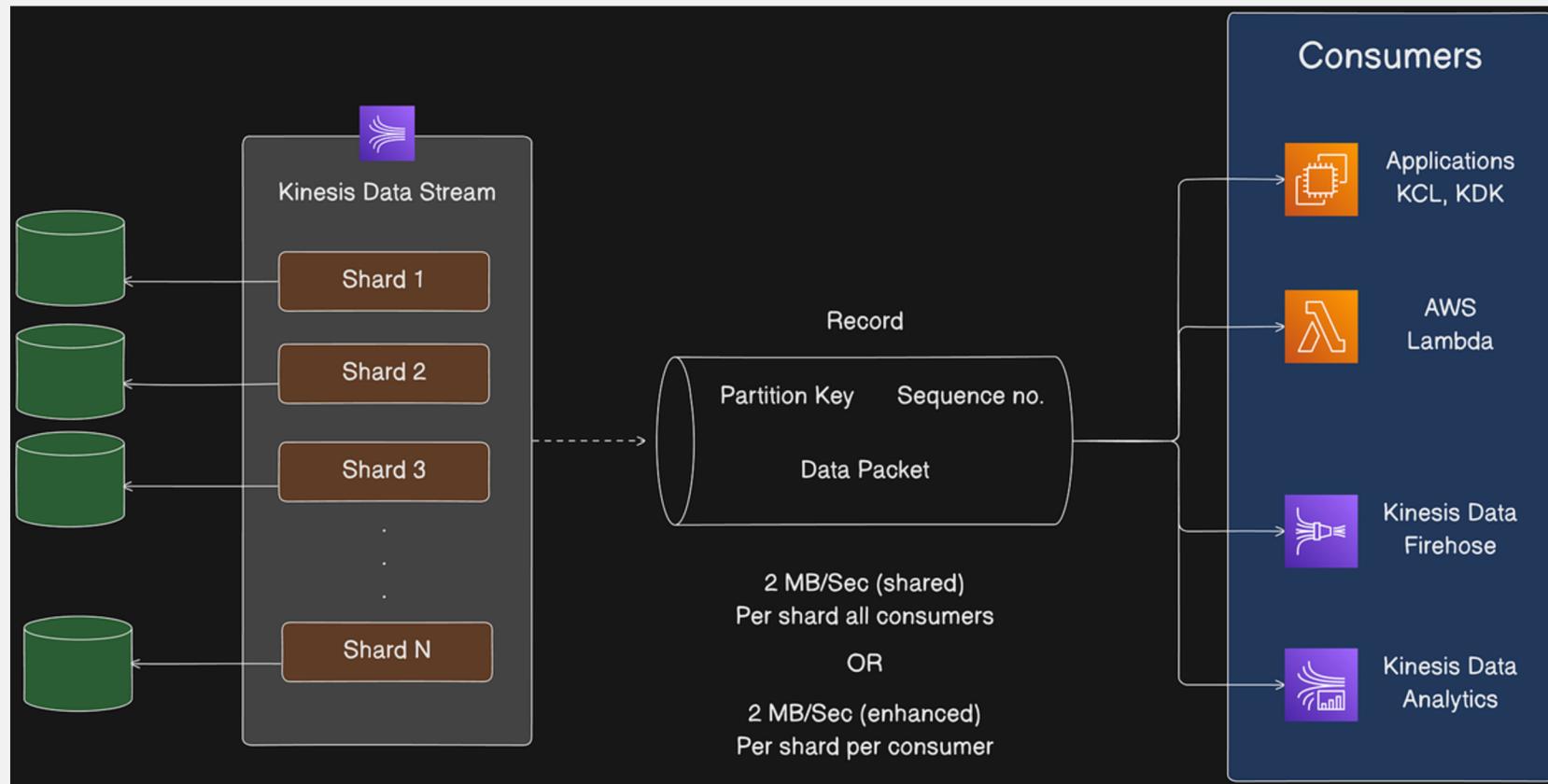
Breakdown

- 01 Kinesis Data Streams: Custom consumers
- 02 Kinesis Video Streams: Video ingestion
- 03 Kinesis Firehose: Delivery to storage
- 04 Kinesis Analytics: Real-time SQL

