

AWS Summit

AWS 技术峰会 · 北京 2014

基于AWS构建高可用应用架构

王毅

解决方案架构师

November 11, 2014



什么是高可用？

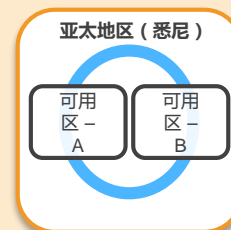
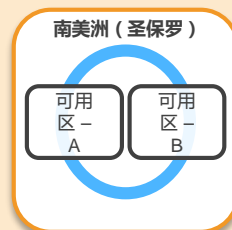
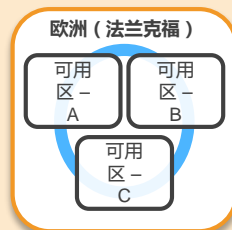
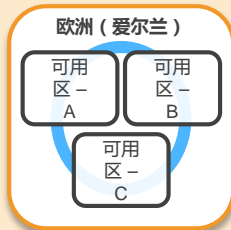
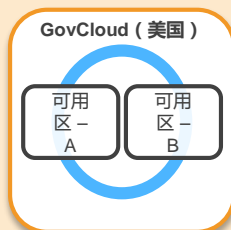
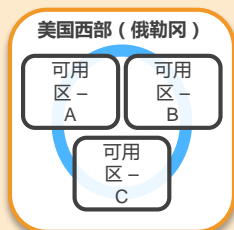
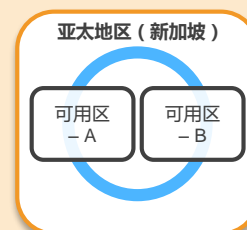
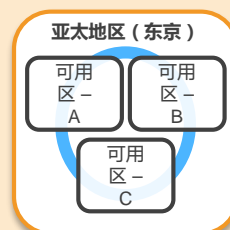
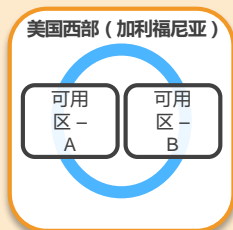
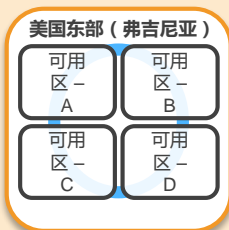
- 可用性
 - 在应用工作周期中可用时间的百分比
- 不可用
 - 应用无法访问，服务中断
 - 应用访问非常缓慢
 - 计划中和非计划中
- 目标
 - 没有宕机，服务一直可用

与可用性相关的事情

- 扩展性
 - 不进行应用设计调整，应用能否满足访问增长
 - 可能会影响可用性
- 容错能力
 - 内建容错能力，应用能够在部分组件失效时继续工作
 - 容错能力对高可用很关键
- 灾备
 - 业务的连续性

AWS的基础设施

全球区域



AWS的服务

天然是高可用和高容错的服务

- ✓ Amazon S3
- ✓ Amazon DynamoDB
- ✓ Amazon CloudFront
- ✓ Amazon Route53
- ✓ Elastic Load Balancing
- ✓ Amazon SQS
- ✓ Amazon SNS
- ✓ Amazon SES
- ✓ Amazon SWF
- ✓ ...

可通过适当的架构设计实现高可用

- ▶ Amazon EC2
- ▶ Amazon EBS
- ▶ Amazon RDS
- ▶ Amazon VPC

设计原则1：假定失效的设计

“Everything fails
all the time”

