Cách thức triển khai Apache Kafka cho khai thác dữ liệu từ Event Streaming:

**Bước 1: Thiết lập môi trường cho Kafka**

Cần tải bản mới nhất của Kafka và Java.

**Bước 2: Tạo Kafka Topic để lưu Kafka Events**

Kafka là một nền tảng phân tán cho việc truyền tin sự kiện (Event Streaming), cho phép bạn quản lý, đọc, ghi và xử lý các sự kiện cũng như thông điệp hoặc bản ghi (Messages hoặc Records). Tuy nhiên, trước khi bạn có thể ghi bất kỳ sự kiện Kafka nào, bạn cần tạo một Topic Kafka để lưu trữ sự kiện đó.

**Bước 3: Tiến hành nạp Kafka Events vào Topic**

Các Kafka Client sử dụng mạng để kết nối vào Kafka Brokers để nạp events. Sau khi nhận được các events, các Brokers sẽ lưu chúng theo một cách đáng tin cậy và chịu lỗi trong suốt thời gian yêu cầu.

**Bước 4: Tiến hành đọc Kafka Events**

Vì Kafka Events được lưu trữ một cách đáng tin cậy trong Kafka, nên chúng có thể được đọc nhiều lần bởi nhiều consumers.

**Bước 5: Xử lý Kafka Events bằng Kafka Streams**

Sau khi dữ liệu đã được lưu trữ trong Kafka dưới dạng Events, chúng ta có thể sử dụng thư viện Kafka Streams cho Java/Scala để xử lý nó. Thư viện này cho phép xây dựng các ứng dụng và dịch vụ thời gian thực cho các yêu cầu. Kafka Streams kết hợp việc tạo và triển khai các ứng dụng client của Java và Scala truyền thống cùng với lợi ích của công nghệ cụm máy chủ Kafka để tạo ra các hệ thống có khả năng mở rộng, chịu lỗi và phân tán.