AWS Kinesis Data Streams vs Kinesis Data Firehose

Kinesis Data Streams		Kinesis Data Firehose
Purpose	Low latency streaming service for ingest at scale	Data transfer service to load streaming data into S3, Redshift, OpenSearch, Splunk and other third-party tools
Message propagation delay	Real time (~200 ms latency for classic shared throughput and ~70 ms for enhanced fan out)	Near real time (depends on the buffer size OR buffer time with min. 60 secs)
Provisioning	Managed service but needs configuration for shards	Fully managed service with no administration
Scaling	Manual Scaling	Automated Scaling - as per the demand
Data Storage	Configurable from 1 day to 365 days	Does not provide data storage
Replay capability	Supports replay capability	Does not support replay capability
Producers	Next to write code for producers. Supports SDK, KPL, Kinesis Agent, CloudWatch, IoT	Need to write code for producer. Supports SDK, KPL, Kinesis Agent, CloudWatch, IoT
Consumers	Open ended. Supports multiple consumers and destinations. Supports KCL and Spark.	Closed ended. Handled by Firehose and supports limited destinations. Does not support KCL or Spark.

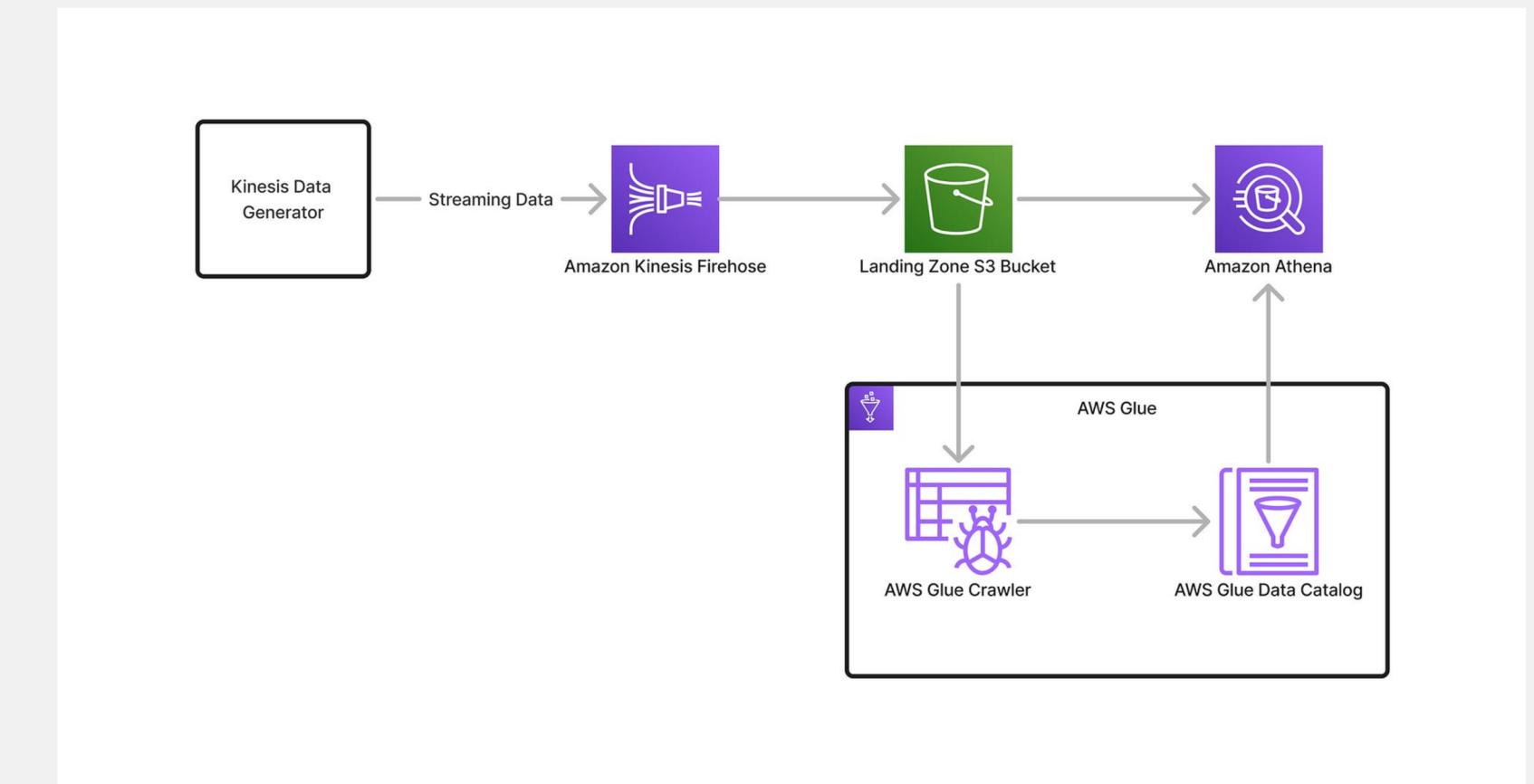
Kinesis Data Streams Architecture



The typical architecture of Kinesis Data Streams consists of data that is split into shards using partition keys, which consumers read in real-time. Checkpointing ensures data isn't lost in case of failure.