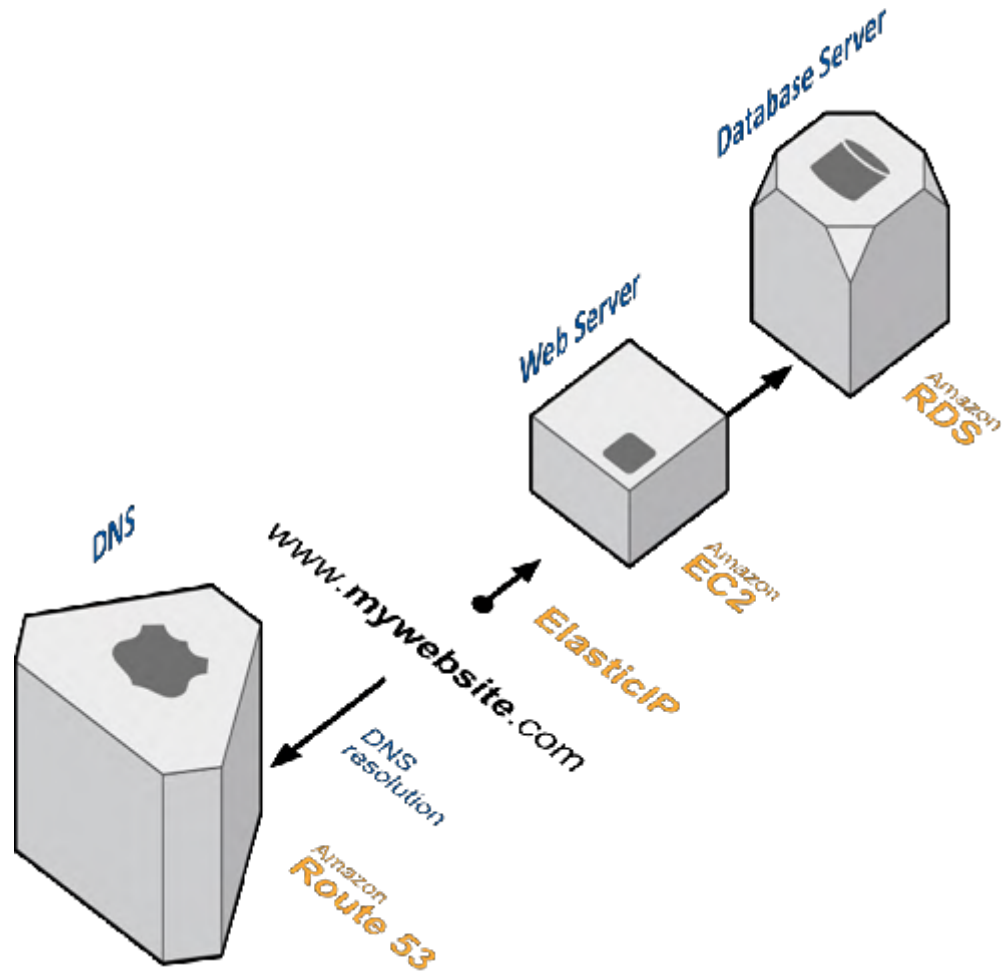


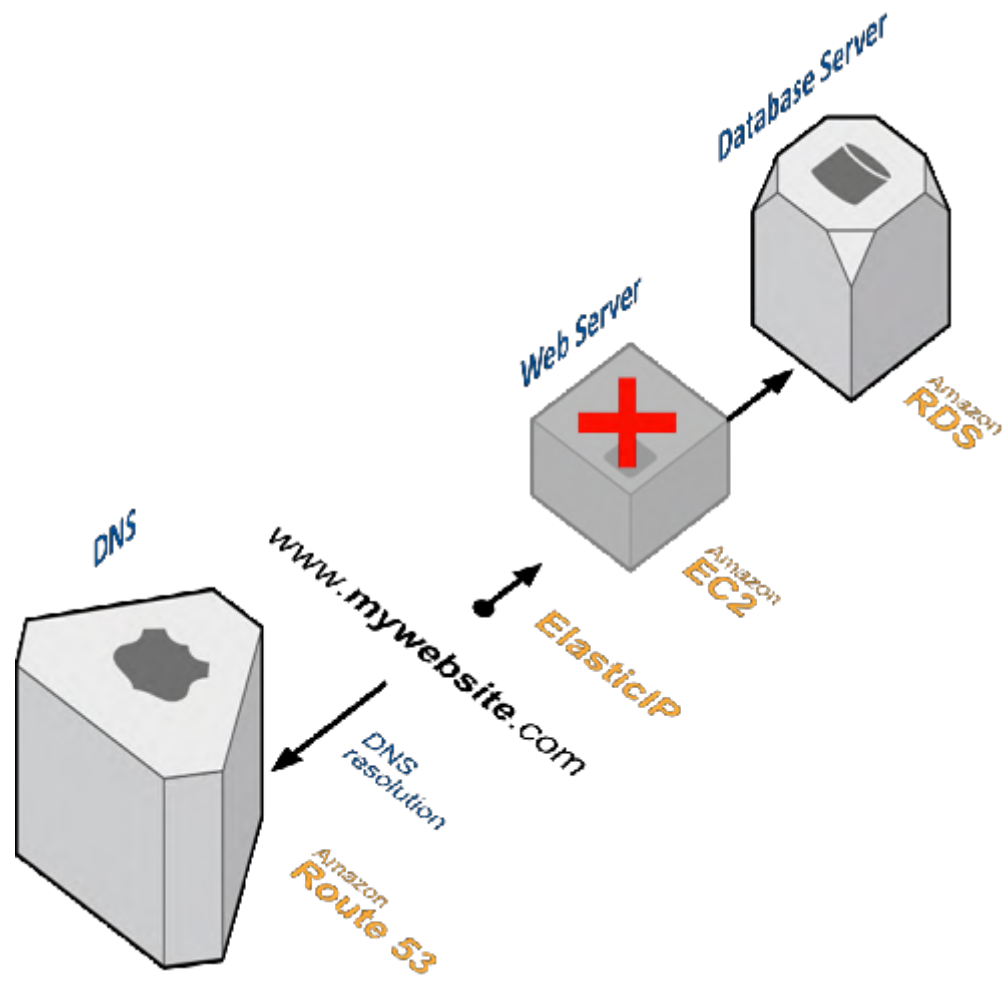
—Werner Vogels
CTO of Amazon

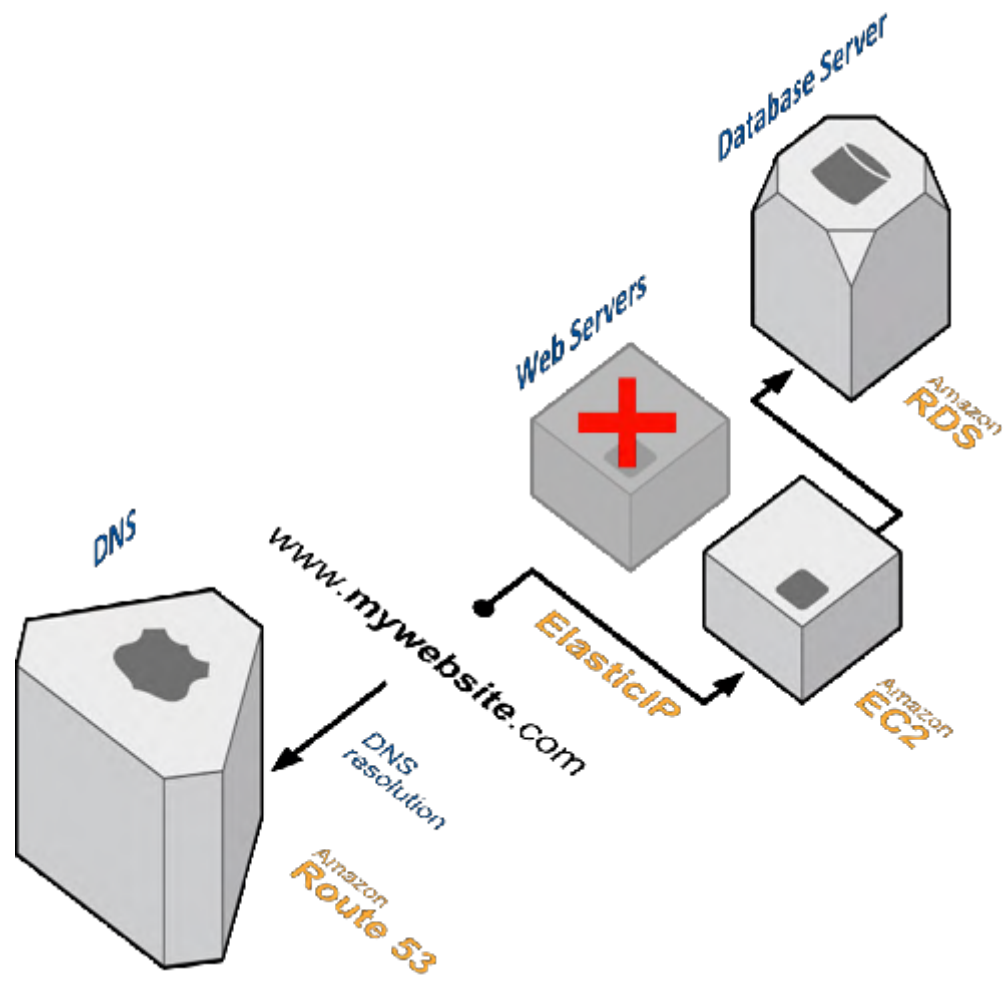
设计原则1：假定失效的设计

- 避免单点故障 (SPoF)
- 假定任何环节都有可能出问题，然后倒推依次设计
- 目标是应用能够连续工作
 - 例如：EIP, EBS, ELB

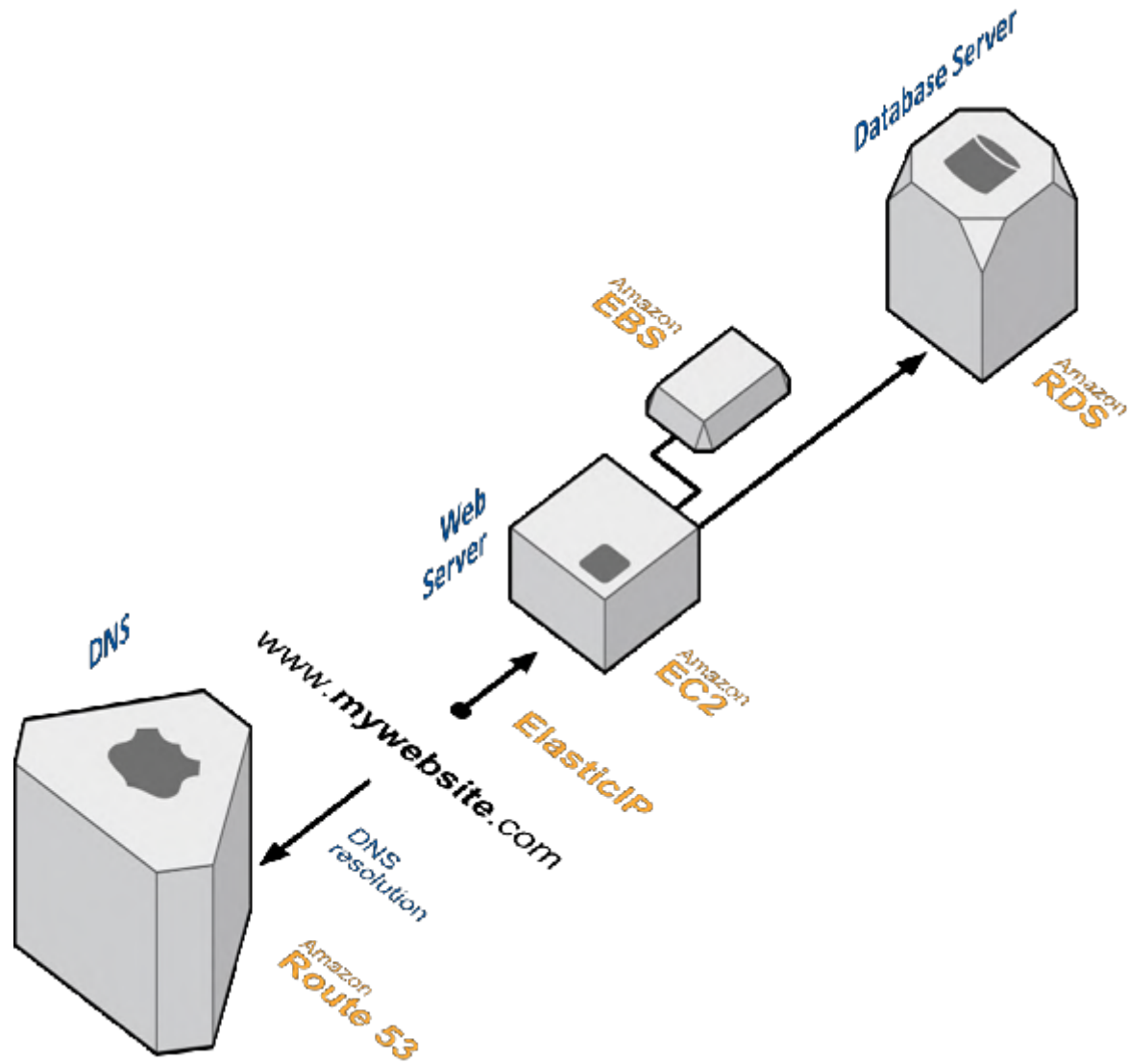
EIP

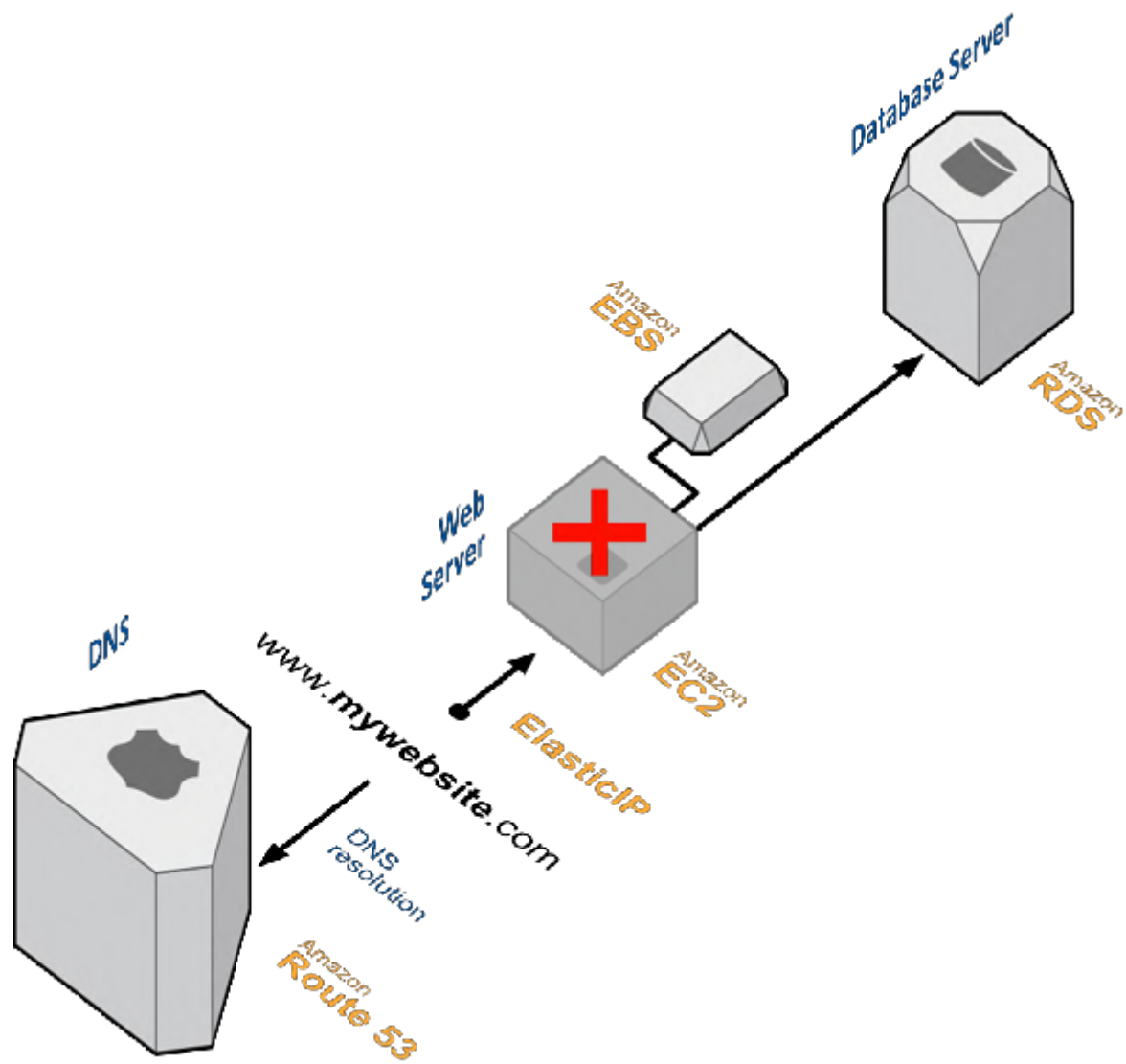


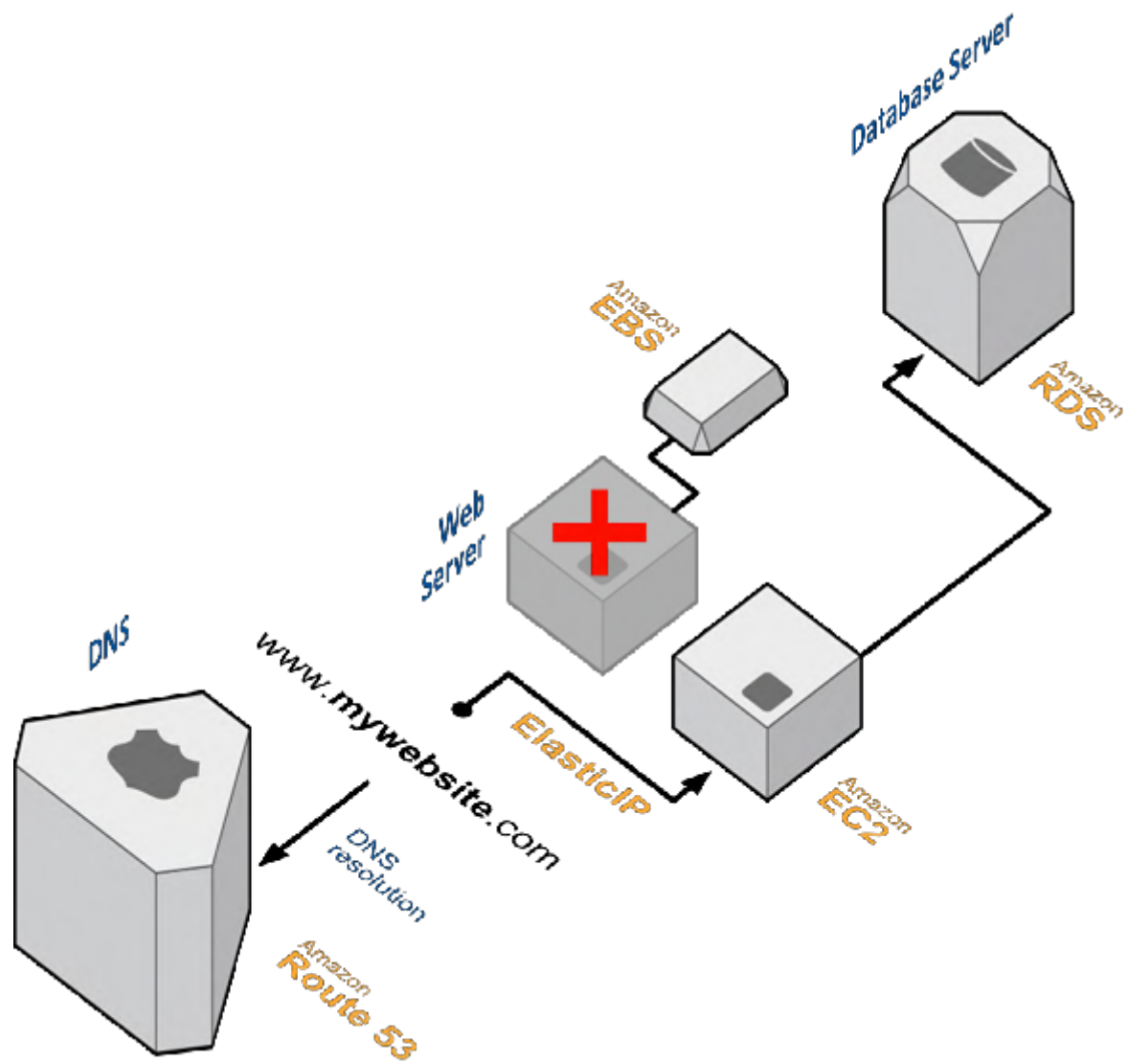


