Frame-based Analysis for Your Videos

AWS Implementation Guide

Paulo Almeida

Bruno Emer

David Reis

Lidio Ramalho

Thiago Morais

November 2017

Notice: Frame-based Analysis for Your Videos has been superseded by the Media Analysis Solution.

In 2017, Frame-based Analysis for Your Videos was launched to help AWS customers more easily search, query, and analyze their video libraries. In 2018, AWS launched the [Media Analysis Solution](#) a new and improved solution that uses serverless, AWS-native artificial intelligence services to automatically extract valuable metadata from media files. We encourage customers to migrate to the Media Analysis Solution for future updates and new features.

Legacy templates, scripts, and documentation for Frame-based Analysis for Your Videos are available in our [GitHub repository](#).

Contents

About This Guide

This implementation guide discusses architectural considerations and configuration steps for deploying the Frame-based Analysis for Your Videos solution on the Amazon Web Services

(AWS) Cloud. It includes links to an AWS CloudFormation template that launches, configures, and runs the AWS services required to deploy this solution using AWS best practices for security and availability.

The guide is intended for IT infrastructure architects and developers who have practical experience architecting on the AWS Cloud.

Overview

Amazon Web Services (AWS) offers powerful and cost-effective services to help customers process, analyze, and extract meaningful data from their video files. Customers who want to obtain a broader understanding of their video libraries can use AWS services to develop solutions that quickly and seamlessly analyze frames in their video files. However, developing these solutions can sometimes require extensive knowledge of deep-learning algorithms, depending on the level of analysis required.

To help customers more easily search, query, and analyze their video libraries, AWS offers the Frame-based Analysis for Your Videos solution. This solution combines Amazon Rekognition with the open source software [FFmpeg](), a fast video converter, to identify and extract frame-level metadata from video files. This AWS solution automatically detects and extracts video frames using machine learning algorithms to identify facial, object, and scene-level metadata in those frames, and stores the resulting metadata for future search and analysis. It also includes a *facial search* feature that uses customer-provided image files to search for matches in the collected video metadata.

Cost

You are responsible for the cost of the AWS services used while running this reference deployment. As of the date of publication, the estimated base Amazon Elastic Cloud Compute (Amazon EC2) and Amazon DynamoDB capacity costs for running this solution with default settings in the US East (N. Virginia) Region is **$0.22 per hour**. This cost estimate can scale up to **$1.12 per hour** based on the number of simultaneous videos being processed.

Video frame processing costs will vary based on the number of frames analyzed. By default, this solution processes a frame every second, which will result in 3,600 frame images per hour of video. Each frame is processed by the Amazon Rekognition `IndexFaces` and `DetectLabels` API calls, resulting in two Amazon Rekognition API calls per frame image.

Please see the [Amazon Rekognition pricing page]() for more information, including details about free tier and volume discount pricing. Note that this estimate does not include charges incurred from Amazon Simple Storage Service (Amazon S3) and Amazon DynamoDB

storage, AWS IoT, Amazon Simple Queue Service (Amazon SQS), or data transfer costs. Prices are subject to change. For full details, see the pricing webpage for each AWS service you will be using in this solution.

Architecture Overview

This solution contains video-processing and face-search components. The *video-processing* component runs uploaded videos through a frame-based analysis workflow to extract image metadata associated with video frames. The *face-search* component uses uploaded images to search video metadata and identify videos that contain matching faces. The following sections describe these components in more detail.

