AWS Elastic Load Balancing

Topics : AWS
Written on December 01, 2023

Amazon Elastic Load Balancing (ELB) is a service provided by Amazon Web Services (AWS) that automatically distributes incoming application traffic across multiple Amazon EC2 instances. This helps ensure that no single instance is overwhelmed with too much traffic and improves the fault tolerance and availability of your application. ELB is an integral part of building scalable and reliable applications in the AWS cloud.

Here are key features and concepts related to AWS Elastic Load Balancing:

1. **Load Balancer Types:**

   - **Application Load Balancer (ALB):** Operates at the application layer (Layer 7) and is best suited for routing HTTP/HTTPS traffic. ALB supports content-based routing, enabling more advanced routing decisions based on the content of the request.
   - **Network Load Balancer (NLB):** Operates at the transport layer (Layer 4) and is designed for handling TCP/UDP traffic. NLB is capable of handling high-throughput and low-latency requirements and is often used for performance-intensive workloads.

2. **Classic Load Balancer:**

   - The Classic Load Balancer provides basic load balancing across multiple EC2 instances. It works at both the application and transport layers.

3. **Listeners and Rules:**

   - A listener is a process that checks for connection requests. For ALB, listeners are configured to check for specific protocols and ports.
   - Rules are conditions defined on the listener to route traffic to different target groups based on factors like the path or host of the URL.

4. **Target Groups:**

   - Target groups are used to route requests to registered targets, such as EC2 instances, IP addresses, or Lambda functions.
   - ALB supports multiple target groups, allowing for more granular routing based on conditions.

5. **Auto Scaling Integration:**

   - ELB can work seamlessly with Auto Scaling groups to automatically register and

deregister instances based on demand.

6. **SSL/TLS Offloading:**

   - ELB can handle SSL/TLS termination, offloading the burden of encrypting and decrypting traffic from the backend instances.

7. **Health Checks:**

   - ELB regularly checks the health of registered instances. Unhealthy instances are automatically taken out of service until they are restored.

8. **Cross-Zone Load Balancing:**

   - Distributes incoming traffic evenly across instances in all enabled Availability Zones. This improves fault tolerance and ensures even load distribution.

9. **Connection Draining:**

   - Ensures that in-flight requests to instances are completed before taking the instances out of service during scaling activities or unhealthy instance detection.

10. **Access Logs:**

    - ELB can generate access logs, which contain detailed information about requests made to the load balancer.

11. **Tagging:**

    - ELB supports tagging, allowing users to label and organize their load balancers for easier management.