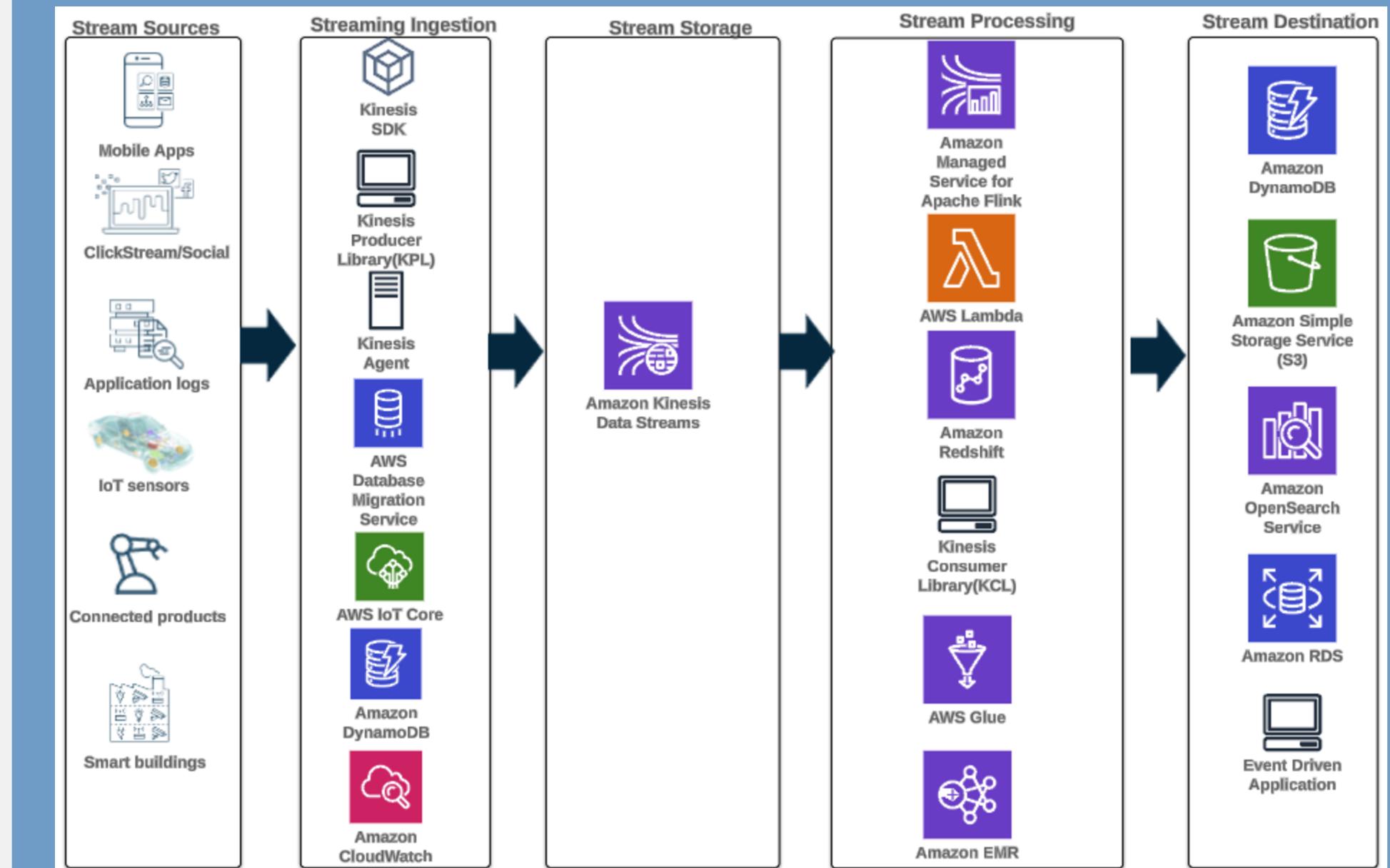
Firehose & Analytics Architecture

Kinesis Firehose abstracts much of the complexity by automatically buffering and delivering data. Analytics lets you run continuous SQL queries on live streams — perfect for alerts or dashboards.