Video Processing Component

Deploying this solution with the **default parameters** builds the following video-processing component in the AWS Cloud.
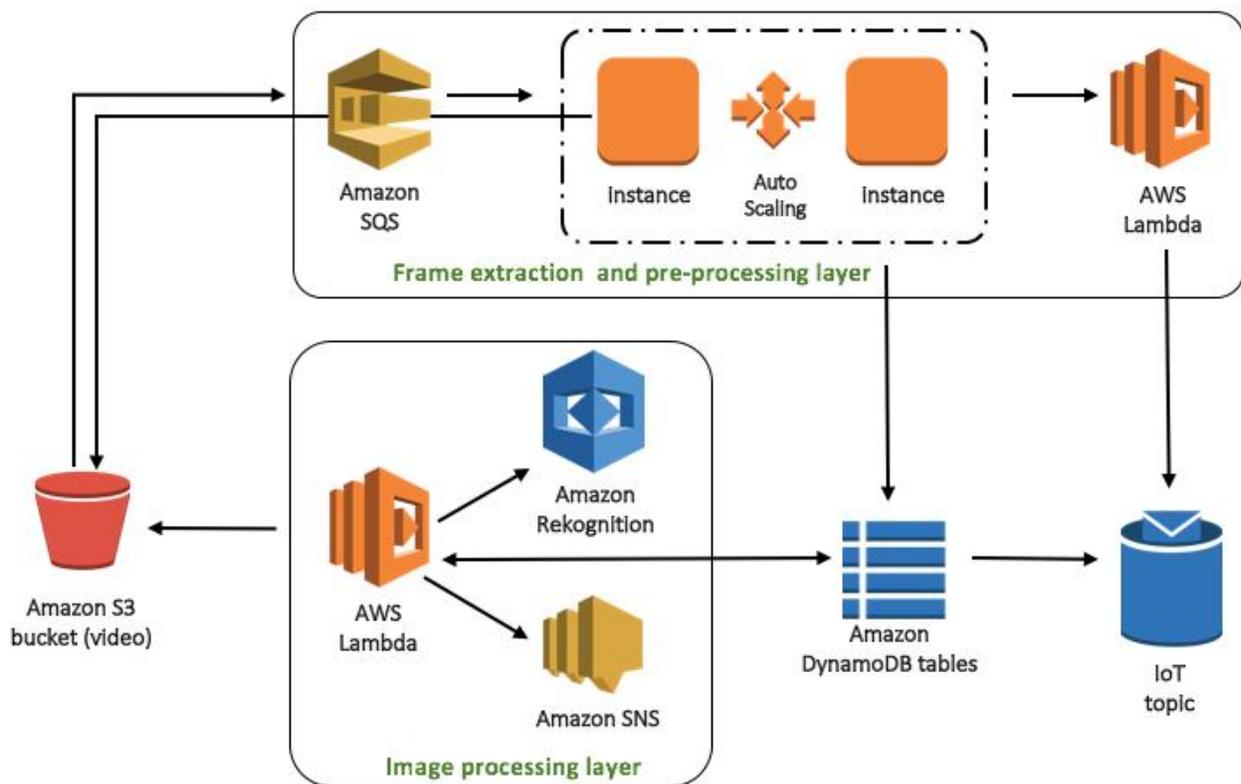

Figure 1: Video Processing Architecture

Amazon S3 Video Bucket

This solution creates an Amazon Simple Storage Service (Amazon S3) bucket for storing videos and extracted frames. When a video is uploaded to this bucket, it triggers the frame analysis workflow. Amazon S3 creates an Amazon Simple Queue Service (Amazon SQS) message, which is used to manage the queue of videos that the solution will process.

Frame Extraction and Preprocessing

To perform frame extraction and preprocessing, the solution launches Amazon Elastic Compute Cloud (Amazon EC2) instances in an Auto Scaling group. Each EC2 instance polls the Amazon SQS queue for new messages, indicating that a new video is ready for processing. The instance uses an Amazon S3 link in the Amazon SQS message to download the new file. FFmpeg is used to extract one frame every second. A Lambda function is triggered and posts the preprocessing results in an AWS IoT topic. After this preprocessing work is completed, the instance uploads every processed frame back to the Amazon S3 video bucket, and stores the frame information in an Amazon DynamoDB table in batches. To optimize processing time and control the parallel Lambda function executions, batch sizes are calculated by the length of video being processed. For example, a thirty-minute video batches in sets of 20, a two-hour video batches in sets of 60, and videos longer than two hours batch in sets of 80.

