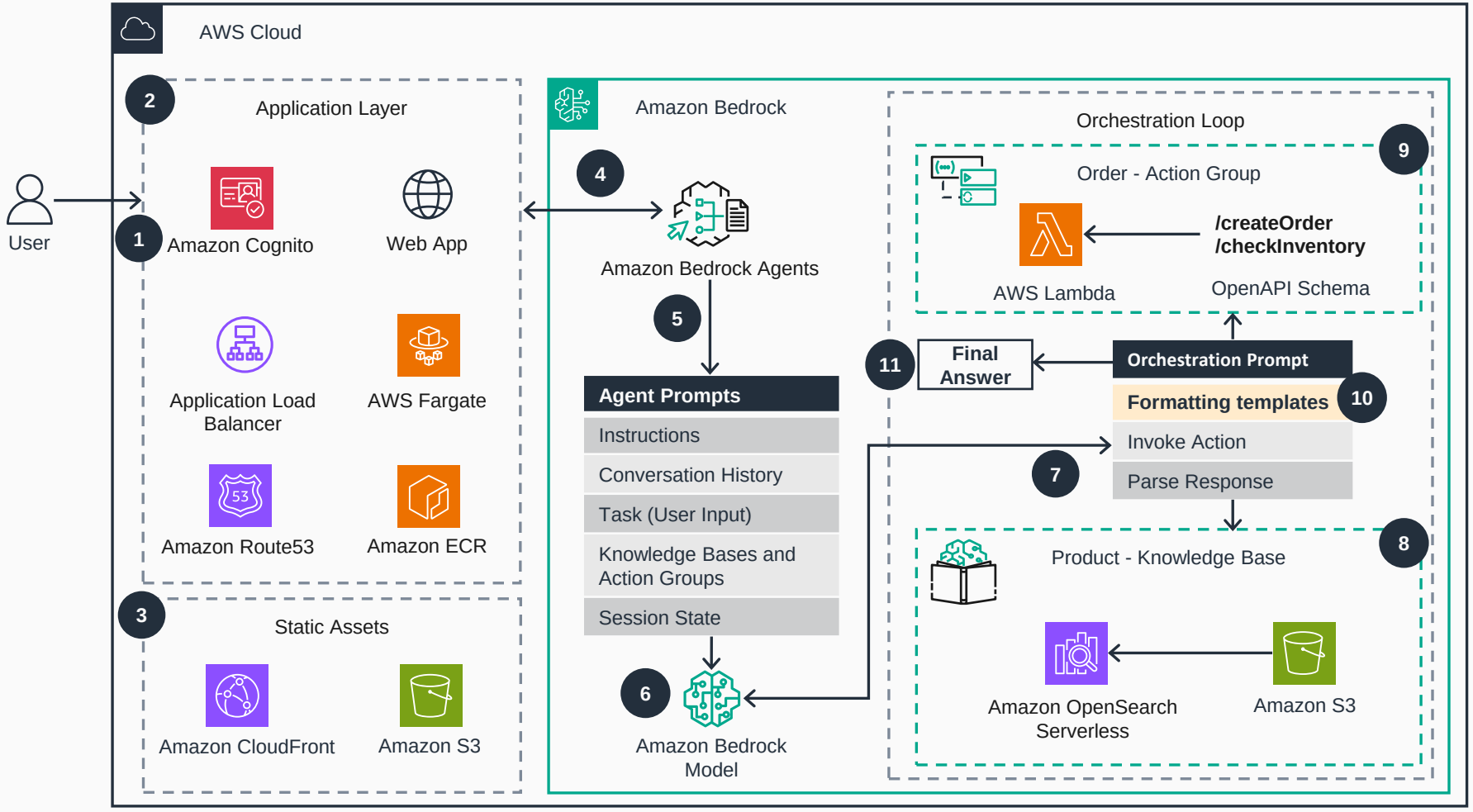
Guidance for Generative AI Shopping Assistants Using Amazon Bedrock Agents

This architecture diagram shows how to use Amazon Bedrock Agents and Amazon Bedrock Knowledge Bases to build a generative AI shopping assistant based on existing resources and automated tasks. This slide details Steps 1-7.



