AWS Global Accelerator

1. What is AWS Global Accelerator?

AWS Global Accelerator is a **networking service** that improves the **availability, performance, and security** of applications running on AWS across multiple Regions. It provides **static IP addresses** and **routes traffic to the optimal endpoint** based on health, geography, and routing policies.

2. Key Features of AWS Global Accelerator

a. Static Anycast IPs

- AWS Global Accelerator provides **two static IP addresses** that act as fixed entry points for applications.
- These IPs don't change, even if backend resources change, making it easier to manage.

b. Traffic Routing Optimization

- Uses **AWS global network** instead of the public internet for lower latency and better reliability.
- Routes traffic to the nearest AWS Region based on **health and performance metrics**.

c. Automatic Health Checks & Failover

- Continuously monitors endpoints and **automatically redirects traffic** in case of failure.

- Uses **EC2 instances, ALBs (Application Load Balancers), or NLBs (Network Load Balancers)** as endpoints.

d. DDoS Protection via AWS Shield

- Integrated with **AWS Shield** for **DDoS protection**.
- Enhances security by minimizing exposure to the public internet.

e. Multi-Region Redundancy

- Distributes traffic across multiple AWS Regions, **increasing fault tolerance**.

f. Weighted Traffic Distribution

- Allows **proportional traffic routing** across multiple endpoints based on predefined weights.

3. How AWS Global Accelerator Works

1. **User Traffic Initiation:**
   - A client (user) requests access to an application using one of the two **static IPs** provided by AWS Global Accelerator.
2. **Optimal Path Selection:**
   - The request is routed through the **AWS global network** rather than the **public internet**, reducing **latency and packet loss**.
3. **Health-Based Routing:**
   - AWS Global Accelerator **monitors the health of endpoints** (ALB, NLB, EC2) and directs traffic to the most responsive endpoint.
4. **Automatic Failover:**
   - If an endpoint becomes **unhealthy**, traffic is **automatically redirected** to the nearest healthy endpoint.
5. **Application Delivery & Response:**
   - The backend service processes the request and sends the response through AWS Global Accelerator, ensuring **low-latency and high-availability**.

4. AWS Global Accelerator vs. Route 53 vs. CloudFront

Feature	AWS Global Accelerator	Amazon Route 53	Amazon CloudFront
Purpose	Improves performance & availability across multiple Regions	DNS-based routing	Content Delivery Network (CDN)
Routing Method	Uses **Anycast IPs** for intelligent routing	Uses **DNS-based routing**	Uses **edge locations** to cache content
Latency Optimization	Uses **AWS global network** to route traffic faster	Depends on DNS resolution time	Caches static content close to users
Automatic Failover	Yes	Yes	No (Cache stays active until TTL expires)
DDoS Protection	Integrated with AWS Shield	Can be combined with AWS Shield	Uses AWS Shield Standard

5. Use Cases for AWS Global Accelerator

✔ **Multi-Region Applications** – Ideal for apps needing **high availability and redundancy**.
✔ **Low-Latency API Endpoints** – Ensures **fast response times** for global users.
✔ **Disaster Recovery (DR) Strategies** – Automatically redirects traffic to another Region in case of failures.
✔ **Gaming Applications** – Reduces lag and improves gaming performance.
✔ **Live Streaming & Media Applications** – Optimizes video/audio delivery globally.

6. Steps to Set Up AWS Global Accelerator

1. **Go to AWS Global Accelerator Console**
2. **Create an Accelerator**
   - Choose **standard or custom routing**.
   - Define **two static IPs** assigned to the accelerator.
3. **Add Listeners**
   - Configure listener ports (e.g., **80 for HTTP, 443 for HTTPS**).
4. **Create Endpoints**
   - Attach **ALBs, NLBs, or EC2 instances** across different AWS Regions.
5. **Enable Health Checks**
   - AWS Global Accelerator monitors and redirects traffic based on health.
6. **Test & Deploy**
   - Verify **latency improvement and failover handling**.

7. Pricing Model

AWS Global Accelerator pricing is based on:

- **Fixed Hourly Rate** – Charged per accelerator (e.g., $0.025 per hour).
- **Data Transfer Rates** – Billed for data transferred through the AWS global network.