Image Processing

When an EC2 instance stores frame information in the Amazon DynamoDB table, an image processing AWS Lambda function is triggered. This Lambda function calls Amazon Rekognition to identify metadata for each frame, and saves the metadata in the DynamoDB table. The function also creates and manages Amazon Rekognition face collections which store faces as searchable image vectors. The results of the image processing are posted in an AWS IoT topic. After all the frames for a video are processed, a final Lambda function processes all of the individual frame metadata for the video to create a consolidated list of tags and labels, which is then stored in Amazon DynamoDB, Amazon S3, and sent to an Amazon SNS topic to alert the subscriber of the new tags.

> **Note:** This solution uses the video filename as the prefix for organizing copies of each frame in Amazon S3. Note this as you consider your video naming convention, and ensure that each video to be processed has a unique name.

AWS IoT Topic

This solution includes an AWS IoT topic as an optional feature that you can use to monitor the video processing status for individual or groups of videos. See AWS IoT Topic for detailed information.

Face Search Component

Deploying this solution with the **default parameters** builds the following face-search component in the AWS Cloud.
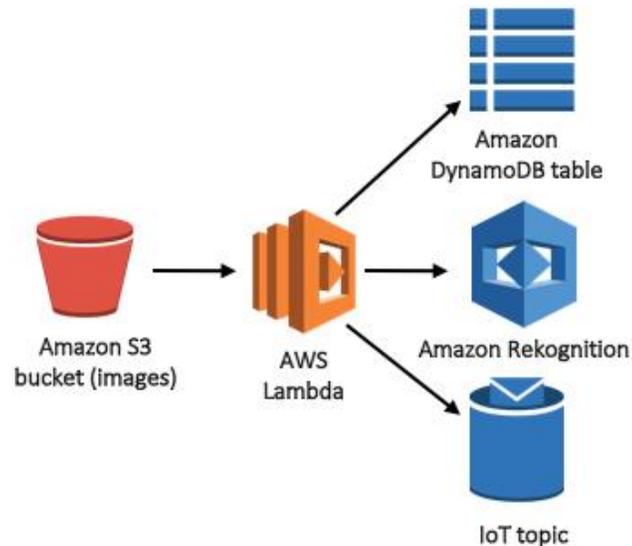

Figure 2: Face Search Architecture

This solution creates another Amazon S3 bucket to store face images for searching. When you upload a face image to this bucket, it triggers an AWS Lambda function that initiates a search to find matching faces in the previously-extracted video metadata. Another AWS Lambda function retrieves the Amazon Rekognition collections from the DynamoDB table and searches those collections for the image. The results of the search are stored in an Amazon DynamoDB table, and the results status of the search are sent to an AWS IoT Topic.

Implementation Considerations

Supported File Types

Currently, the Frame-based Analysis for Your Videos solution supports MP4 videos for frame processing and JPG images for facial searching. All face search images must use a .jpg extension. The solution will *not* process JPG images with a .jpeg extension.

Amazon S3 Server-Side Encryption (SSE)

AWS highly recommends that customers encrypt sensitive data in transit and at rest. This solution creates bucket policies on the video and face search buckets requiring objects to be encrypted at rest with Amazon S3 SSE. As a result, SSE must be enabled when uploading

videos and images, such as using the `--sse AES256` option with the AWS CLI. The following is a complete example for uploading a video using the AWS CLI:

```
aws s3 cp video_file.mp4 s3://$video_bucket/ --metadata topic=iot-
topic-name --sse AES256
```

AWS IoT Topic

Each stage of the image search and video processing workflows are designed to provide status messages to an AWS IoT topic. This is an optional feature that you can use to monitor the video processing status for your videos.

