



INNOVATE

ONLINE CONFERENCE

分会场六：大数据分析

利用 Amazon Kinesis 搭建实时数 据分析平台

史天，AWS 解决方案架构师

议程

Amazon Kinesis 基本介绍

Amazon Kinesis Analytics 功能特性

基于 Amazon Kinesis 构建实时数据分析链条

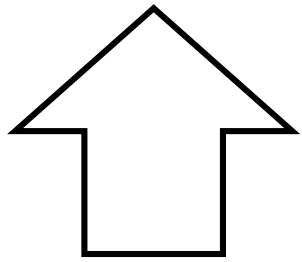
Amazon Kinesis 基本介绍

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

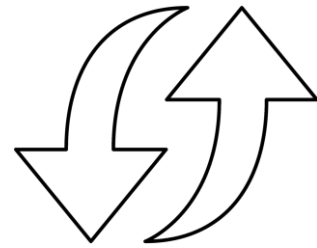


AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

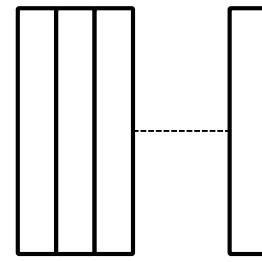
什么是流数据



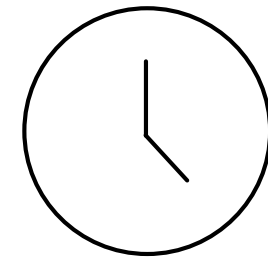
高容量



持续的

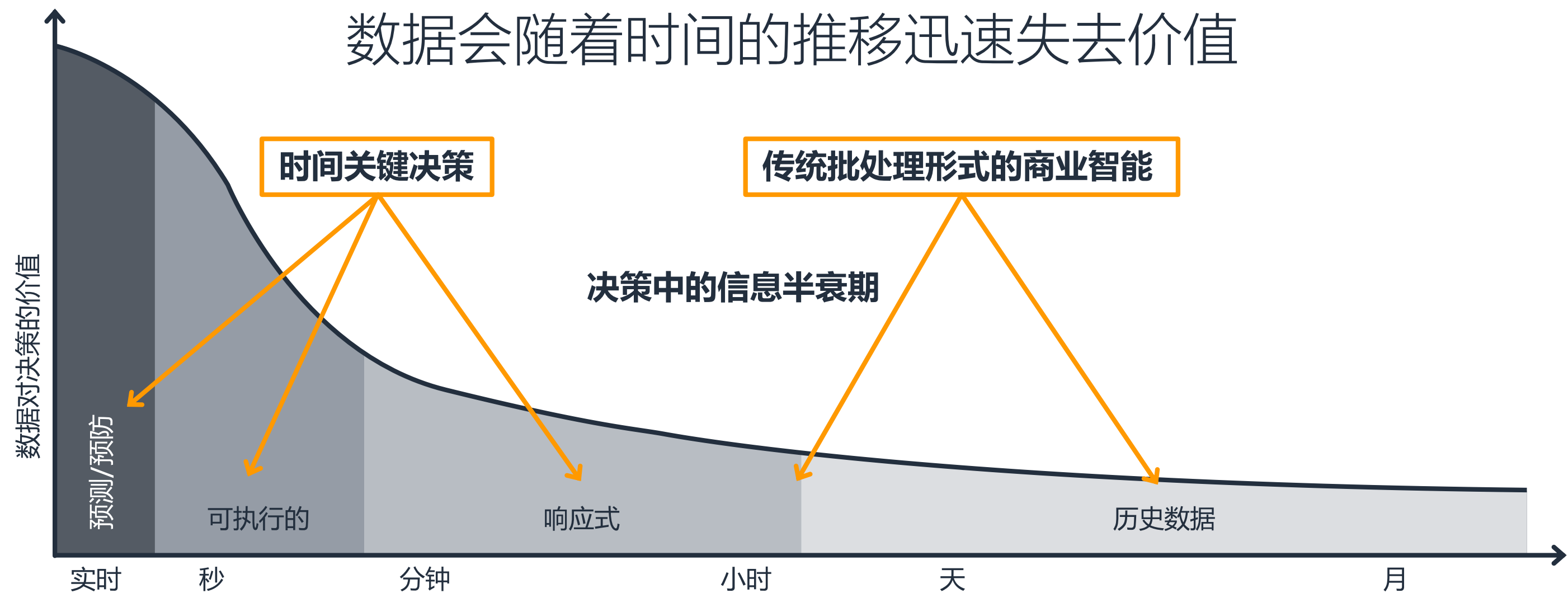


有序的增量



低延时