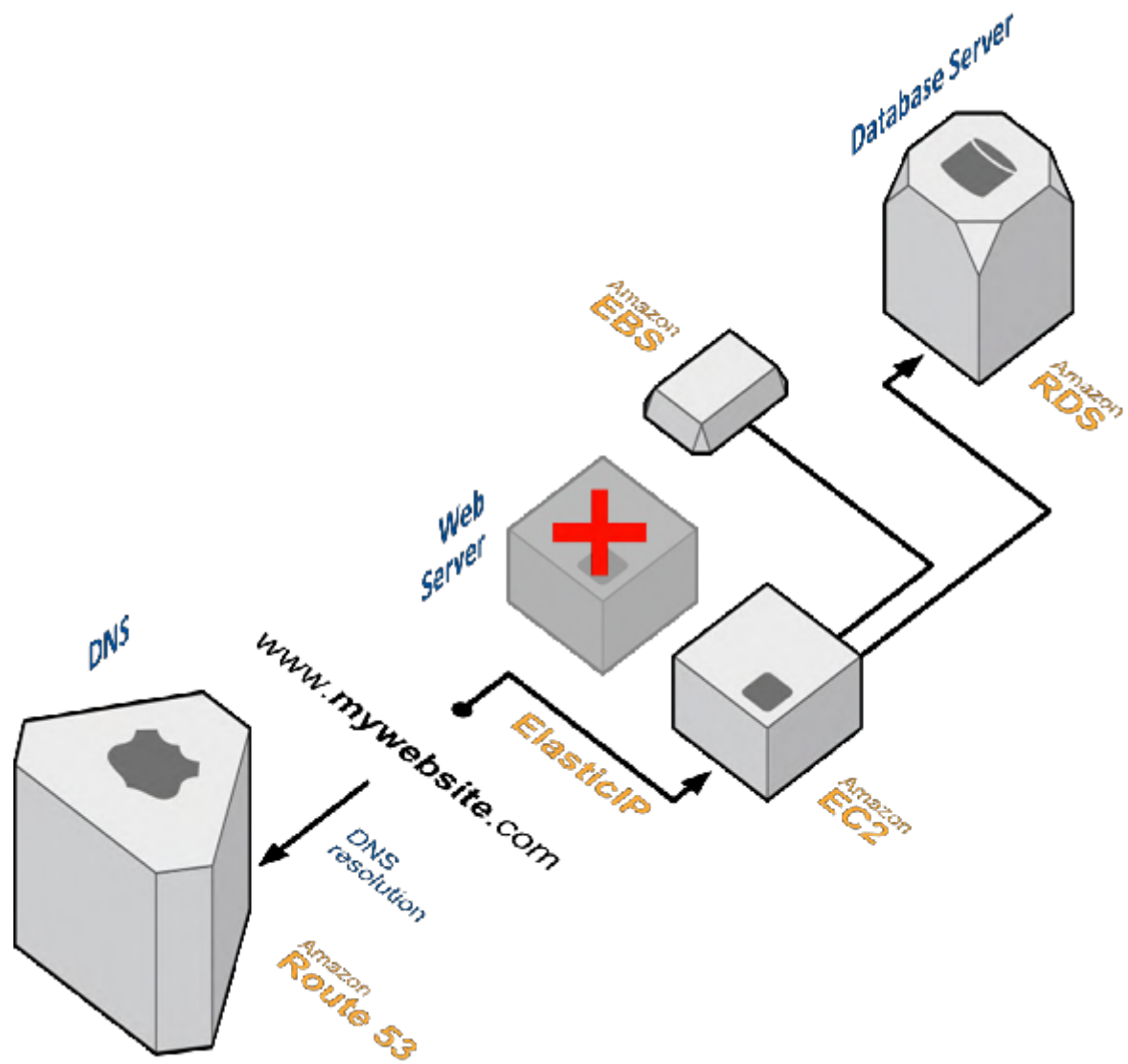


AMAZON EBS

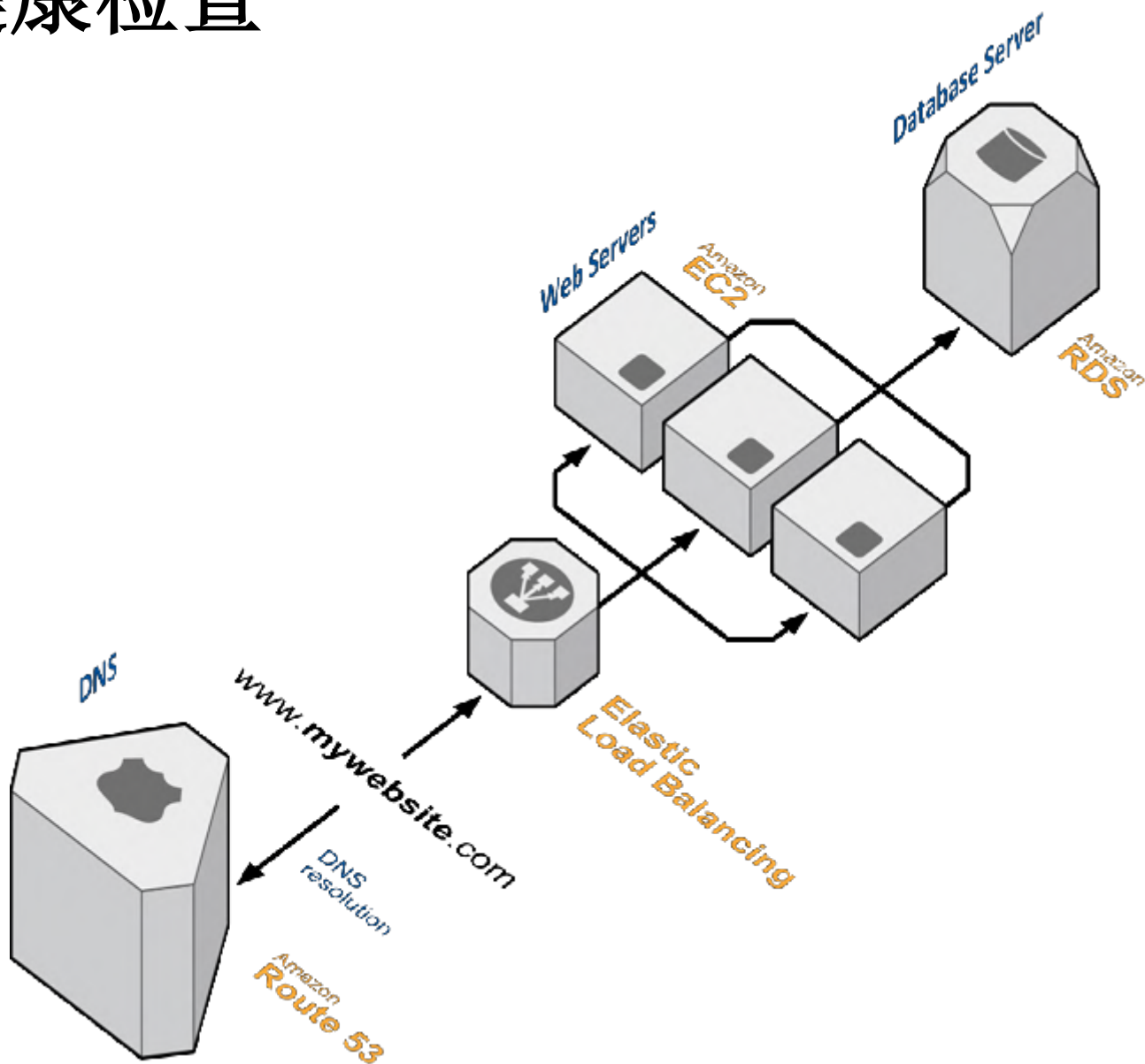








ELB健康检查



Load Balancers

Create Load Balancer

Viewing: All Load Balancers

	Load Balancer Name
<input checked="" type="checkbox"/>	VRLAB-ElasticLoadB...
<input type="checkbox"/>	VRLAB-ElasticLoadB...

1 Load Balancer selected

Load Balancer: VRLAB-ElasticLoadB...

Description Instances

Ping Target:

Timeout:

Interval:

Unhealthy Threshold:

Healthy Threshold:

Edit Health Check

Configure Health Check Cancel

Your load balancer will automatically perform health checks on your EC2 instances and only route traffic to instances that pass the health check. If an instance fails the health check, it is automatically removed from the load balancer.