To use this feature, create an `x-amz-meta-topic` header with a value of the desired AWS IoT topic name for status messages when you upload an image or video for processing. This header could contain a different topic name for each video, for each user, or for batches of videos to be processed, depending on the desired amount of video processing status granularity. This header can be configured manually through the Amazon S3 console or when uploading files with the AWS CLI with the `--metadata topic=iot-topic-name` option.

```
aws s3 cp video_file.mp4 s3://$video_bucket/ --metadata topic=iot-
topic-name --sse AES256
```

Limits

This solution automatically uses managed Amazon Rekognition collections. It stores 100,000 face identifiers per collection. Once this number is reached, a new collection will be automatically created to index subsequent facial frames. When an image file is submitted to the face search feature, the face search Lambda function will search any Amazon Rekognition collections in parallel. If the face matches a facial frame in the collection, the percentage of the match will be shown in the `svbp results` table. The match threshold is currently set at 85%. If you want to increase or decrease the number of face identifiers per collection, you can modify the **Max Collection Size** parameter in the solution template.

Currently, this solution has been configured to support up to two parallel FFmpeg EC2 instances. While the **Max Video** parameter is configurable in the solution's AWS CloudFormation template, we do not recommend changing this parameter. However, if you choose to change this parameter, you must request an increase to your AWS Lambda or Amazon Rekognition API limits and configure your Amazon DynamoDB capacity to meet your parallel video processing requirements.

Regional Deployment

This solution uses the Amazon Rekognition service which is currently available in specific AWS Regions only. Therefore, you must launch this solution in an AWS Region where Amazon Rekognition is available. For the most current Amazon Rekognition availability by region, see the [AWS service offerings by region](#).

AWS CloudFormation Template

This solution uses AWS CloudFormation to automate the deployment of the Frame-based Analysis for Your Videos solution on the AWS Cloud. It includes the following CloudFormation template which you can download before deployment:

View template   **video-frame-based-analysis.template:** Use this template to launch the solution and all associated components. The default configuration deploys Amazon Simple Storage Service (Amazon S3) buckets, Amazon DynamoDB tables, AWS Lambda functions, an Amazon Simple Queue Service (Amazon SQS) queue, Amazon Elastic Compute Cloud (Amazon EC2) instances, an Amazon Simple Notification Service (Amazon SNS) topic, an Auto Scaling group, an Amazon Virtual Private Cloud (VPC) network topology, and Amazon Rekognition, but you can also customize the template based on your specific needs.

Automated Deployment

Before you launch the automated deployment, please review the architecture, configuration, and other considerations discussed in this guide. Follow the step-by-step instructions in this section to configure and deploy the Frame-based Analysis for Your Videos solution into your account.

Time to deploy: Approximately 20 minutes

What We'll Cover

The procedure for deploying this architecture on AWS consists of the following steps. For detailed instructions, follow the links for each step.

[Step 1. Launch the stack](#)

- Launch the AWS CloudFormation template into your AWS account.
- Enter values for required parameters: **Video Bucket, Face Search Bucket, EC2 Key Name**, and **SSH location**.
- Review the other template parameters and adjust if necessary.

- Upload a video to Amazon S3 and monitor the results.

- Upload a facial image to search video metadata for matches.

- Search for tags in the Amazon DynamoDB table.

Step 1. Launch the Stack

This automated AWS CloudFormation template deploys Frame-based Analysis for Your Videos on the AWS Cloud.

> **Note**: You are responsible for the cost of the AWS services used while running this solution. See the Cost section for more details. For full details, see the pricing webpage for each AWS service you will be using in this solution.

1. Sign in to the AWS Management Console and click the button to the right to launch the `video-frame-based-analysis` AWS CloudFormation template.
   You can also download the template as a starting point for your own implementation.

   Launch Solution

2. The template is launched in the US East (N. Virginia) Region by default. To launch the solution in a different AWS Region, use the region selector in the console navigation bar.