为什么需要数据实时分析



来源: Perishable insights, Mike Gualtieri, Forrester

使用 Amazon Kinesis 进行数据实时分析



**Amazon
Kinesis Video
Streams**

捕获、处理和存储视
频流



**Amazon
Kinesis Data
Streams**

捕获、处理和存储数
据流



**Amazon
Kinesis Data
Firehose**

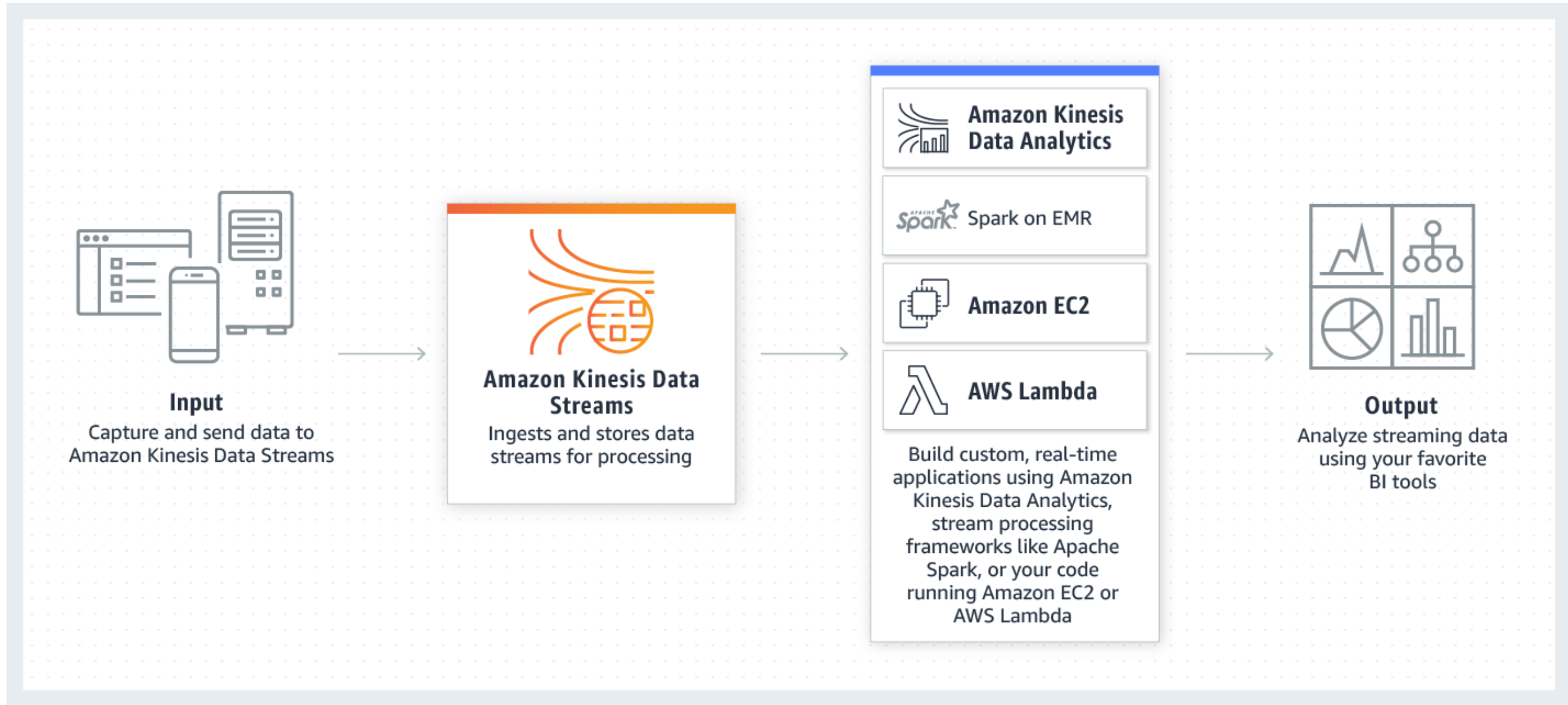
将数据流加载到数据
存储



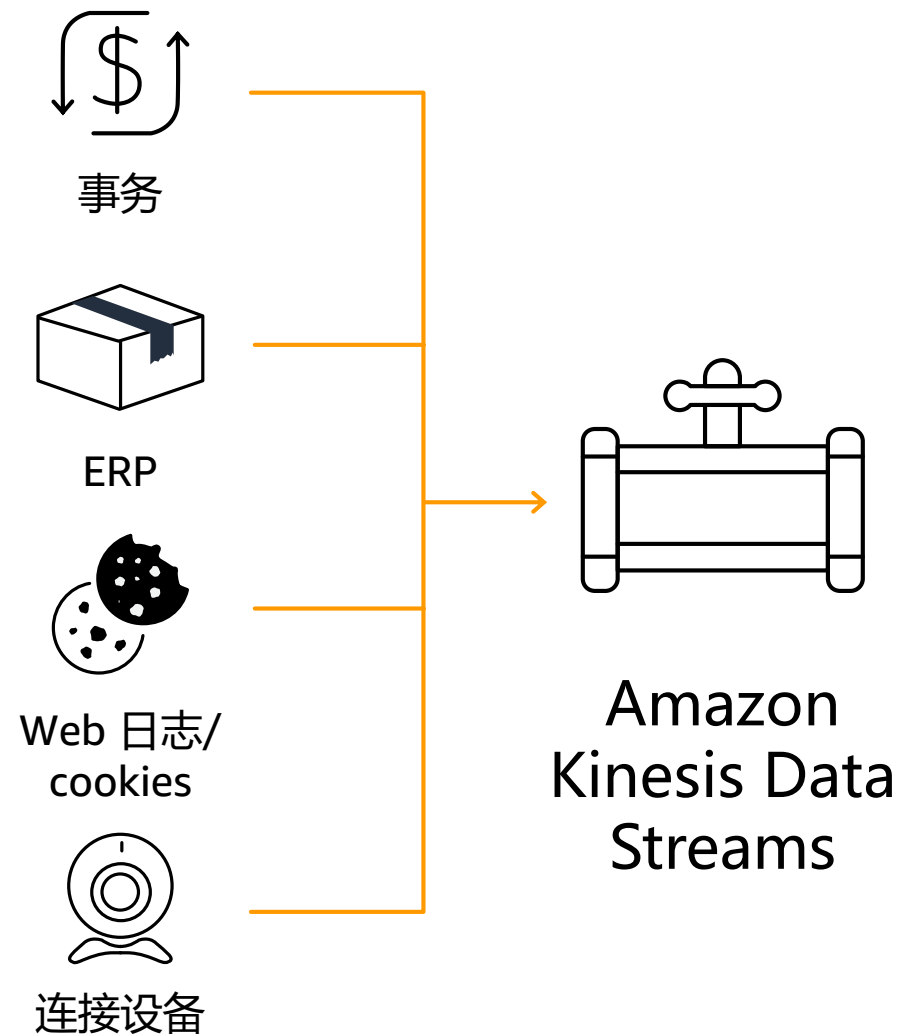
**Amazon
Kinesis Data
Analytics**

通过 SQL 和 Java 程
序分析数据流

Amazon Kinesis Data Streams 概览



从各种来源获取数据



AWS SDKs

- 在程序代码中直接通过 API 推送
- AWS Mobile SDK
- AWS 服务: Amazon CloudWatch 日志, AWS IoT, Amazon Kinesis Data Analytics 等

