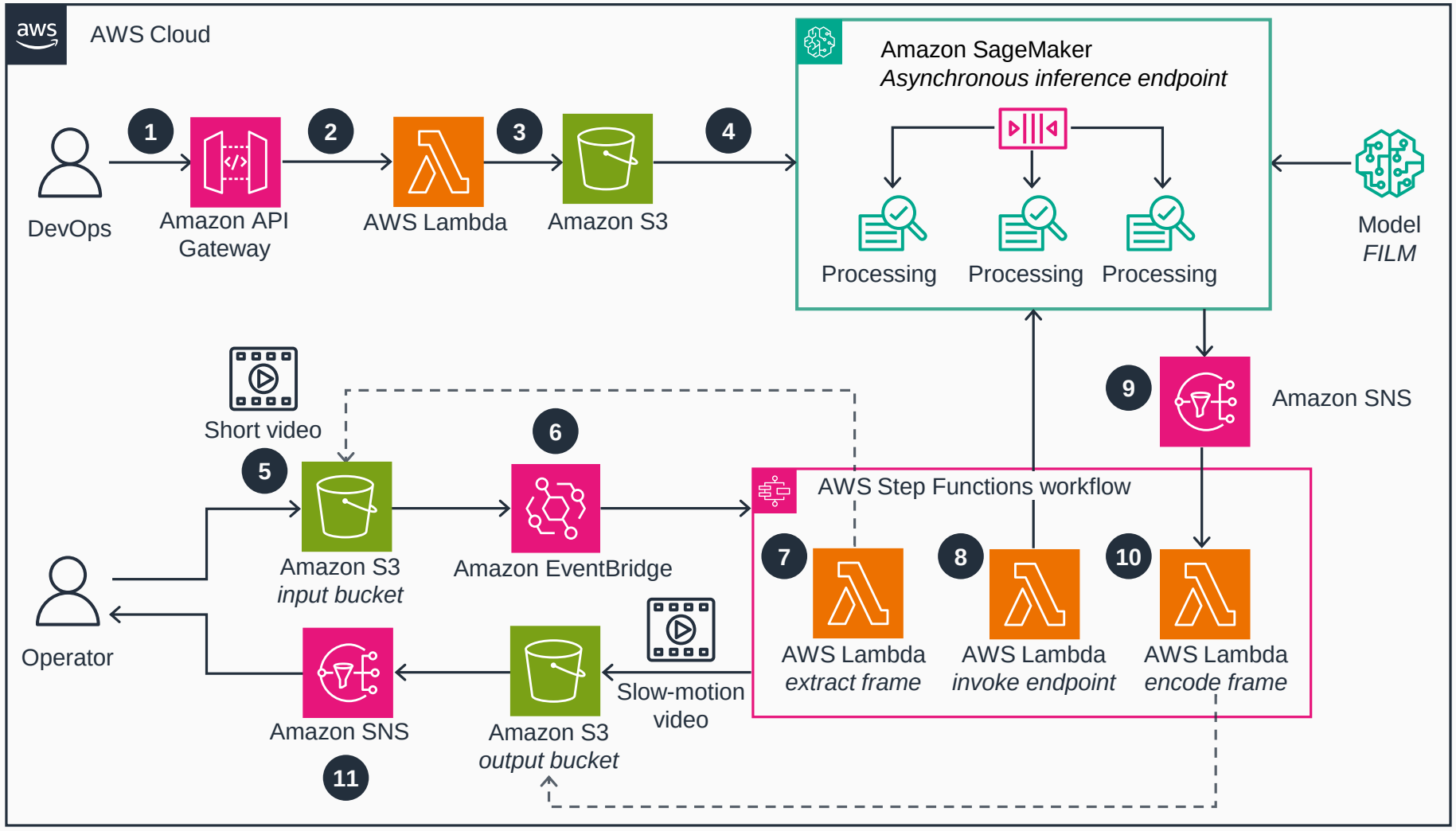
Guidance for Creating Super Slow-Motion Videos Using Generative AI on AWS

This architecture diagram shows how to create a super slow-motion video by interpolating synthetic frames between real frames using generative artificial intelligence (AI) at scale.



1. A DevOps engineer calls an **Amazon API Gateway** endpoint to create a model endpoint.
2. **API Gateway** invokes an **AWS Lambda** function to process the request.
3. The **Lambda** function uploads model artifacts. In this case, the Frame Interpolation for Large Motion (FILM) model and the endpoint configuration are uploaded to an **Amazon Simple Storage Service (Amazon S3)** bucket, creating an endpoint.
4. On endpoint creation, **Amazon SageMaker** creates an asynchronous inference endpoint with an autoscaling feature.
5. The operator uploads a short video to an **Amazon S3** bucket for processing.
6. An **Amazon S3** event invokes an **AWS Step Functions** state machine through **Amazon EventBridge** to process the request.
7. A **Lambda** function extracts frames from the video and stores them in an **Amazon S3** bucket.
8. A **Lambda** function creates an inference job by invoking the **SageMaker** Asynchronous Inference endpoint, where the FILM model interpolates new frames. The state machine processing is paused and waits for a job completion status.
9. The **SageMaker** Asynchronous Inference endpoint sends the job status to **Amazon Simple Notification Service (Amazon SNS)**.
10. The state machine processing job resumes, where a **Lambda** function encodes all new frames to create a slow motion video, storing the frames in an **Amazon S3** bucket.
11. An **Amazon S3** event sends the status to **Amazon SNS** to notify the operator that the slow-motion video is complete.

Reviewed for technical accuracy April 1, 2024

AWS Reference Architecture