> **Note**: This solution uses the Amazon Rekognition service, which is currently available in specific AWS Regions only. Therefore, you must launch this solution in an AWS Region where Amazon Rekognition is available. [1]

3. On the **Select Template** page, keep the default settings for **Choose a template** and select **Next**.

4. On the **Specify Details** page, assign a name to your solution stack.

5. Under **Parameters**, review the parameters for the template and modify them as necessary. This solution uses the following default values.

---

[1] For the most current Amazon Rekognition availability by region, see https://aws.amazon.com/about-aws/global-infrastructure/regional-product-services/

Parameter	Default	Description
New Video Bucket	*<Requires input>*	Enter a unique bucket name for the Amazon S3 bucket the template will create to host videos for processing.
New Face Search Bucket	*<Requires input>*	Enter a unique bucket name for the Amazon S3 bucket that the template will create to host images used to search video metadata.
EC2 Key Name	*<Requires input>*	The Amazon EC2 key pair that you will use to securely connect to your instance after it launches. When you created an AWS account, this is the key pair you created in your preferred AWS Region.
SSH Location	*<Requires input>*	IP address range that can access the EC2 instances
Min Videos	1	The minimum number of simultaneous videos that you want the solution to support. The template uses this value to configure the minimum size of the Auto Scaling group, which correlates to the minimum compute capacity required when scaling down.
Max Videos	2	The maximum number of simultaneous videos that this solution should support. The template uses this value to configure the maximum size of the Auto Scaling group, which correlates to the maximum compute capacity required when scaling up. **Note:** AWS recommends that you **do not** increase the default parameter beyond `2`. If you choose to increase this parameter, you must request an increase to your AWS Lambda or Amazon Rekognition API limits, and configure your Amazon DynamoDB capacity to meet your parallel video processing requirements.
Max Lambda Executions	20	The maximum number of Lambda functions used to process each video
Lambda Log Level	INFO	The logging level for the Lambda functions. `Info` level includes error messages. `Debug` level shows information regarding Lambda function execution and includes info and error messages.
Max Collection Size	100,000	The maximum number of faces to be indexed in each Amazon Rekognition collection

6. Choose **Next.**

7. On the **Options** page, choose **Next**.

8. On the **Review** page, review and confirm the settings. Be sure to check the box acknowledging that the template will create AWS Identity and Access Management (IAM) resources.

9. Choose **Create** to deploy the stack.

aws

You can view the status of the stack in the AWS CloudFormation Console in the **Status** column. You should see a status of CREATE_COMPLETE in approximately 20 minutes.

> **Note:** In addition to the primary AWS Lambda functions, this solution includes the `solution-helper` Lambda function, which runs only during initial configuration or when resources are updated or deleted.
>
> When running this solution, the `solution-helper` function is inactive. However, do not delete this function as it is necessary to manage associated resources.

Step 2. Upload a Video for Processing

After the solution successfully launches, you can start uploading MP4 video files for processing. The solution will process videos that you upload to the video Amazon S3 bucket, which you named in the **New Video Bucket** parameter in Step 1. You can upload videos using the AWS Console or AWS CLI.

> **Note:** Do not use a prefix when you upload videos to the S3 bucket.

The following example demonstrates how to use the AWS CLI to upload a video:

```
aws s3 cp video_file.mp4 s3://$video_bucket/ --metadata topic=iot-
topic-name --sse AES256
```

After you upload a video, you can check the results in several different places. Processing log files are available in Amazon CloudWatch Logs under the `video-frame-based-analysis-nodes` log group. Check the `preprocess.out.log` and `preprocess.err.log` log streams for log information about FFmpeg preprocessing. You can also check for extracted frames; they are stored in the same bucket as your videos and have the prefix `/images/$video_name.mp4/`.

Video processing results are stored in the `RVA_VIDEOS_RESULTS_TABLE`, `RVA_PROCESS_TABLE`, and `RVA_FRAMES_RESULTS_TABLE` Amazon DynamoDB tables.