Kinesis Agent

- 监控日志文件并转发到 Amazon Kinesis Data Streams

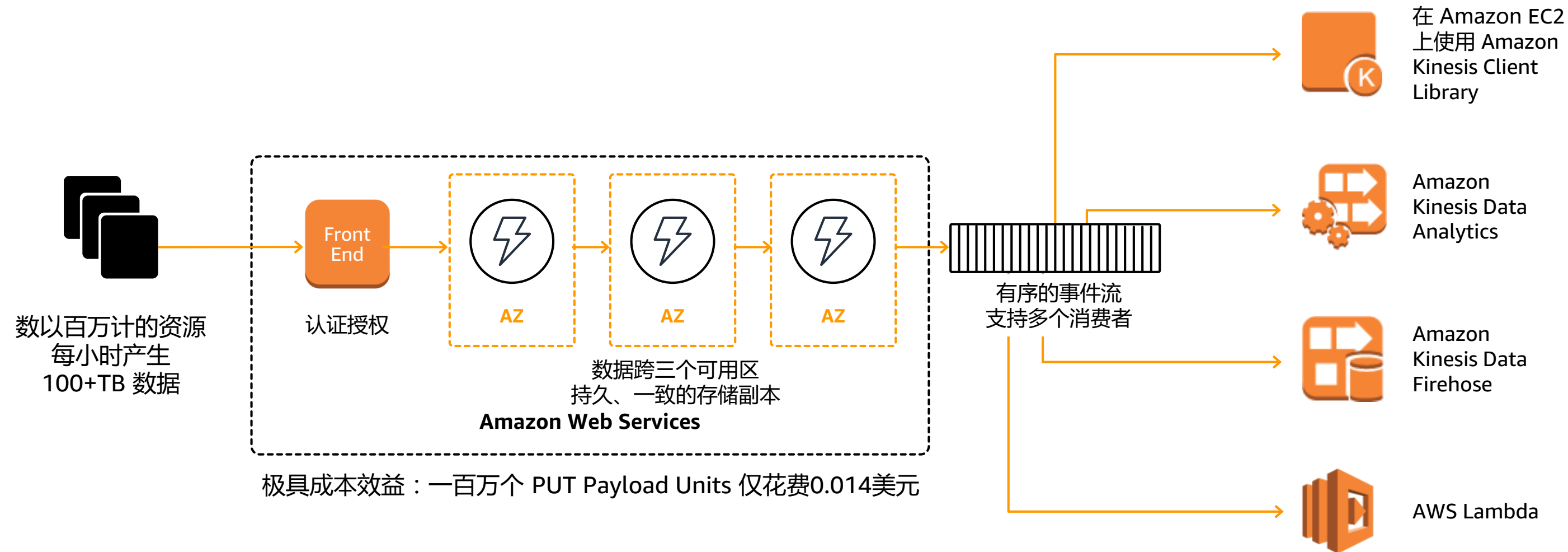
Kinesis Producer Library (KPL)

- 后台进程聚合和批处理消息

第三方和开源工具

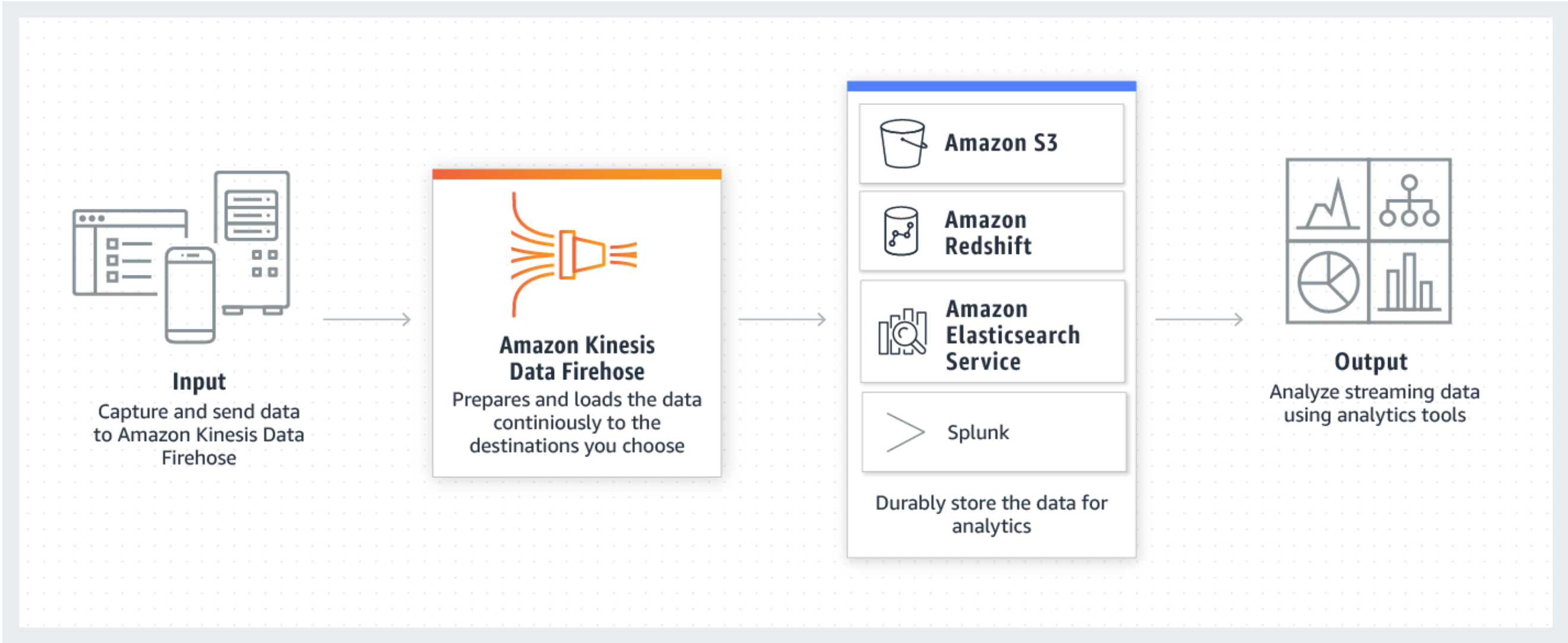
- Log4j appender
- Apache Kafka
- Flume, fluentd 等