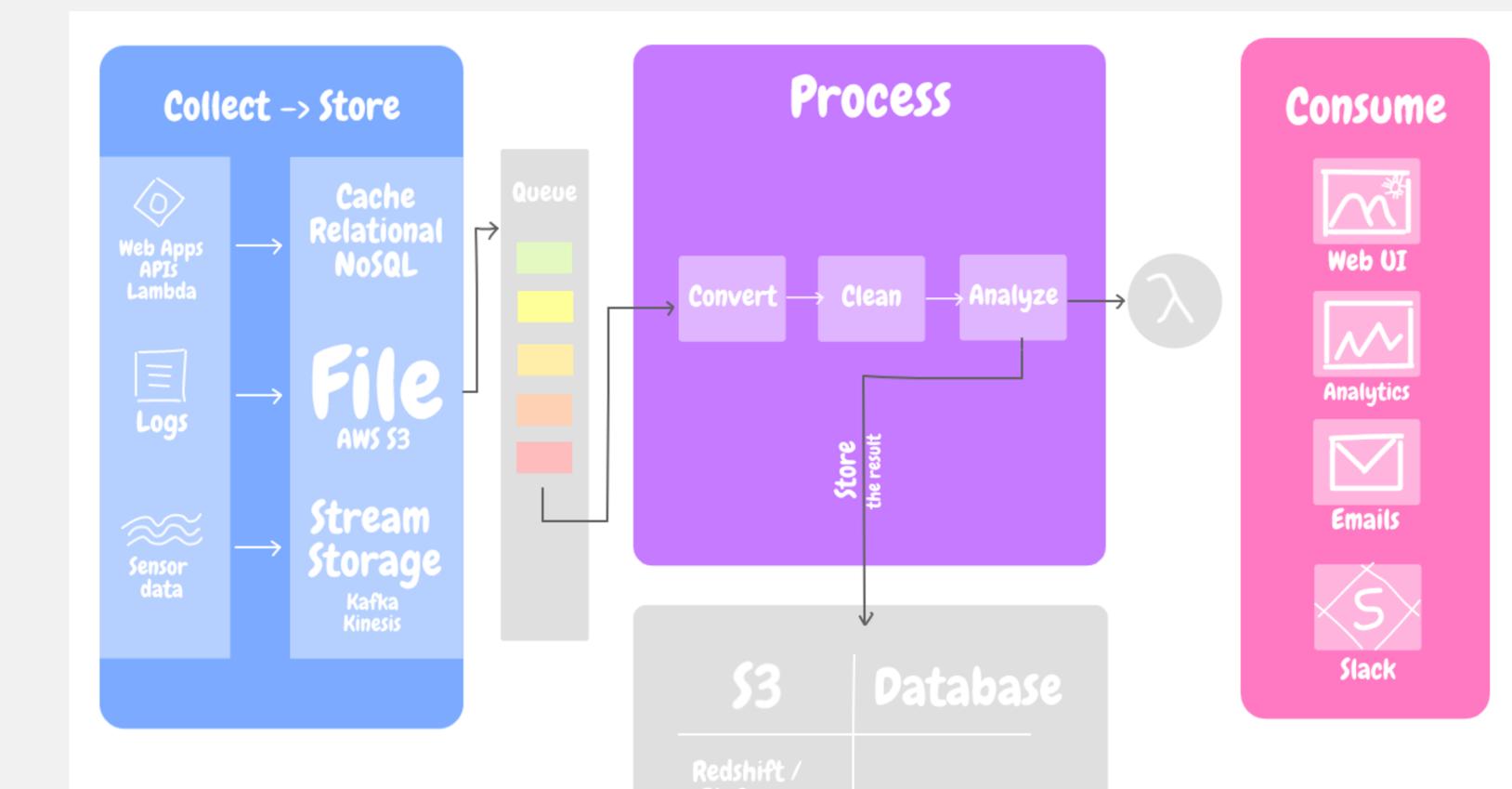
AWS Kinesis Integration

Kinesis connects deeply with the AWS ecosystem. You can ingest data from IoT or logs, process with Lambda, store in S3, and then visualize it in QuickSight – all in real time.



Kinesis in Action

Kinesis is used for log monitoring, detecting fraud in financial systems, IoT telemetry, and even video analysis. Its real-time nature supports many use cases where decisions must be immediate.



Case Studies

Major companies use Kinesis at scale. Netflix processes millions of logs to monitor uptime. Zillow personalizes property recommendations, while Comcast streams video from smart devices.