Step 3. Perform a Facial Search

After a video file is successfully processed, you have the option to upload a JPG image of a face and perform a facial search on your video metadata. The solution will process JPG images that you upload to the face search Amazon S3 bucket, which you named in the **New Face Search Bucket** parameter in Step 1. You can upload images using the AWS Console or AWS CLI.

> **Note:** This solution is configured to process images uploaded with an `upload/` prefix.

The following example demonstrates how to use the AWS CLI to upload an image:

```
aws s3 cp video_face-image.jpg s3://$image_bucket/upload/ --metadata
topic=iot-topic-name --sse AES256
```

Image search results are stored in Amazon DynamoDB tables. The `svbp_processing` table contains a mapping between a solution-created object ID and the original S3 image path. The `svbp_results` table contains image search results, including a list of the Amazon Rekognition collection, face ID, and Similarity score, for all video face collections containing faces that match the image. Empty results indicate that the image does not match any collections.

Step 4. Search by Tag

The solution creates tags with the timeline for each frame that is processed and consolidates the information. You can use these tags with your own custom search by tag user interface, or store them in Amazon Elasticsearch for future analysis.

Once a video is processed, navigate to the `RVA_VIDEOS_LABELS_TABLE` in Amazon DynamoDB. You can use this table to see the timeline of the tags in each video that is processed. The scan and filter options allow you to search all the uploaded videos for specific tags. Each Amazon DynamoDB item contains the video file, frame image name, labels detected by Amazon Rekognition, and the time (in milliseconds) where the frame image appears in the video. The tag results are uploaded to the Amazon S3 video bucket with a `/labels/` prefix. Once the files are uploaded, an Amazon SNS topic notifies subscribed users of the location of the `labels` files.

Security

When you build systems on AWS infrastructure, security responsibilities are shared between you and AWS. This shared model can reduce your operational burden as AWS operates, manages, and controls the components from the host operating system and virtualization layer down to the physical security of the facilities in which the services operate. For more information about security on AWS, visit the [AWS Security Center](#).

Image and video files are uploaded using Amazon Simple Storage Service (Amazon S3) SSL endpoints using the HTTPS protocol. [Server-Side Encryption](#) (SSE) is required to ensure all files are 256-bit Advanced Encryption Standard (AES-256) encrypted at rest.

Additional Resources

AWS services

- [AWS CloudFormation](#)
- [Amazon DynamoDB](#)
- [Amazon SQS](#)
- [AWS Lambda](#)
- [Amazon Rekognition](#)
- [Amazon EC2](#)
- [Amazon S3](#)
- [AWS IoT](#)

Appendix: Collection of Anonymous Data

This solution includes an option to send anonymous usage data to AWS. We use this data to better understand how customers use this solution and related services and products. When enabled, the following information is collected and sent to AWS each time the AWS Lambda function runs:

- **Solution ID:** The AWS solution identifier

- **Unique ID (UUID):** Randomly generated, unique identifier for each deployment

- **Timestamp:** Data-collection timestamp

- **Processed frames:** The number of frames the solution processes

- **Processed videos:** The number of videos the solution processes

- **Processed images:** The number of images the solution processes

Note that AWS will own the data gathered via this survey. Data collection will be subject to the [AWS Privacy Policy](). To opt out of this feature, complete one of the following tasks:

a) Modify the AWS CloudFormation template mapping section as follows:

```
"Send" : {
"AnonymousUsage" : { "Data" : "Yes" }
},
```

 to

```
"Send" : {
"AnonymousUsage" : { "Data" : "No" }
},
```

OR

b) After the solution has been launched, find the `metrics` function in the Lambda console and set the **SEND_ANONYMOUS_DATA** environment variable to `No`.

Send Us Feedback

We welcome your questions and comments. Please post your feedback on the [AWS Solutions Discussion Forum](#).

You can visit our [GitHub repository](#) to download the templates and scripts for this solution, and to share your customizations with others.

Document Revisions

Date	Change	In sections
November 2017	*Initial Release*	-