处理来自各种消费者的数据

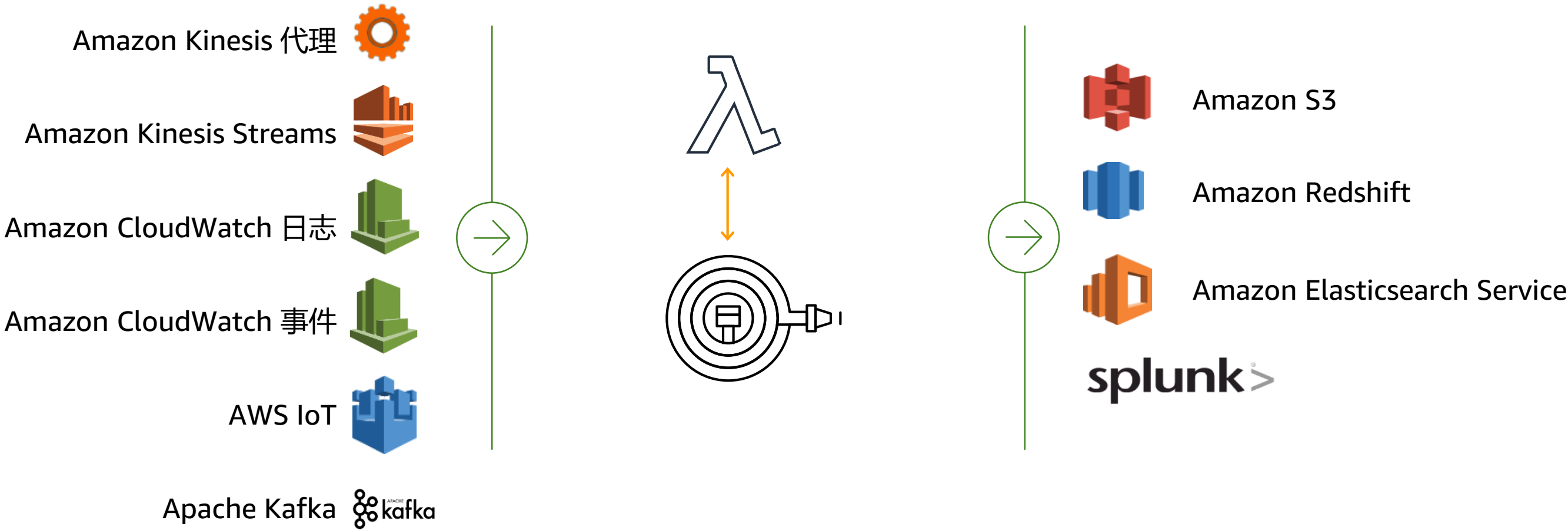


注：上述价格是按美国东部（弗吉尼亚北部）区域计算。

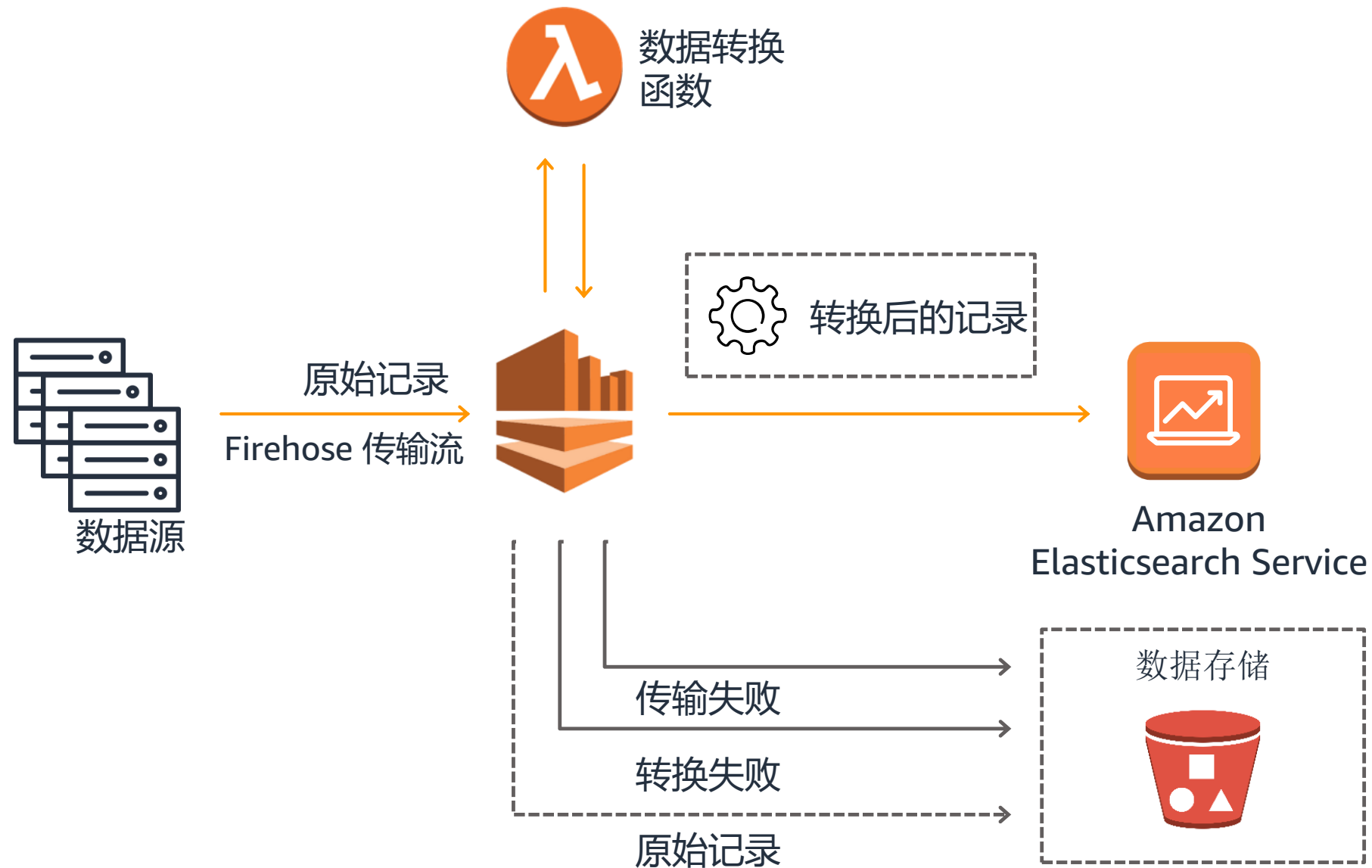
Amazon Kinesis Data Firehose 概览



Amazon Kinesis Data Firehose 基本原理

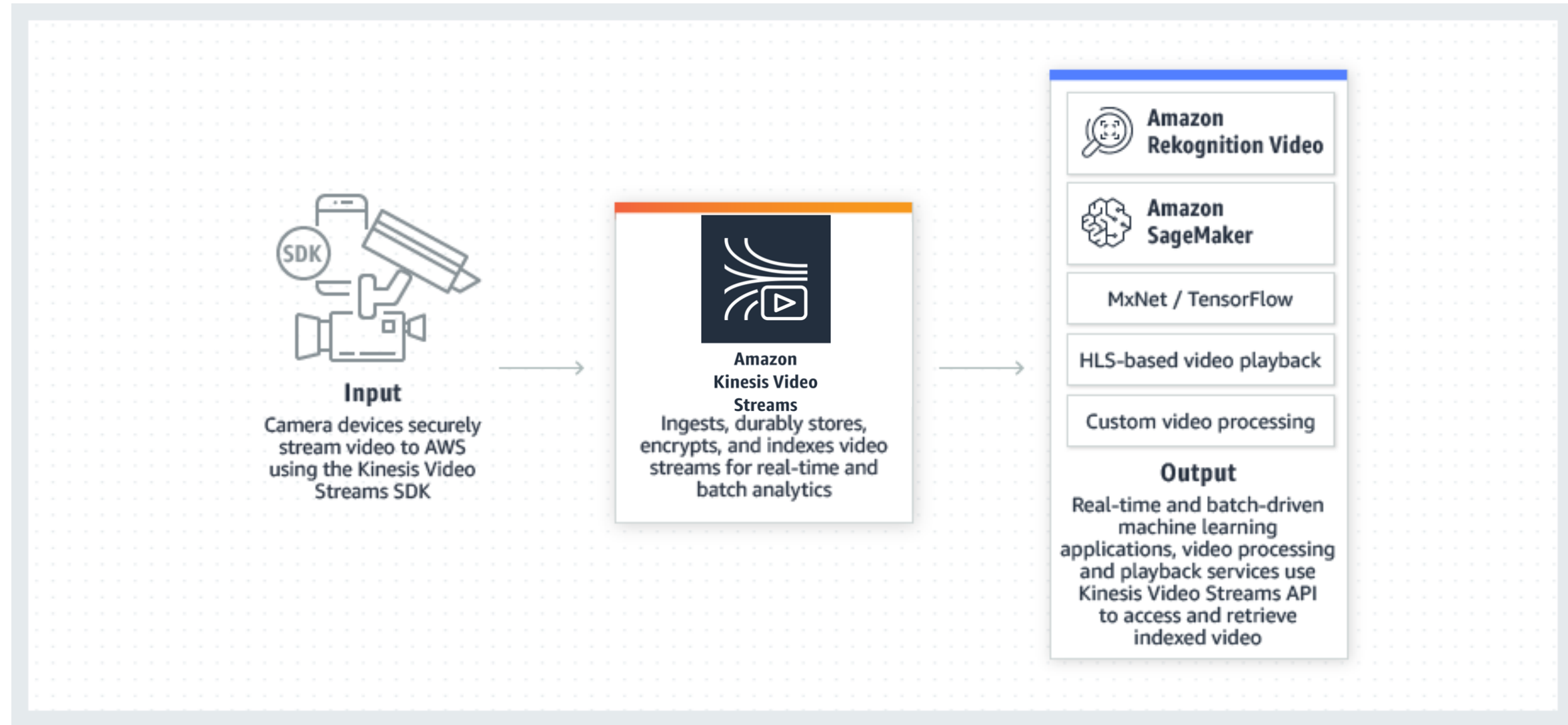


Amazon Kinesis Data Firehose 数据传输



- 针对公开访问域
- 轻松转换数据
- 内置的 Serverless 批处理, Amazon Elasticsearch 索引轮换, 错误管理