01 Netflix



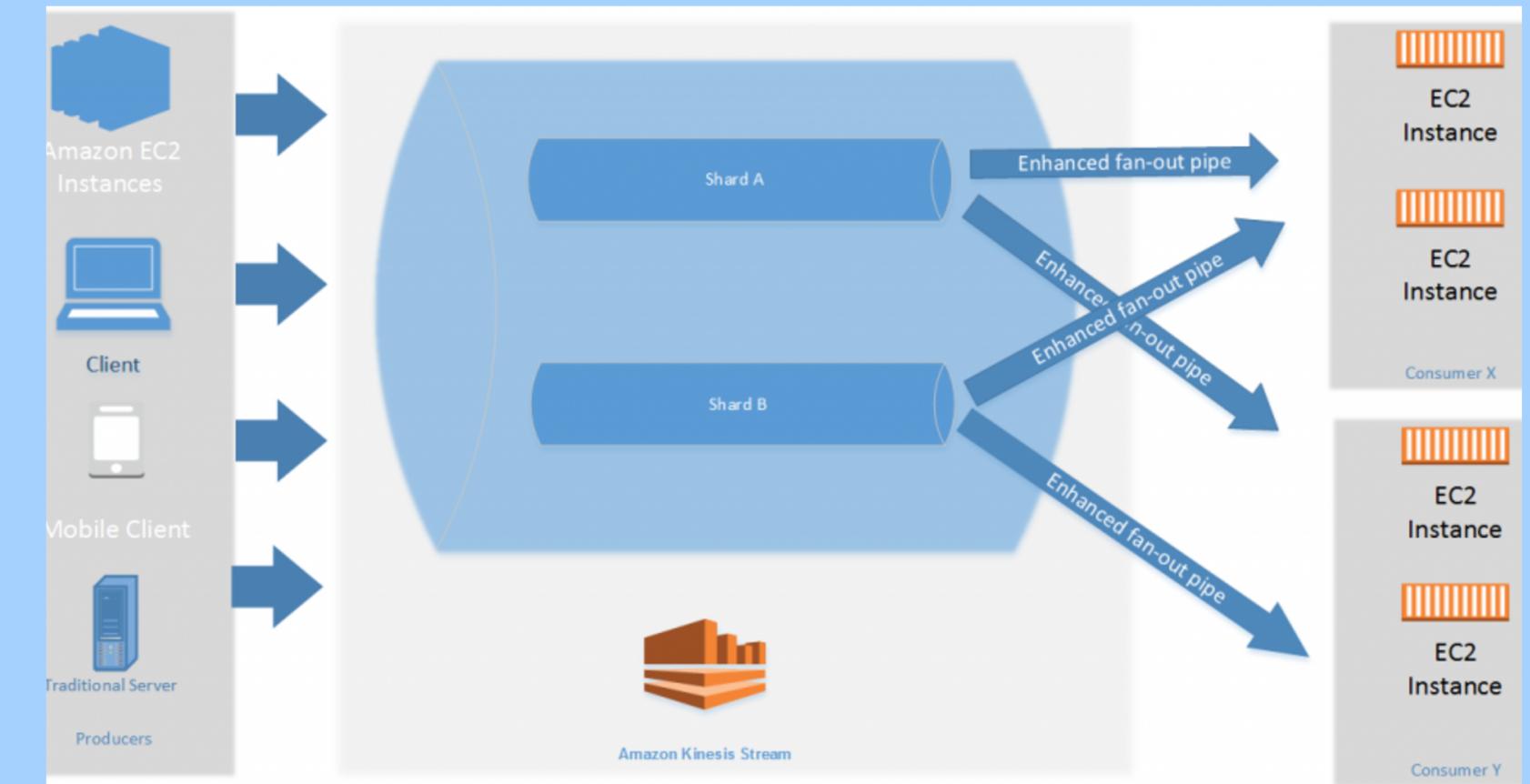
02 Zillow



03 Comcast

Performance and Scalability

Kinesis offers sub-second latency with features like enhanced fan-out. You can scale horizontally by adding shards, and it works beautifully with serverless tools like AWS Lambda.



Cost and Billing

Costs are usage-based – by shards, data volume, and processing time. It's efficient, but can get expensive at high scale, especially with enhanced features like Lambda transformations