Ping Protocol: HTTP

Ping Port: 80

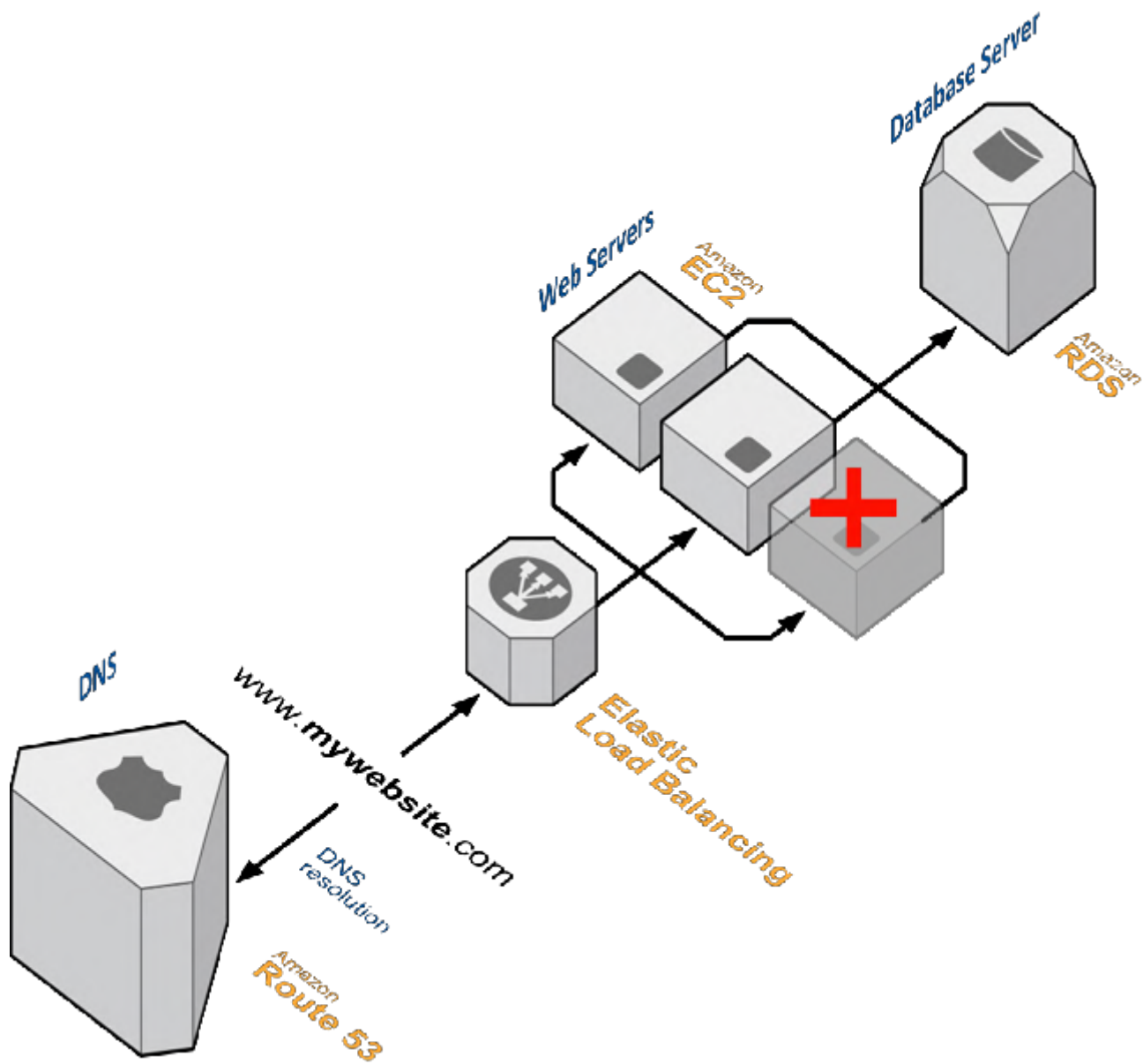
Ping Path: /check.php

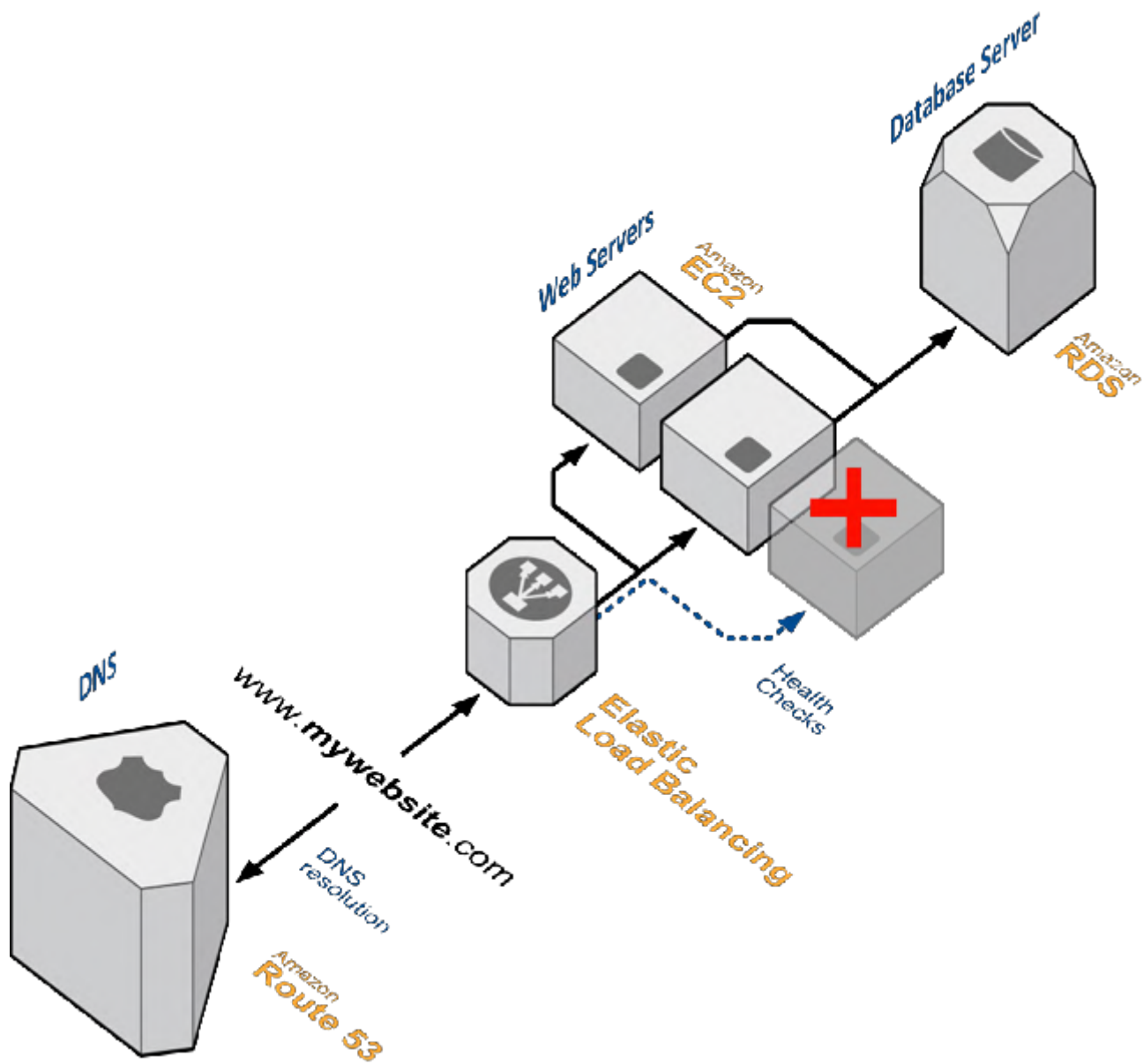
Response Timeout: 5 Seconds

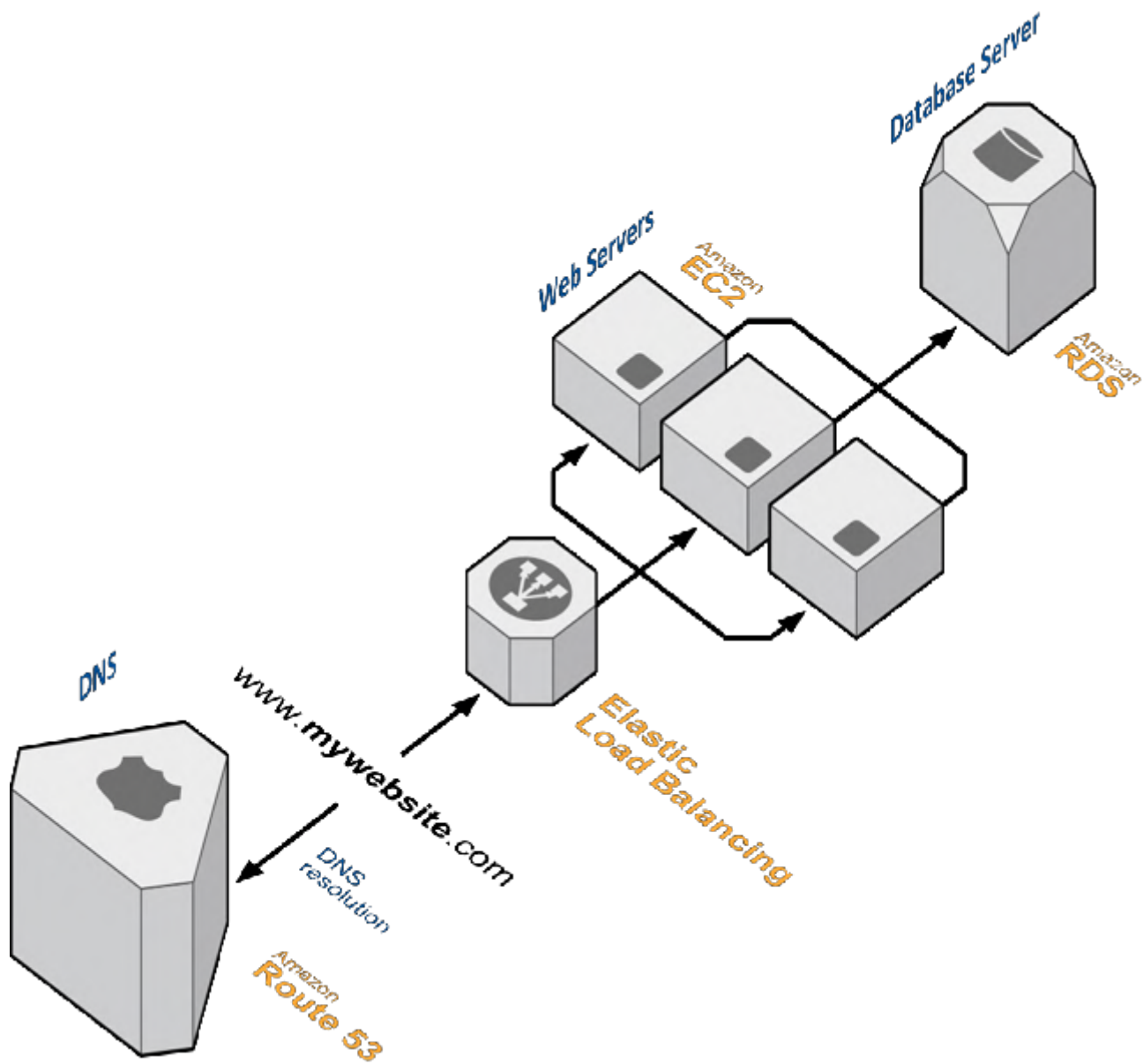
Health Check Interval: 0.5 Minutes

Unhealthy Threshold:

Healthy Threshold:

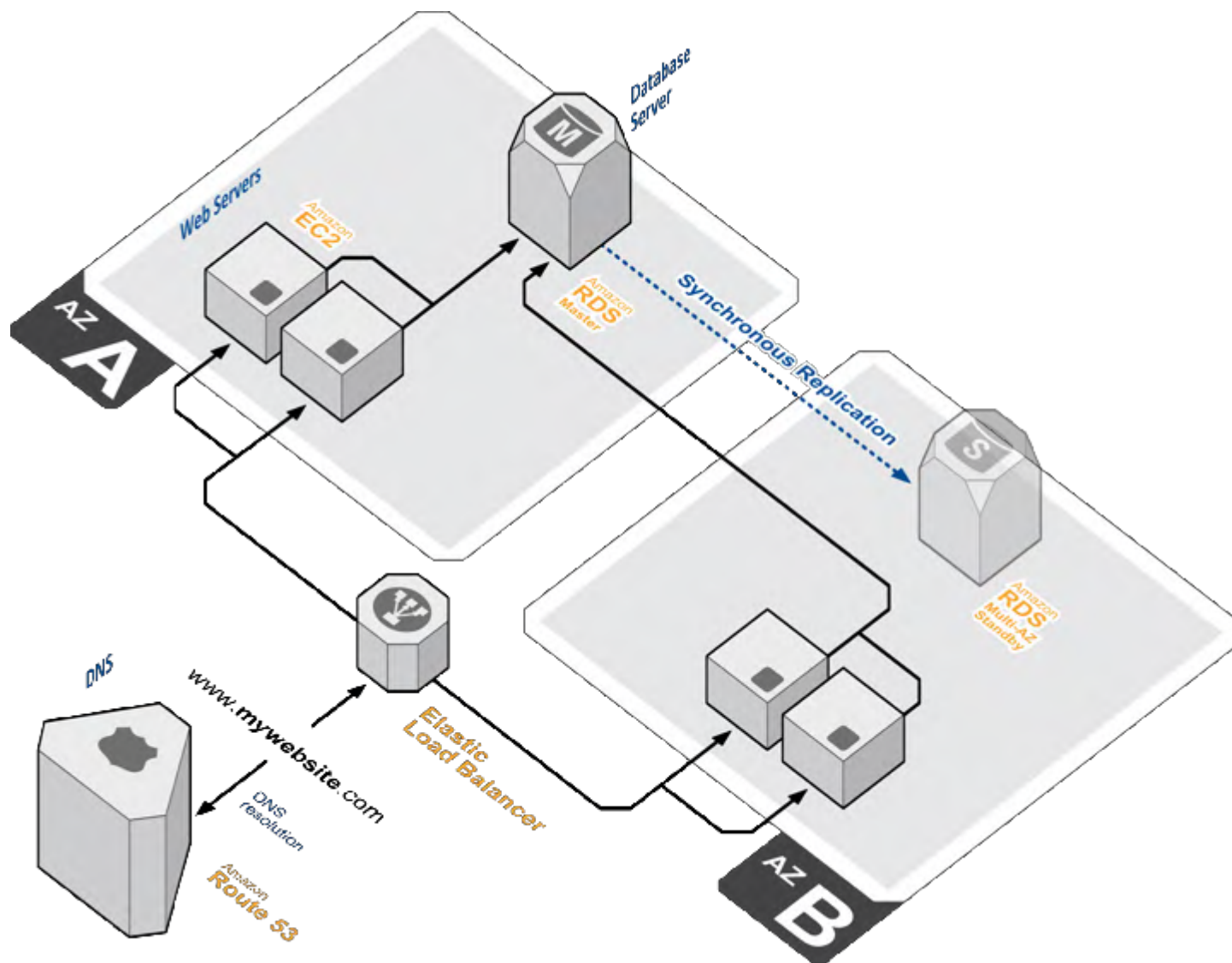




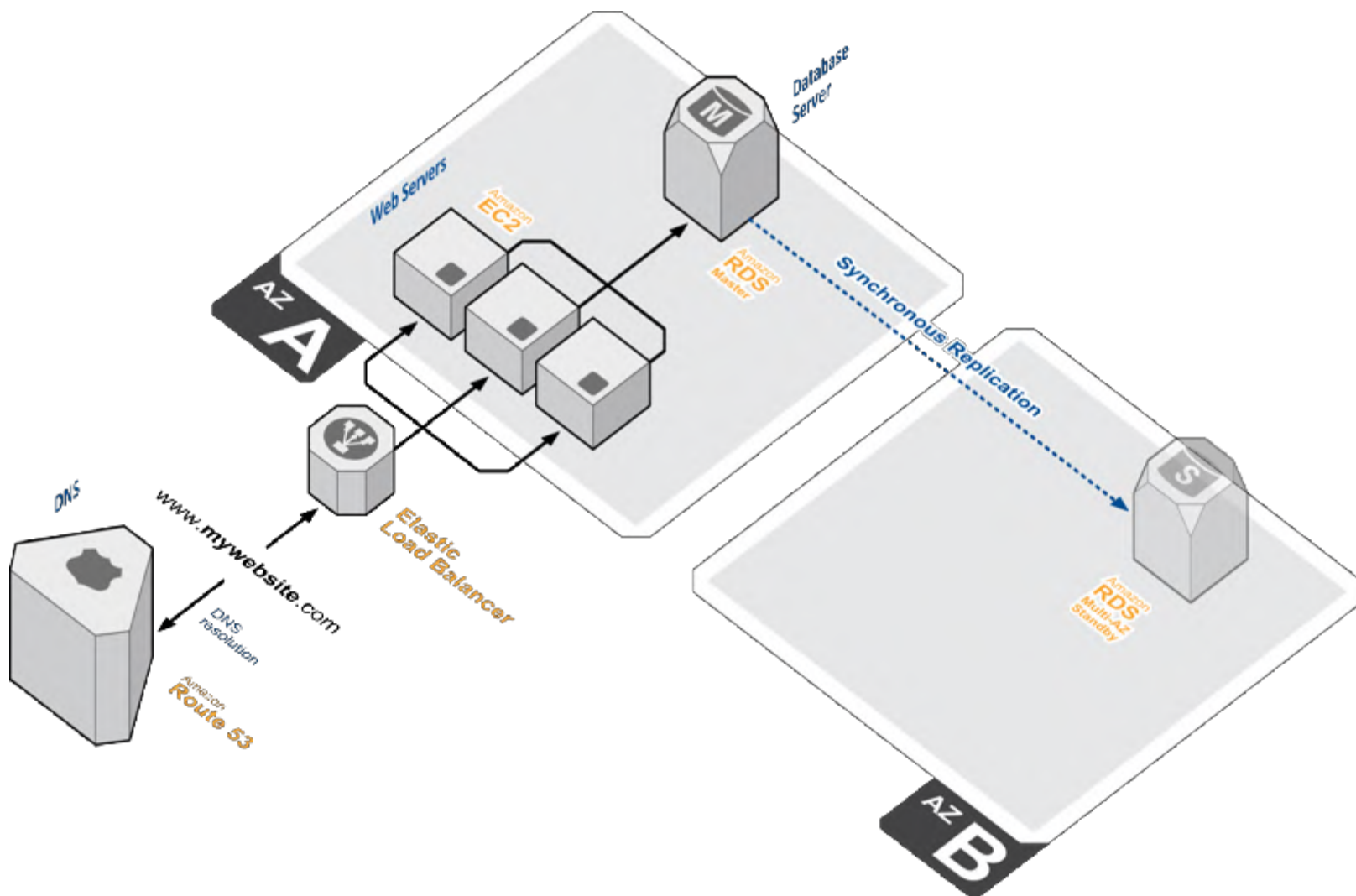


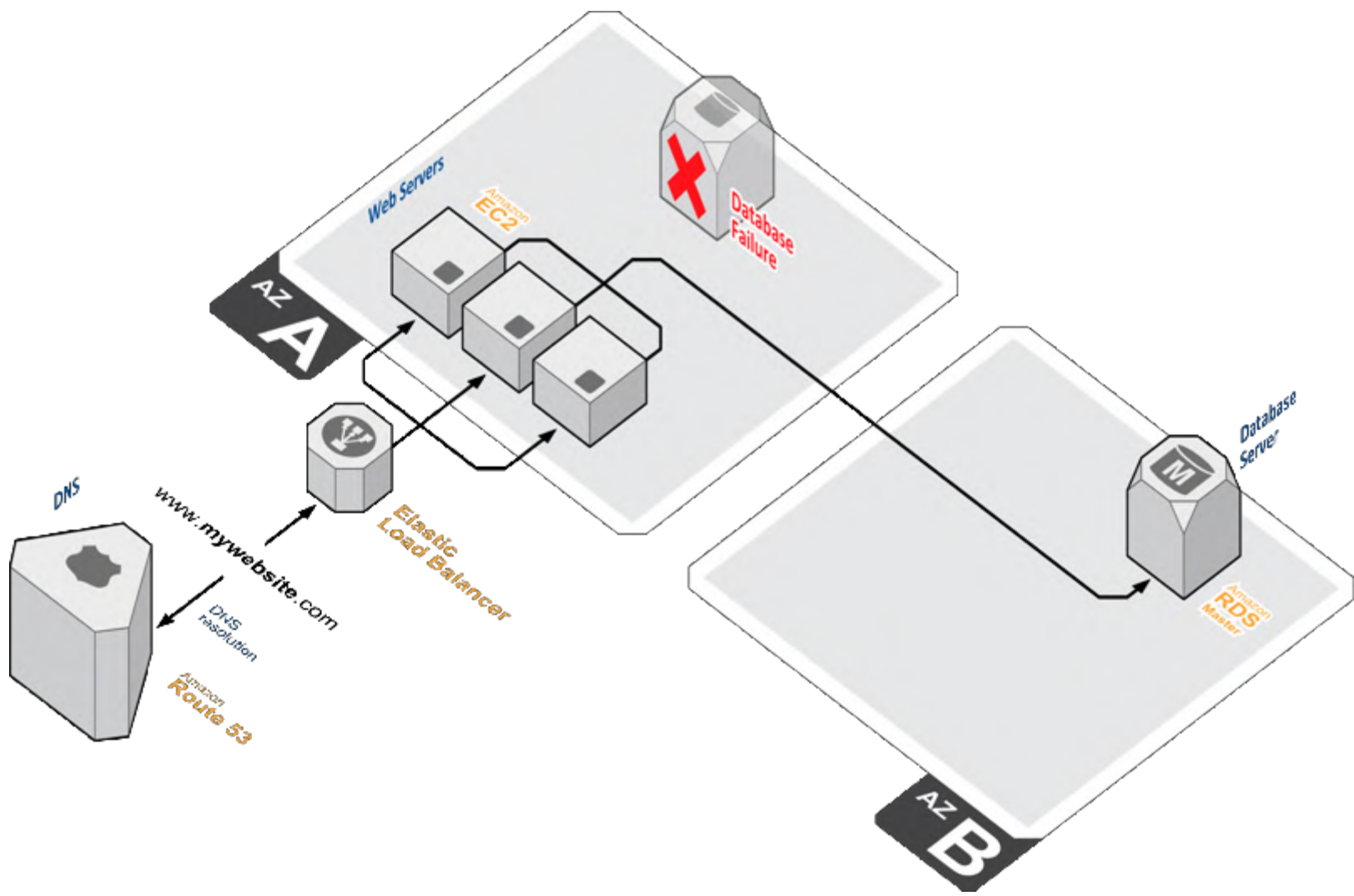
设计原则2：多可用区（AZ）设计

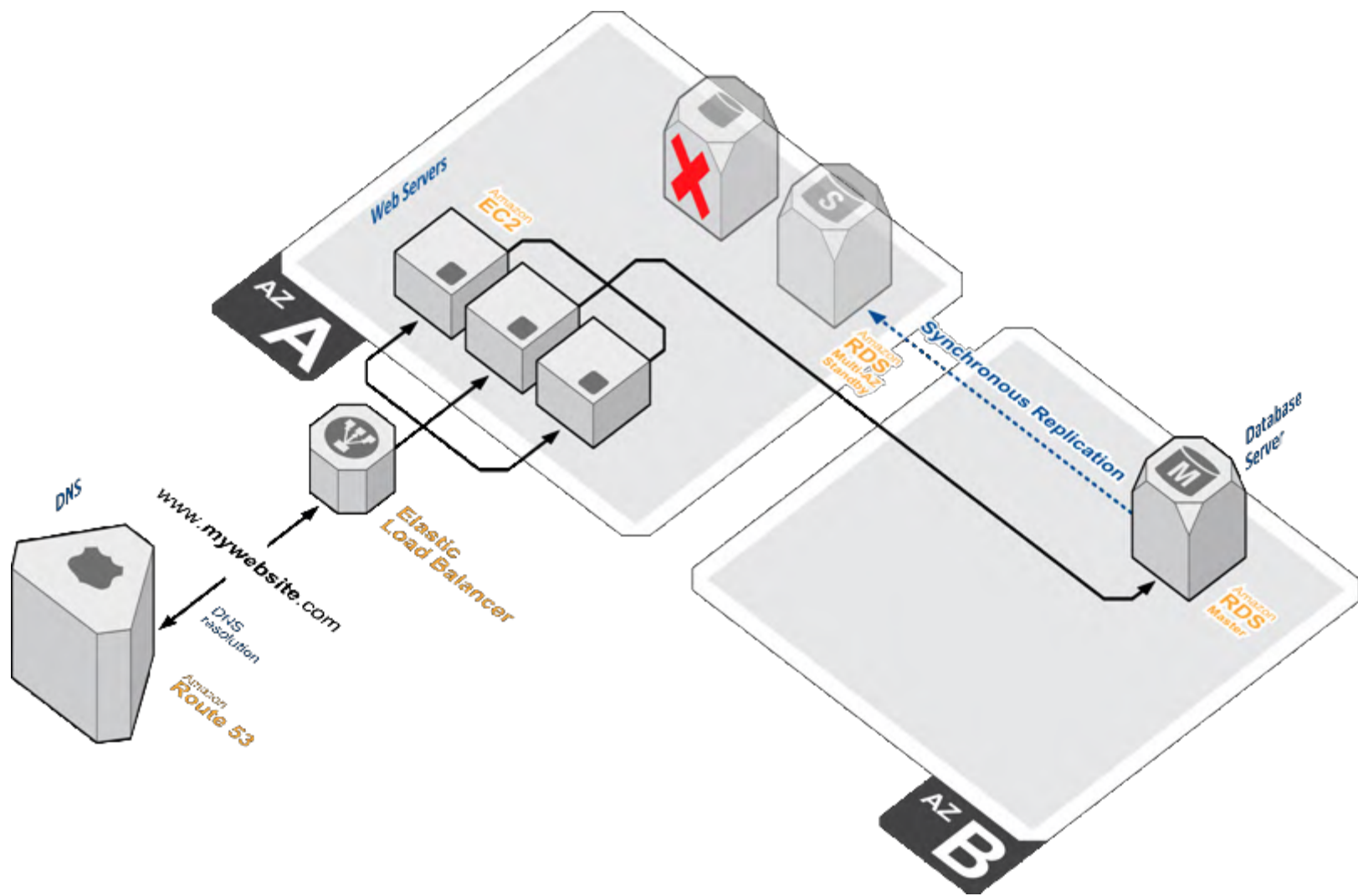
Web应用例子



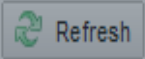
RDS多可用区 (Multi-AZ) 部署







You are using the following Amazon RDS region



Launch DB Instance Wizard

Cancel X

ENGINE SELECTION DETAILS ADDITIONAL CONFIGURATION MANAGEMENT OPTIONS REVIEW

To get started, choose an engine below and click **Continue**



License: General Public License

DB Engine: MySQL 5.5.27 (default)

DB Instance Class: db.m1.xlarge

Multi-AZ Deployment: - Select One -

Auto Minor Version Upgrade: - Select One -
No
Yes

Provide the details for your RDS Database Instance

Allocated Storage: 500 GB (Minimum: 5 GB, Maximum: 1024 GB) Higher allocated storage may improve IOPS performance.