Amazon Kinesis Data Video Streams 概览



Amazon Kinesis Data Video Streams 场景

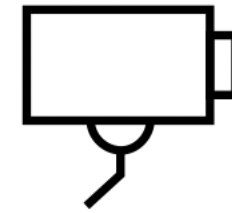
智能家居



智慧城市



安防监控



工业自动化



计算机视觉



Amazon Go 无人商店



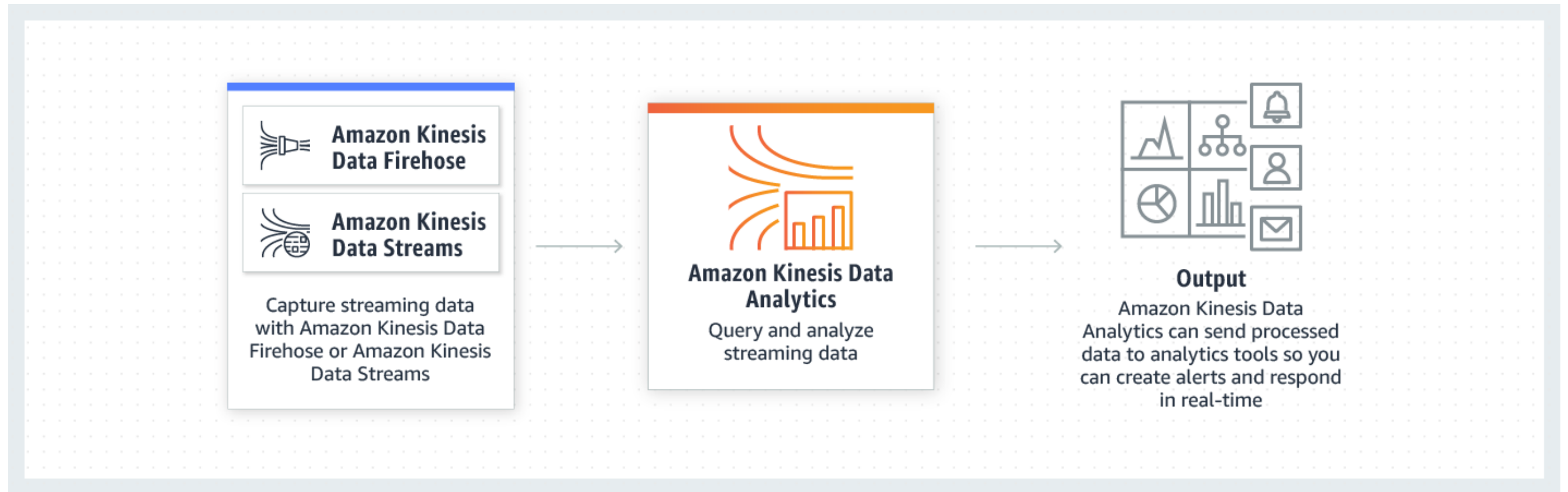
Amazon Kinesis Analytics 功能特性

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.



