DocuPanda: Network Security and Deployment

DocuPanda Network Security and Deployment

Overview

DocuPanda processes documents securely using AWS and Azure. This document outlines the architecture, security measures, and deployment of intrusion detection and prevention systems to safeguard the network.

Key Components

1. **API Gateway**
   - **Hosted in:** AWS
   - **Purpose:** Entry point for external document processing requests.
   - **Protection:** AWS Web Application Firewall (WAF)
     - Rate limiting to prevent abuse.
     - IP blocklisting and allowlisting.
     - SQL injection and XSS filtering.
2. **Load Balancers and Autoscaling Groups**
   - **AWS and Azure Load Balancers** distribute traffic to auto-scaling instances.
   - **Isolated** within private networks to prevent direct internet access.

3. **Document Processing Workflow**
    - **Input:** Raw documents are uploaded through the API Gateway.
    - **Storage:**
        - **Raw documents:** Stored in Azure Blob (primary) and AWS S3 (backup).
        - **Insights and metadata:** Stored in AWS DocumentDB.
    - **Long-term storage:** Processed documents and metadata are archived in EU regions.

Security Configurations

- **Network Access Control**
  Security Groups and Network ACLs restrict inbound/outbound traffic to trusted IPs.
- **Encryption**
    - **Data in Transit:** TLS 1.2+
    - **Data at Rest:** AES-256 using AWS KMS and Azure Key Vault.
- **Identity and Access Management (IAM)**
  Minimal privileges for roles interacting with DocuPanda resources.
- **Intrusion Detection and Prevention**
    - **AWS GuardDuty**:
        * Monitors and detects unauthorized activities (e.g., reconnaissance, instance compromise).
        * Integrated with AWS WAF and API Gateway for automated threat response.
    - **Azure Security Center**:
        * Detects vulnerabilities and provides real-time threat alerts.
        * Automatically applies preventive measures to Azure Load Balancers and Blob Storage.
    - **Azure Firewall with Threat Intelligence**:
        * Prevents known malicious traffic based on real-time threat intelligence feeds.

Regional Compliance

- **Primary Processing:** AWS and Azure (US regions for low-latency operations).
- **Long-term Archival:** All sensitive data is transferred and stored in EU regions for regulatory compliance.

Network Diagram

Below is the network diagram for DocuPanda's system, now including IDS/IPS:

here's the image:

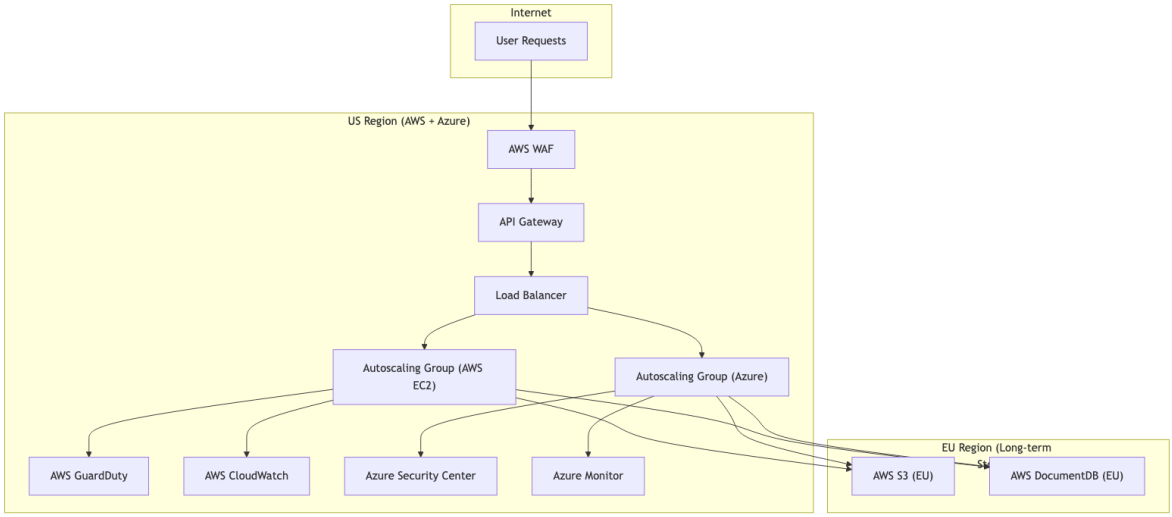

Figure 1: Network Diagram