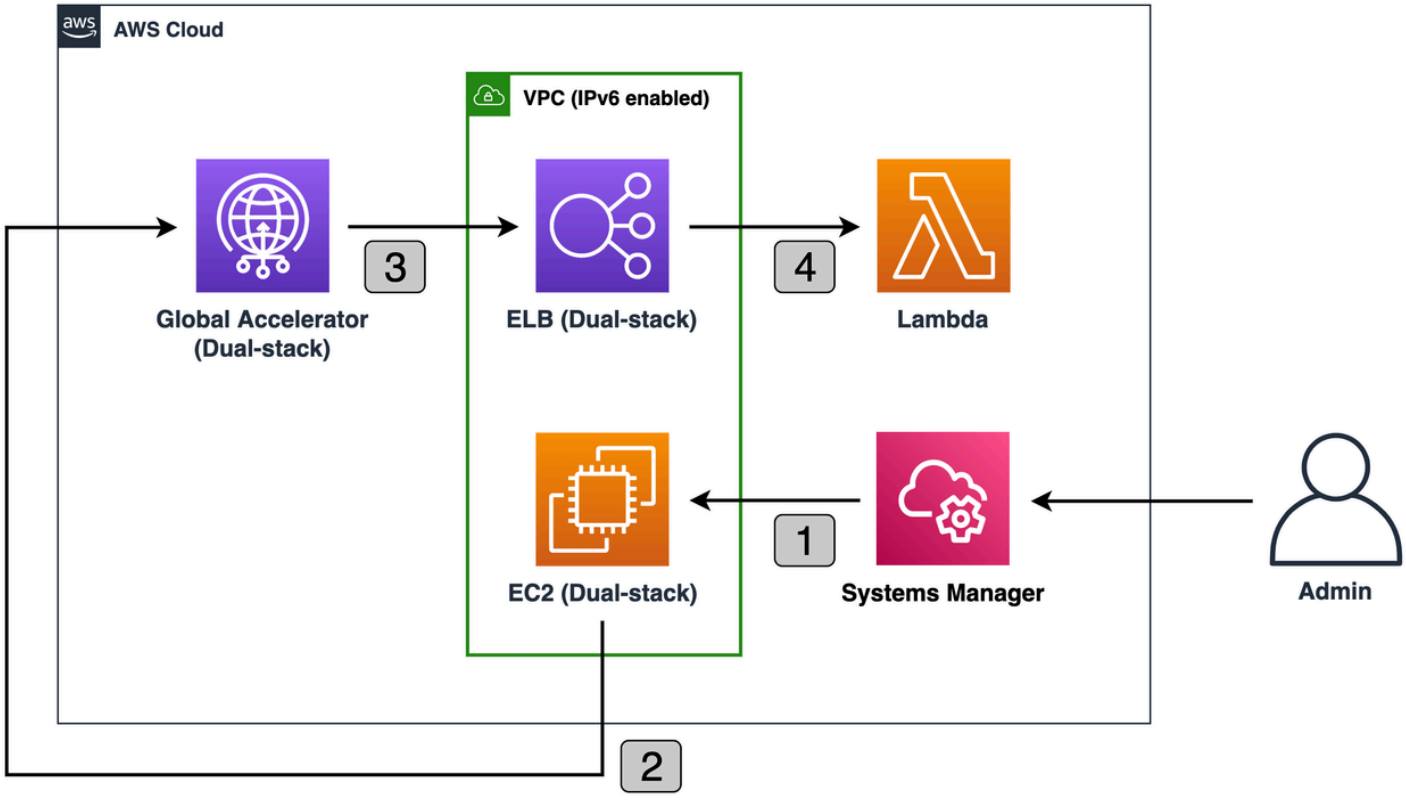
- **Free Tier** – No free tier; charges apply based on usage.

🔹 **Example Calculation**: If you run AWS Global Accelerator for 30 days, the estimated cost is: **$0.025/hour × 24 hours × 30 days = $18/month** + data transfer charges.

8. Pros and Cons of AWS Global Accelerator

✅ **Pros:**
✔ Improves **global application performance**.
✔ Provides **automatic failover & health checks**.
✔ Uses **static IPs** for easier management.
✔ Protects against **DDoS attacks** with AWS Shield.

❌ **Cons:**
✗ **Additional cost** compared to using Route 53 or CloudFront alone.
✗ **Not a CDN replacement** (CloudFront is better for caching).
✗ **Limited to AWS services** (can't route to external endpoints).



9. Conclusion

AWS Global Accelerator is a **powerful service** for organizations needing **low-latency, high-availability, and DDoS protection** across multiple AWS Regions. It's ideal for **multi-region applications, disaster recovery, and API acceleration** but may not be necessary for **simple regional applications**.