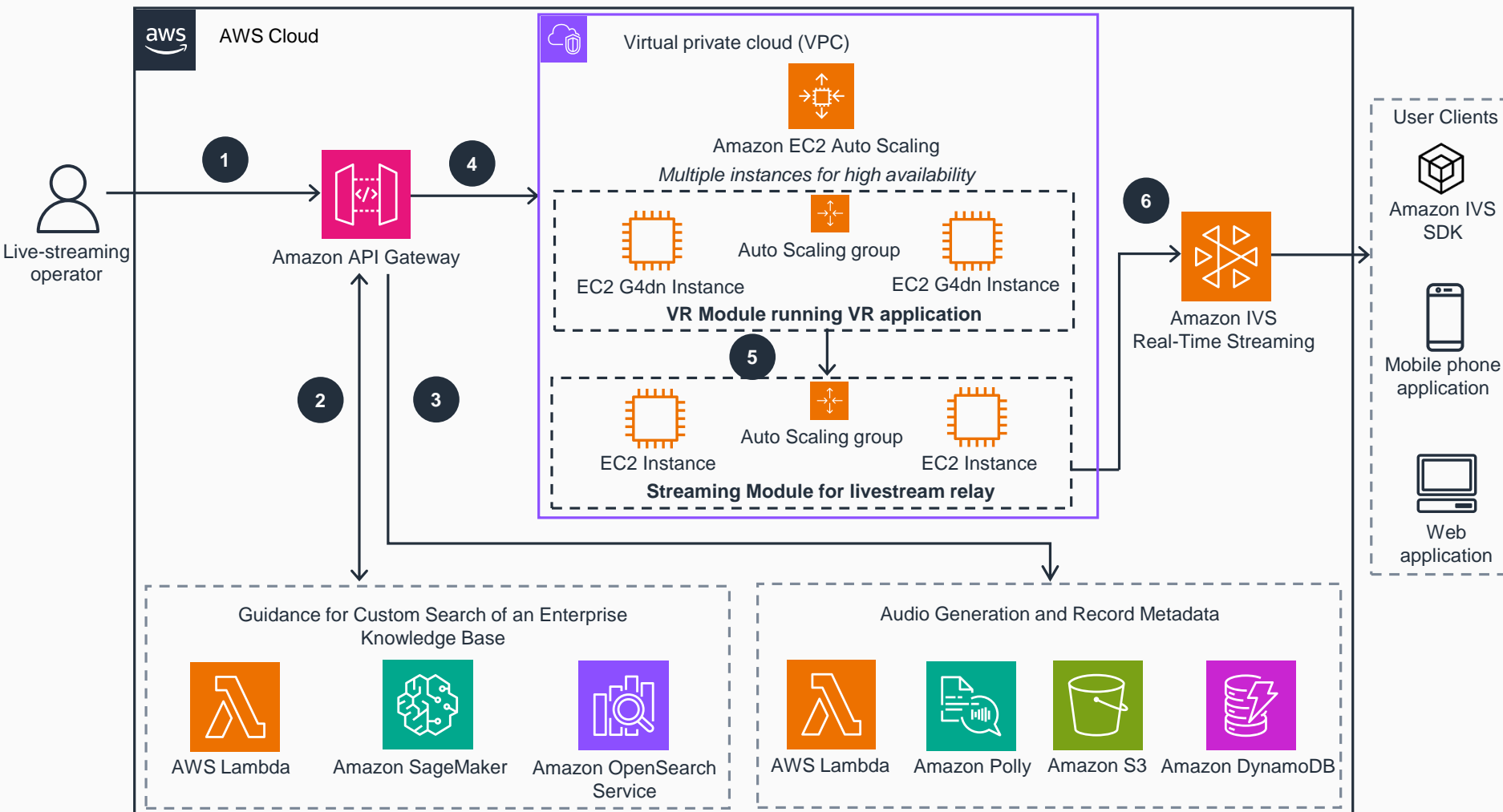


# Guidance for Livestreams Hosted with Digital Humans on AWS

This architecture diagram helps you build an end-to-end virtual human livestreaming solution with cloud rendering, cloud livestreaming, and AI components.



- 1 The livestreaming operator uses APIs or frontend pages that encapsulate APIs to send control commands of the digital human streamer or viewers' questions to **Amazon API Gateway**.
- 2 **API Gateway** passes the questions to the Search & Question Answering (QA) system, powered by a large language model (LLM), and gets back suggested answers. The QA system can be built based on **Guidance for Custom Search of an Enterprise Knowledge Base on AWS**. All APIs are configured with authentication. **Amazon CloudWatch** monitors for **AWS Lambda** functions and API calls.
- 3 **Lambda** uses **Amazon Polly** to convert answers into a voice file, stores the voice file in **Amazon Simple Storage Service (Amazon S3)**, and saves the metadata in **Amazon DynamoDB**.
- 4 **API Gateway** passes the control commands of the digital human to a virtual reality (VR) module running a VR application hosted on **Amazon Elastic Compute Cloud (Amazon EC2)**. We recommend using **EC2 G4dn** instances. **Amazon EC2 Auto Scaling** enhances availability.
- 5 The VR module runs the digital human module, renders images into video streams, and pushes streams to the **EC2** instances hosting the streaming module.
- 6 **Amazon Interactive Video Service (Amazon IVS)** Low-Latency Streaming distributes the livestream to **Amazon IVS SDK** or viewers' mobile phone and web applications. We recommend **Amazon IVS** real-time streaming to reduce latency.

