**Proposed Solution for Hybrid Working Environment for Faculty and Students**

The existing architecture comprises a basic setup with a single application load balancer, application servers, and a database. This setup faces significant challenges during high-traffic events and DDoS attacks, leading to service disruptions and customer dissatisfaction.

**Tasks & Deliverables**

**1. Improving Scalability and Availability**

**Scalability Improvements:**

* **Auto-Scaling Groups:**
  + Implement auto-scaling groups for application servers. This ensures the system can automatically scale up during high traffic and scale down during low traffic, optimizing resource usage and costs. Auto-scaling policies will be defined based on CPU usage, request rates, and other relevant metrics.
* **Elastic Load Balancer (ELB):**
  + Use an Elastic Load Balancer to distribute incoming traffic across multiple instances, improving availability and fault tolerance. This helps balance the load efficiently, preventing any single instance from becoming a bottleneck.

**Availability Improvements:**

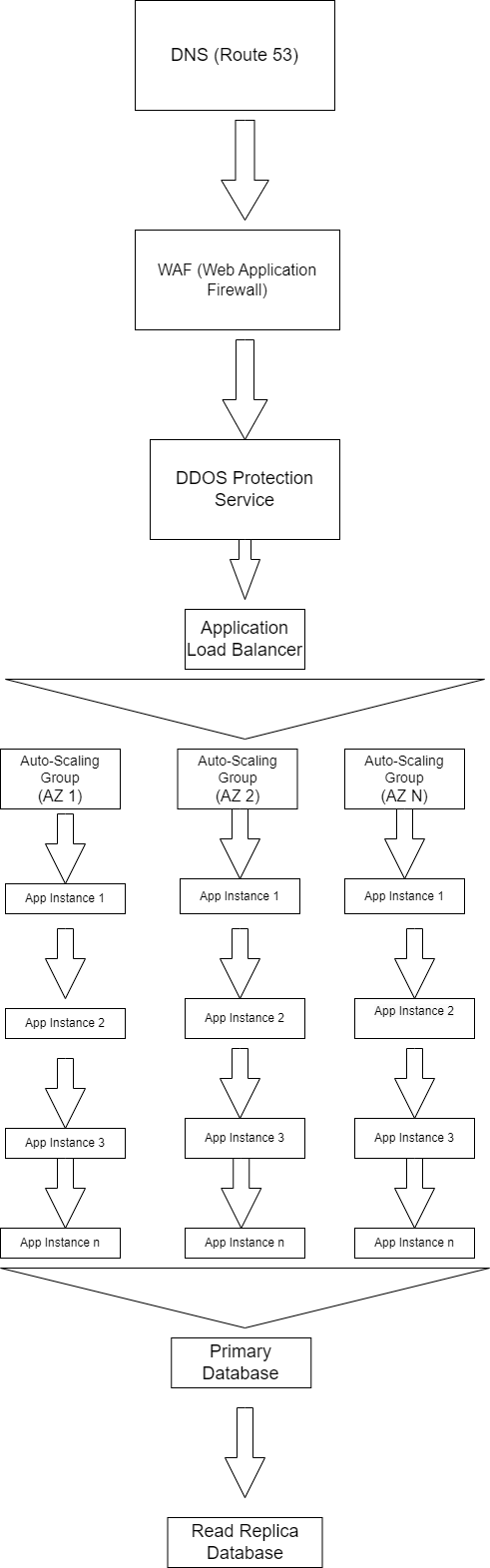
* **Multi-AZ Deployment:**
  + Deploy application servers and databases across multiple Availability Zones (AZs) to ensure high availability and fault tolerance. In the event of an AZ failure, traffic can be redirected to instances in another AZ, ensuring continuous service availability.
* **Database Replication:**
  + Implement a read replica and failover mechanism for the database to ensure high availability and quick recovery in case of a primary database failure. The read replica can also help balance read traffic, reducing the load on the primary database.

**Cost Efficiency:**

* **Spot Instances:**
  + Utilize spot instances for non-critical tasks to reduce costs. Spot instances are significantly cheaper than on-demand instances, although they can be terminated by the cloud provider with little notice.
* **Auto-Scaling Policies:**
  + Define auto-scaling policies to adjust the number of instances based on demand. This ensures cost efficiency during non-peak times while still providing the necessary resources during peak traffic.