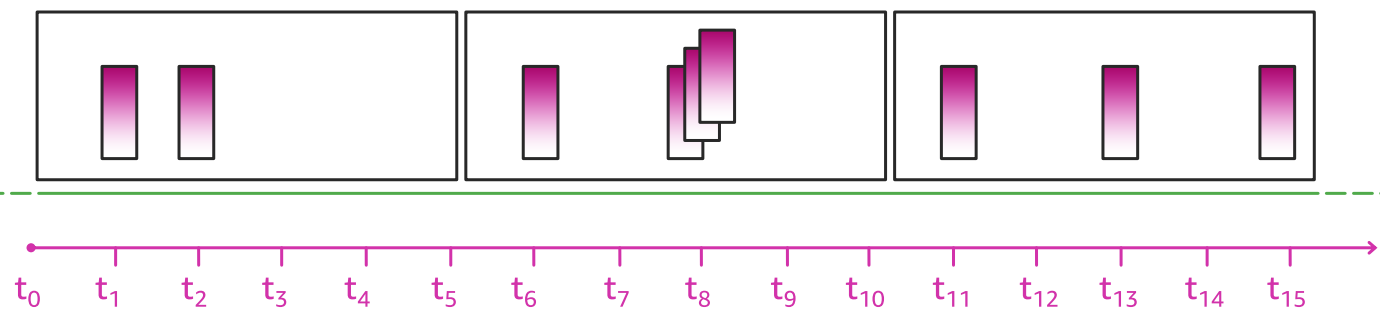
AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Amazon Kinesis Data Analytics 概览



SQL 流数据实时分析

```
CREATE OR REPLACE PUMP calls_per_ip_pump AS
INSERT INTO calls_per_ip_stream
SELECT STREAM source_ip_address,
              COUNT(*)
FROM source_sql_stream_001
WINDOWED BY STAGGER(
  PARTITION BY source_ip_address
  RANGE INTERVAL '1' MINUTE);
```



- 提供模板和交互式编辑器，以及时间窗口聚合、过滤器等丰富的 SQL 查询
- 只需根据实际需求，选择适合的分析任务的模板，然后使用 SQL 编辑器编写数据自定义的分析代码

Amazon Kinesis Data Analytics for Java

Kinesis Analytics - Create application

Kinesis Analytics applications continuously read and analyze data from a connected streaming source in real-time. To enable interactivity with your data during configuration you will be prompted to run your application. Kinesis Analytics resources are not covered under the [AWS Free Tier](#), and **usage-based charges apply**. For more information, see [Kinesis Analytics pricing](#).

Application name*


aws-sample

Description

Runtime

☐ SQL

☒ Flink 1.6



Access permissions

Create or choose IAM role with the required permissions. Kinesis Analytics will assume this role when you later execute your code. [Learn more](#)

Access permissions*

☒ Create / update IAM role **kinesis-analytics-aws-sample-us-east-1**

☐ Choose from IAM roles that Kinesis Analytics can assume

* Required

Cancel

Create application

- 包含基于 Apache Flink 的开源库，能够在数小时而不是数天或者数周内构建应用程序
- 使用操作习惯的 IDE，然后连接到 AWS 并安装 Java 库
- 包括超过25个内置运算符的可扩展库，用于过滤、聚合和转换流数据，以及与多种 AWS 服务的集成