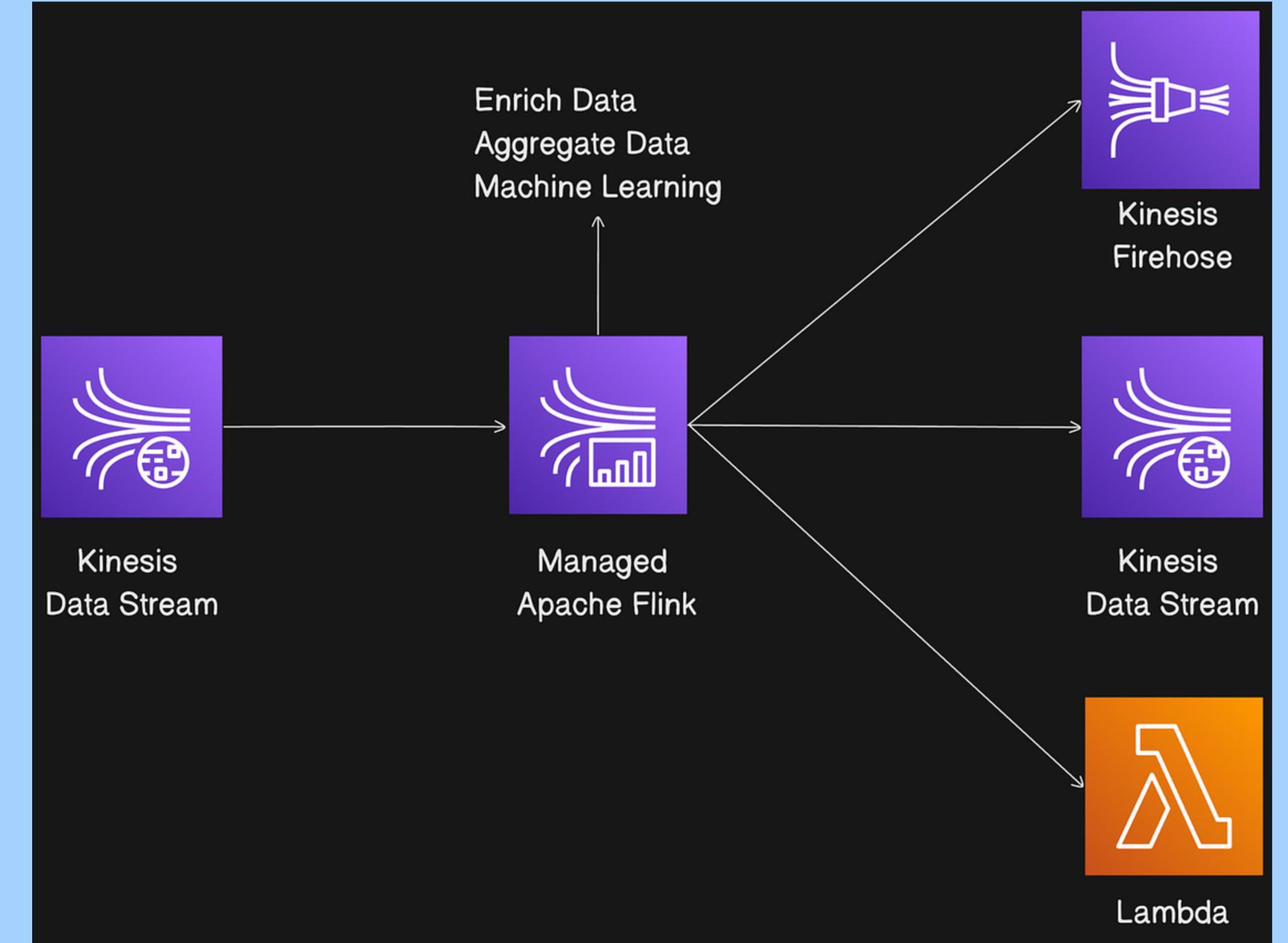
Challenges and Limitations

POINTS OF DIFFERENCE	SQS	KINESIS
WORKING	Producer and consumer architecture.	Amazon Kinesis Video Streaming, Data Stream, Data Firehouse and Data Analytics.
FEATURES	Message timer, delay queue and unlimited queues.	Real-time processing and easy to use.
ADVANTAGES	Reliable, secures sensitive data and eliminates administrative overhead.	Fast, management and scalable.

Despite its power, Kinesis has limitations. Ordering is only guaranteed within shards, and improper partitioning can cause hot spots. Firehose also introduces latency that may not work for all use cases.

Looking Ahead

As real-time data grows with edge devices and AI, Kinesis is likely to evolve with better automation and tighter integration into AWS's ML ecosystem.





amazon KINESIS

Conclusion

AWS Kinesis is a powerful, flexible, and scalable platform that plays a central role in enabling real-time intelligence for modern data-driven organizations.

Cost and Billing

Amazon Kinesis Data Streams.

Amazon Kinesis Data Analytics

Amazon Kinesis Data Firehose

Thank you!