**2. New Diagram with Proposed Design Improvements**

**Proposed Architecture Diagram (Flowchart)**



**3. Explanation of Choices**

**Auto-Scaling Groups:**

* **Advantages:**
  + Automatically adjusts resources based on traffic, ensuring optimal performance and cost efficiency.
  + Provides a seamless user experience by maintaining service levels during peak times.
* **Risks:**
  + Misconfigured auto-scaling policies could lead to over or under-provisioning, impacting performance or incurring unnecessary costs.

**Elastic Load Balancer (ELB):**

* **Advantages:**
  + Distributes traffic efficiently, providing fault tolerance and high availability.
  + Prevents any single instance from becoming a bottleneck.
* **Risks:**
  + Single point of failure if not properly configured, though mitigated by multi-AZ deployment.

**Multi-AZ Deployment:**

* **Advantages:**
  + Ensures high availability and disaster recovery capabilities.
  + Increases fault tolerance by distributing instances across different geographical locations.
* **Risks:**
  + Slightly higher costs due to additional instances in multiple AZs.

**Database Replication:**

* **Advantages:**
  + Provides high availability and quick recovery from failures.
  + Balances read traffic, reducing the load on the primary database.
* **Risks:**
  + Data inconsistency if replication lag is not properly managed.

**Cost Efficiency:**

* **Advantages:**
  + Reduces operational costs by using spot instances and optimizing resource usage.
  + Ensures resources are allocated efficiently based on demand.
* **Risks:**
  + Spot instances can be terminated by the cloud provider, so they should be used for non-critical tasks.

**4. Research on DDOS Attacks**

**How DDOS Attacks Occur:**

* **Volume-Based Attacks:**
  + Overwhelm the network bandwidth with a flood of traffic, making it difficult for legitimate traffic to pass.
* **Protocol Attacks:**
  + Exploit vulnerabilities in network protocols, causing service disruptions by exhausting server resources.
* **Application Layer Attacks:**
  + Target specific applications with requests to exhaust resources, leading to slowdowns or crashes.

**5. Vulnerabilities and Resilience Strategies**

**Potential Vulnerabilities:**

* **Volume-Based Attacks:**
  + Can overwhelm network bandwidth, causing service outages.
* **Protocol Attacks:**
  + Can exploit weaknesses in network protocols, leading to disruptions.
* **Application Layer Attacks:**
  + Can exhaust server resources, causing slowdowns or crashes.

**Resilience Strategies:**

* **Web Application Firewall (WAF):**
  + Protects against application layer attacks by filtering malicious traffic and blocking common attack patterns.
* **DDoS Protection Service:**
  + Uses cloud provider’s DDoS protection services to detect and mitigate DDoS attacks, ensuring continuous service availability.
* **Rate Limiting:**
  + Controls the number of requests from a single IP address, preventing any one user from overwhelming the system.
* **Traffic Analysis:**
  + Continuously monitors traffic patterns to detect and respond to anomalies, enabling proactive defense against attacks.

**Advantages of Proposed Solution:**

* **Improved Availability:**
  + Multi-AZ deployment and database replication ensure high availability and quick recovery, minimizing downtime and service disruptions.
* **Enhanced Scalability:**
  + Auto-scaling groups and ELB handle burst traffic efficiently, providing a seamless user experience during peak times.
* **Increased Security:**
  + WAF, DDoS protection, and rate limiting mitigate the risk of attacks, ensuring a secure environment for both users and the application.

**Risks:**

* **Configuration Complexity:**
  + Properly configuring and managing the new components requires expertise, and misconfigurations could lead to vulnerabilities or inefficiencies.
* **Cost Implications:**
  + Additional resources and services may increase costs, though these can be optimized by using scaling policies and cost-effective instance types.

By implementing these improvements, the start-up will be better prepared to handle high traffic during sales events and mitigate potential DDoS attacks, ensuring a reliable and secure shopping experience for customers.