Amazon Kinesis Data Analytics for Java

1

使用熟悉的 IDE 环境构建 Java 程序

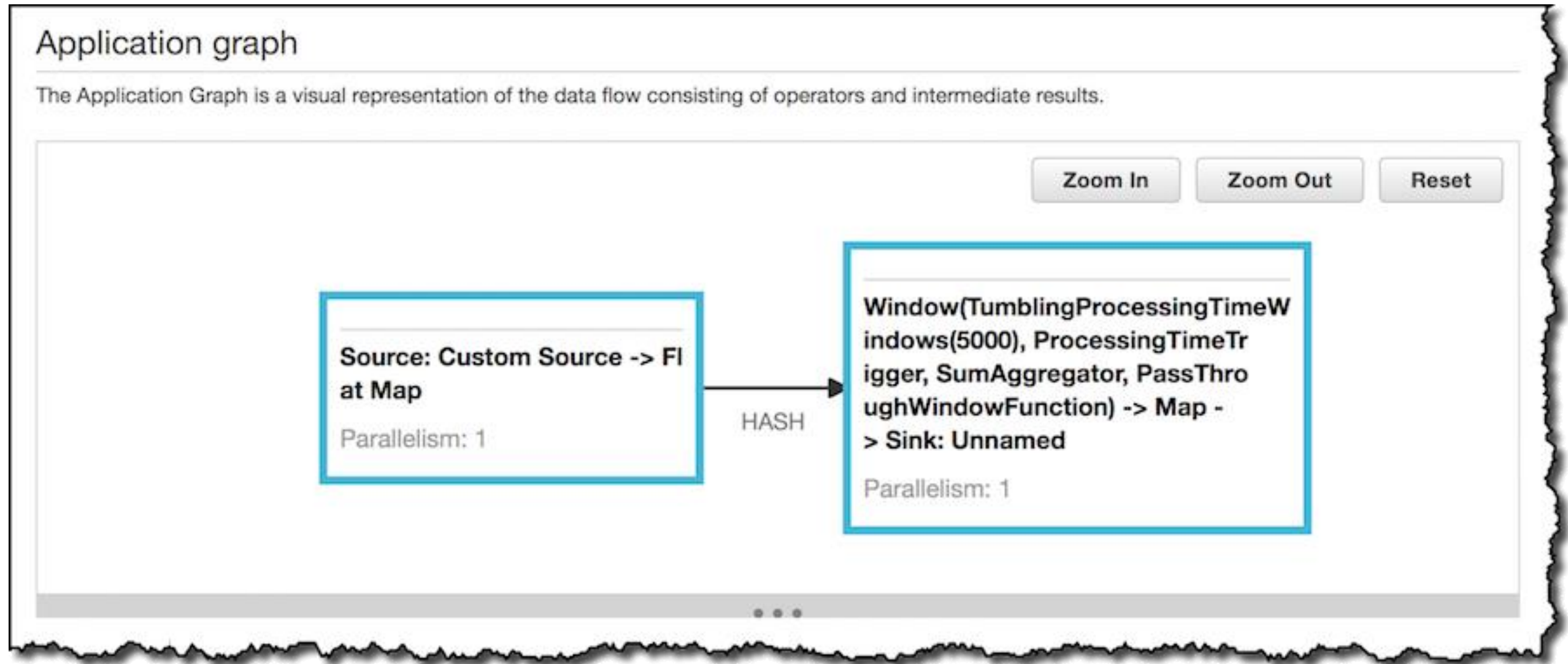
2

上传应用程序代码到
Amazon Kinesis Data
Analytics

3

在可扩展的托管服务中
运行应用程序

Amazon Kinesis Data Analytics for Java 可视化



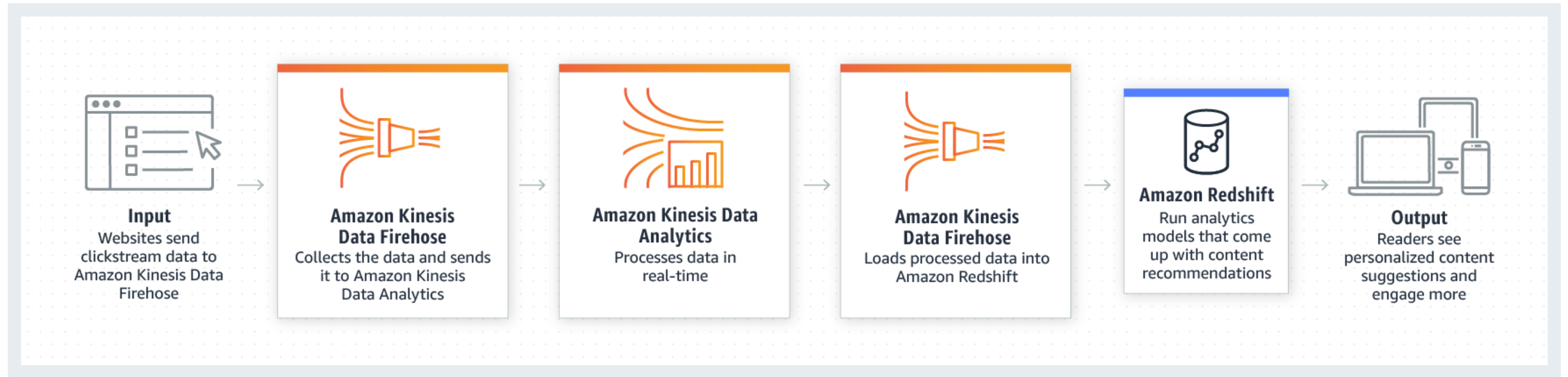
基于 Amazon Kinesis 构建实时数据分析链条

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.



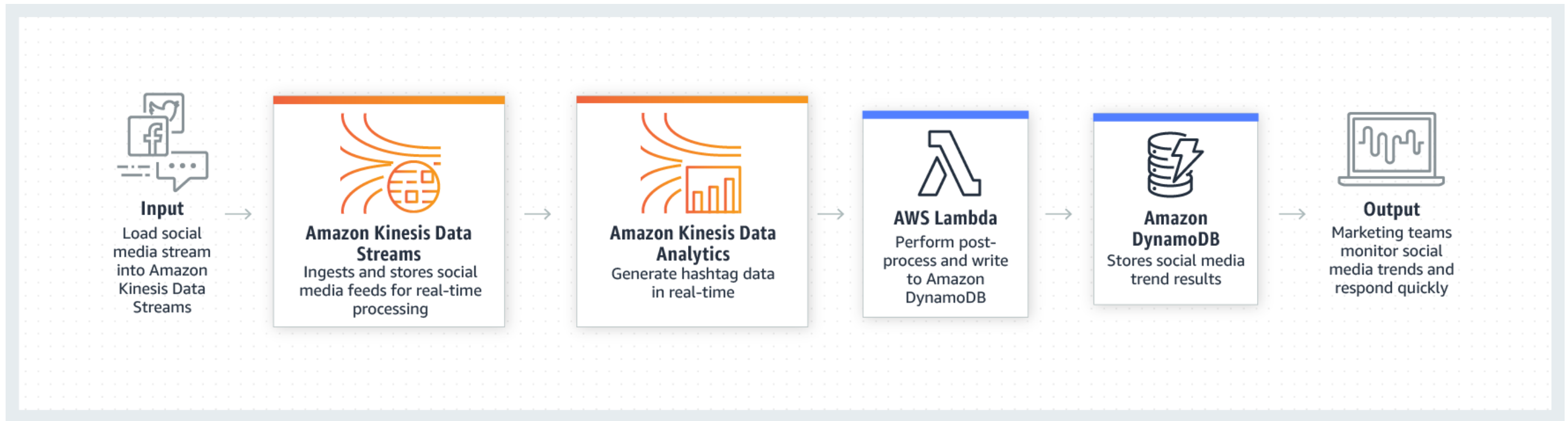
AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