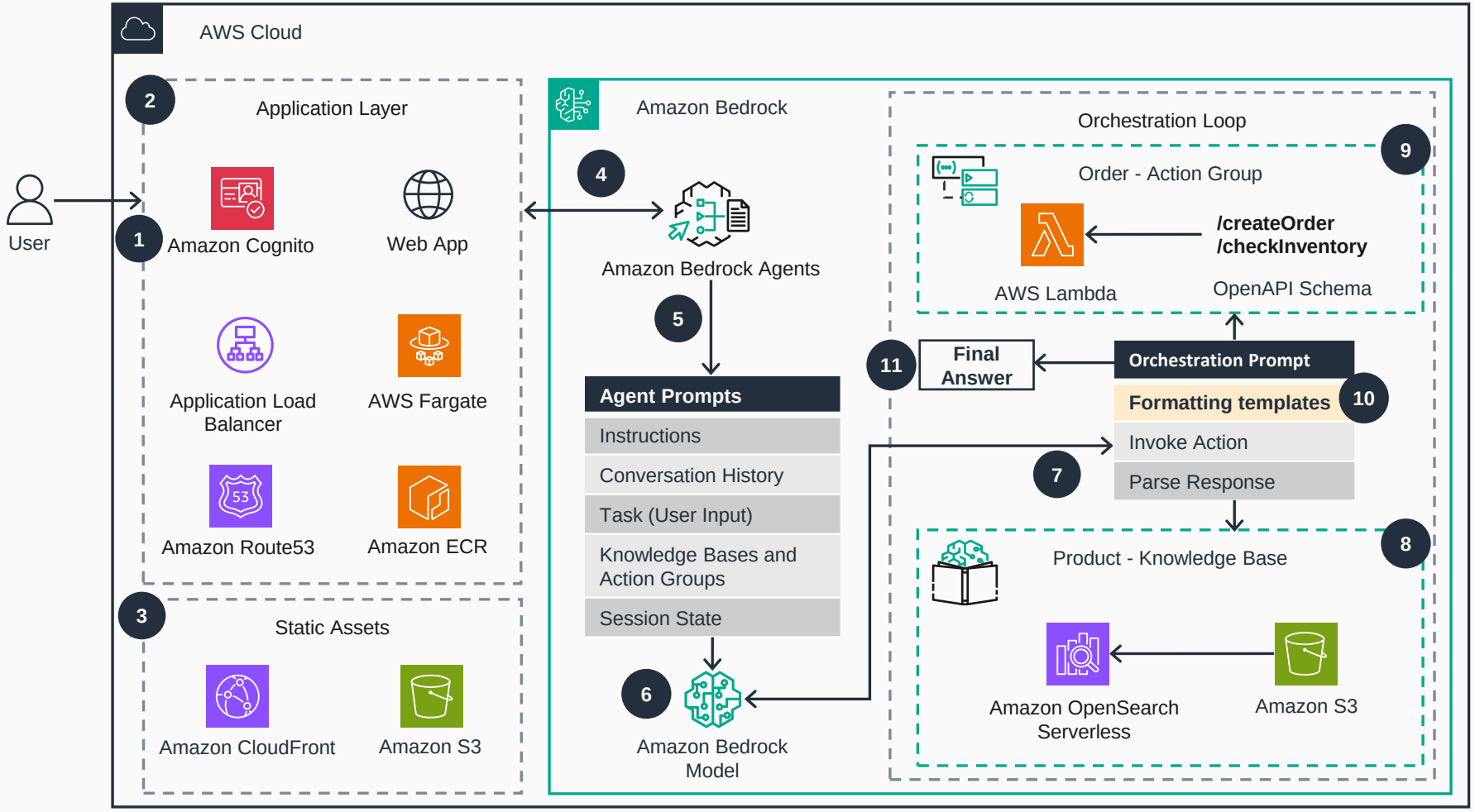
1. A user provides natural language queries to find products or place orders through the application interface.
2. The application is hosted using serverless containers on **AWS Fargate**, **Amazon Elastic Container Registry (Amazon ECR)** for managing container images, **Elastic Load Balancing (ELB)** for traffic distribution, **Amazon Route 53** for DNS, and **Amazon Cognito** for authentication.
3. Static assets, such as product images, are stored in **Amazon Simple Storage Service (Amazon S3)** with **Amazon CloudFront** distribution to provide low latency content delivery for an enhanced user experience.
4. The user query is sent to **Amazon Bedrock Agents** using an **Amazon Bedrock API**.
5. **Amazon Bedrock Agents** interprets user input using conversation history, agent instructions, and other configurations, in addition to the underlying foundation model (FM) offered by **Amazon Bedrock** for contextual understanding of the request, resulting in accurate and relevant responses.
6. The FM uses the given agent instructions including the role of the assistant, specific tasks, step-by-step instructions for completing tasks, and examples, to orchestrate actions for assisting the user in finding and purchasing products.
7. During orchestration, **Amazon Bedrock Agents** uses ReAct prompting with the Orchestration prompt to invoke relevant actions for finding products or placing orders.

Reviewed for technical accuracy October 9, 2024

AWS Reference Architecture

Guidance for Generative AI Shopping Assistants Using Amazon Bedrock Agents

This architecture diagram shows how to use Amazon Bedrock Agents and Amazon Bedrock Knowledge Bases to build a generative AI shopping assistant based on existing resources and automated tasks. This slide details Steps 8-11.



8 To find relevant products from the catalog, **Amazon Bedrock Knowledge Bases** offers fully managed retrieval augmented generation (RAG) for **Amazon Bedrock Agents**. It uses product documents stored in **Amazon S3** as a data source for embedding product information and **Amazon OpenSearch Serverless** for vector store and semantic search.

9 To manage API operations such as order creation, inventory checks, and sending email confirmations, **Amazon Bedrock Agents** uses an Order action group built using an **AWS Lambda** function and the OpenAPI schema.

10 Using Advanced prompts for Orchestration, **Amazon Bedrock Agents** is configured with templates for formatting email confirmations and output responses to users for clear and consistent communication.

11 The **Amazon Bedrock Agents** reasoning process continues until the agent finds relevant products, creates an order with email confirmation, or wants to request additional details from the customer. The final answer is sent back to the application, providing a seamless shopping experience for the user.

Reviewed for technical accuracy October 9, 2024

AWS Reference Architecture