点击流分析 - 从批处理到实时分析



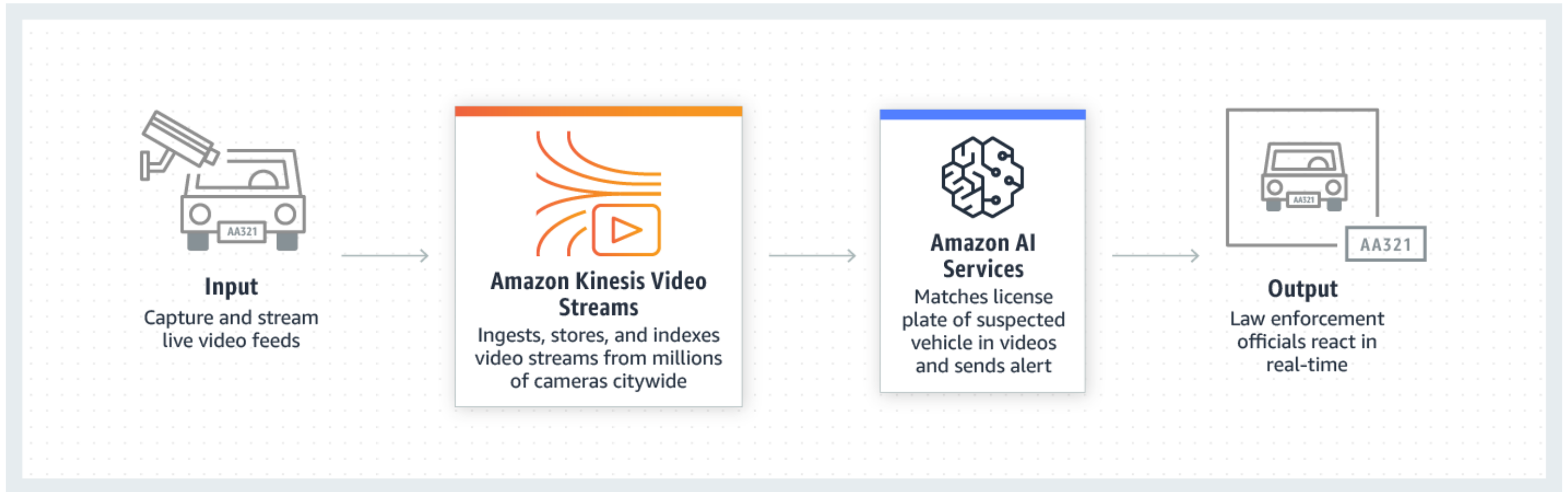
很多数据以前通过数据仓库批量处理的方式或通过 Hadoop 框架进行分析，而借助 Amazon Kinesis，您可以对此类数据进行实时分析。最常见的使用案例包括数据湖、数据科学和机器学习。您可以使用 Amazon Kinesis Data Firehose 将数据持续加载到 Amazon S3 数据湖中。您也可以在获得新数据时更频繁地更新机器学习模型，确保结果的准确性和可靠性。

社交媒体流数据分析 - 构建实时应用程序



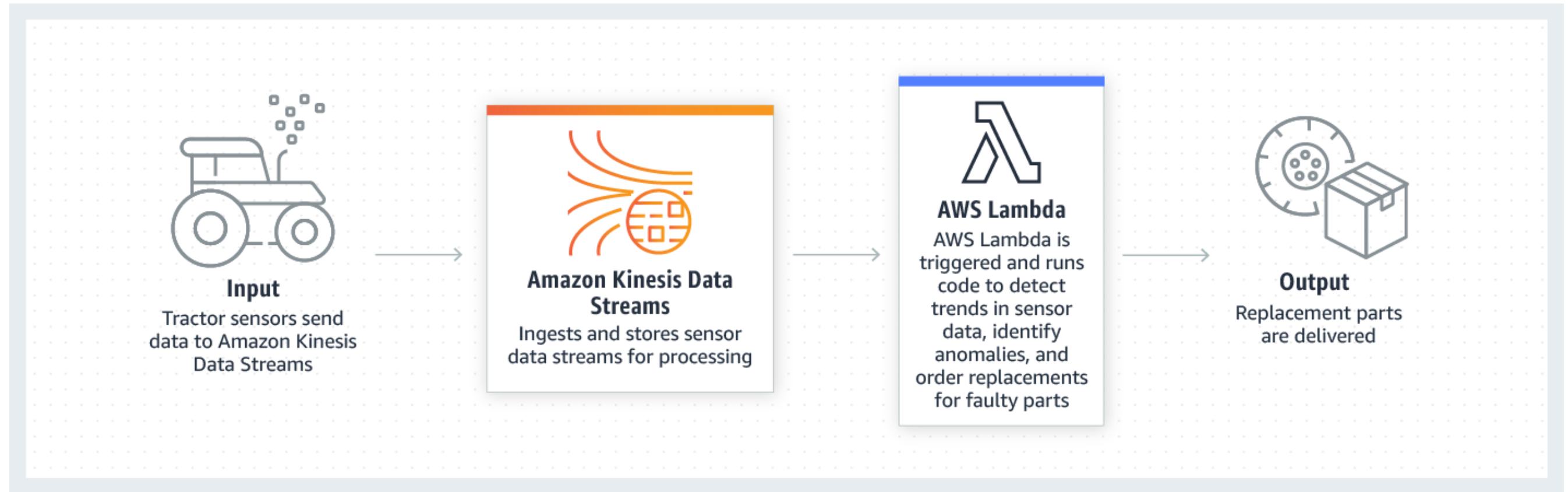
您可以将 Amazon Kinesis 用于实时应用程序，例如监控、欺诈识别和直播排行榜应用程序。您可以使用 Amazon Kinesis Data Streams 接收流数据，使用 Amazon Kinesis Data Analytics 处理流数据，然后使用 Amazon Kinesis Data Streams 将结果发送到任何数据存储或应用程序，端到端延迟只有几毫秒。这可以帮助您了解自己的客户、应用程序和产品的当前状态并迅速做出反应。

视频监控警报系统 - 构建视频分析应用程序



您可以使用 Amazon Kinesis Video Streams 将视频从家庭、办公室、工厂和公共场所的配备摄像头的设备安全地传输到 AWS。然后，可以将这些视频流用于视频播放、安全监控、面部检测、机器学习和其他分析。

传感器检测 - 分析 IoT 设备数据



您可以使用 Amazon Kinesis 来处理来自消费电器、嵌入式传感器和电视机顶盒等 IoT 设备的流数据。然后您可以使用编程方式使用这些数据，以便在传感器超过特定运行阈值时发送实时提醒或进行其他操作。使用 AWS 提供的 IoT 分析示例代码还可以帮助构建您的应用程序，无需从头开始。

Amazon Kinesis 实时数据分析



**Amazon
Kinesis Video
Streams**

捕获、处理和存储视
频流



**Amazon
Kinesis Data
Streams**

捕获、处理和存储数
据流



**Amazon
Kinesis Data
Firehose**

将数据流加载到数据
存储



**Amazon
Kinesis Data
Analytics**

通过 SQL 和 Java 程
序分析数据流

<https://aws.amazon.com/kinesis/>

感谢参加 AWS INNOVATE 2019 在线技术大会

我们希望您在这里找到感兴趣的内容！

也请帮助我们完成**投票打分**和**反馈问卷**。

欲获取关于 AWS 的更多信息和技术内容，可以通过以下方式找到我们：



微信公众号：AWSChina



新浪微博：<https://www.weibo.com/amazonaws/>



领英：<https://www.linkedin.com/company/aws-china/>



知乎：<https://www.zhihu.com/org/aws-54/activities/>



视频中心：<http://aws.amazon.bokecc.com/>



更多线上活动：<https://aws.amazon.com/cn/about-aws/events/webinar/>