Use Provisioned IOPS:

DB Instance Identifier: (e.g. mydbinstance)

Master Username: (e.g. awsuser)

Master Password: (e.g. mypassword)

< Back

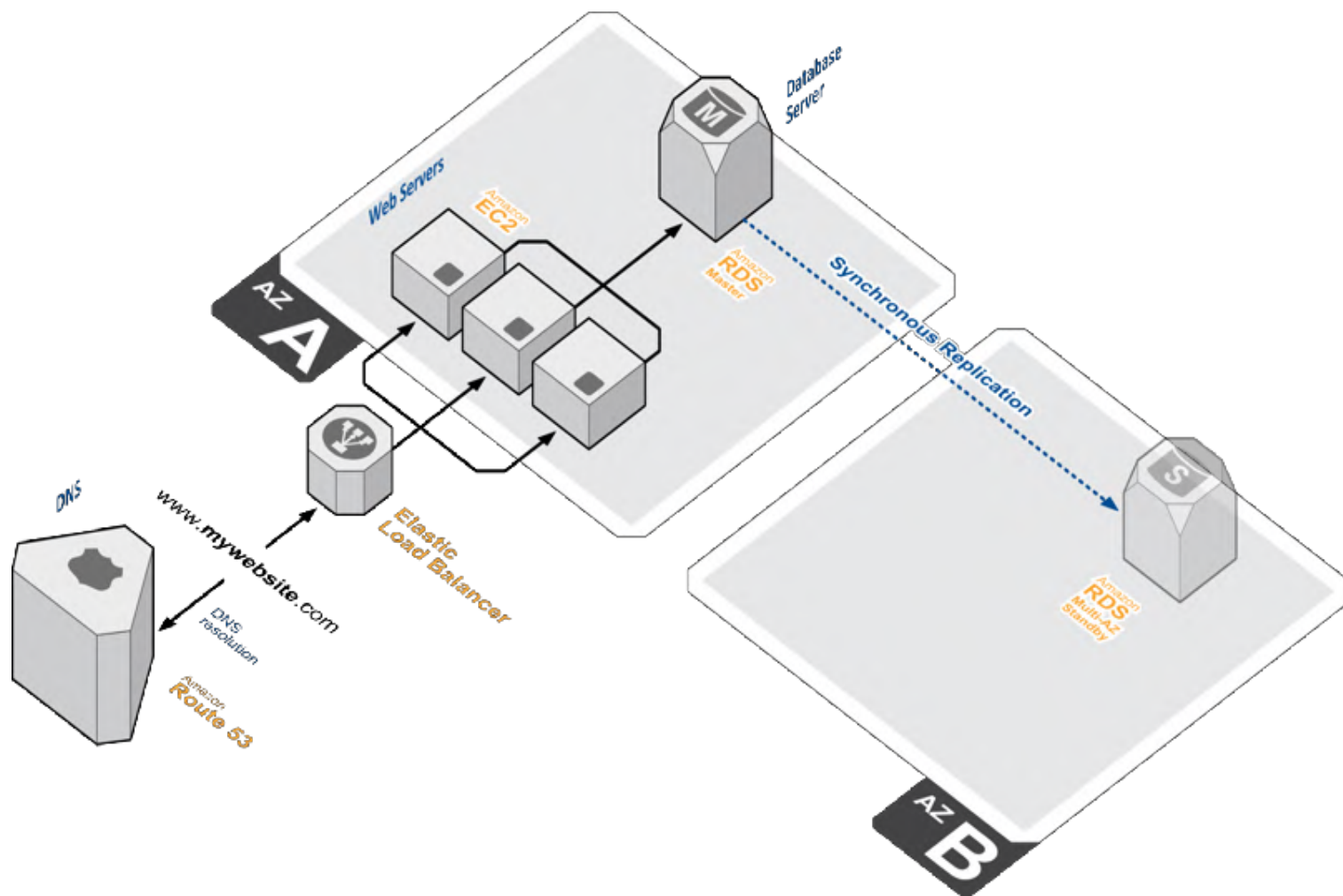
Continue >

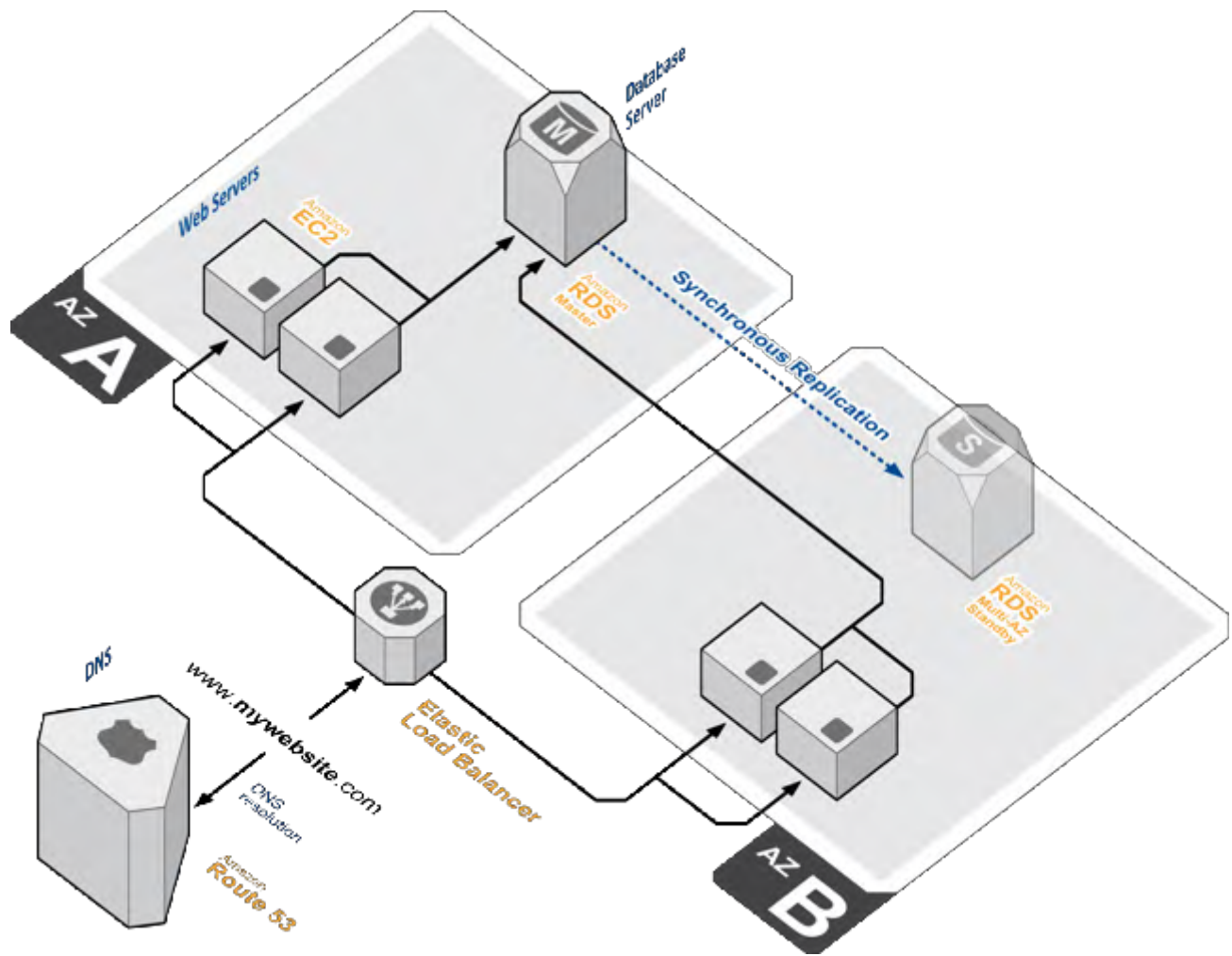
Current Status

Details

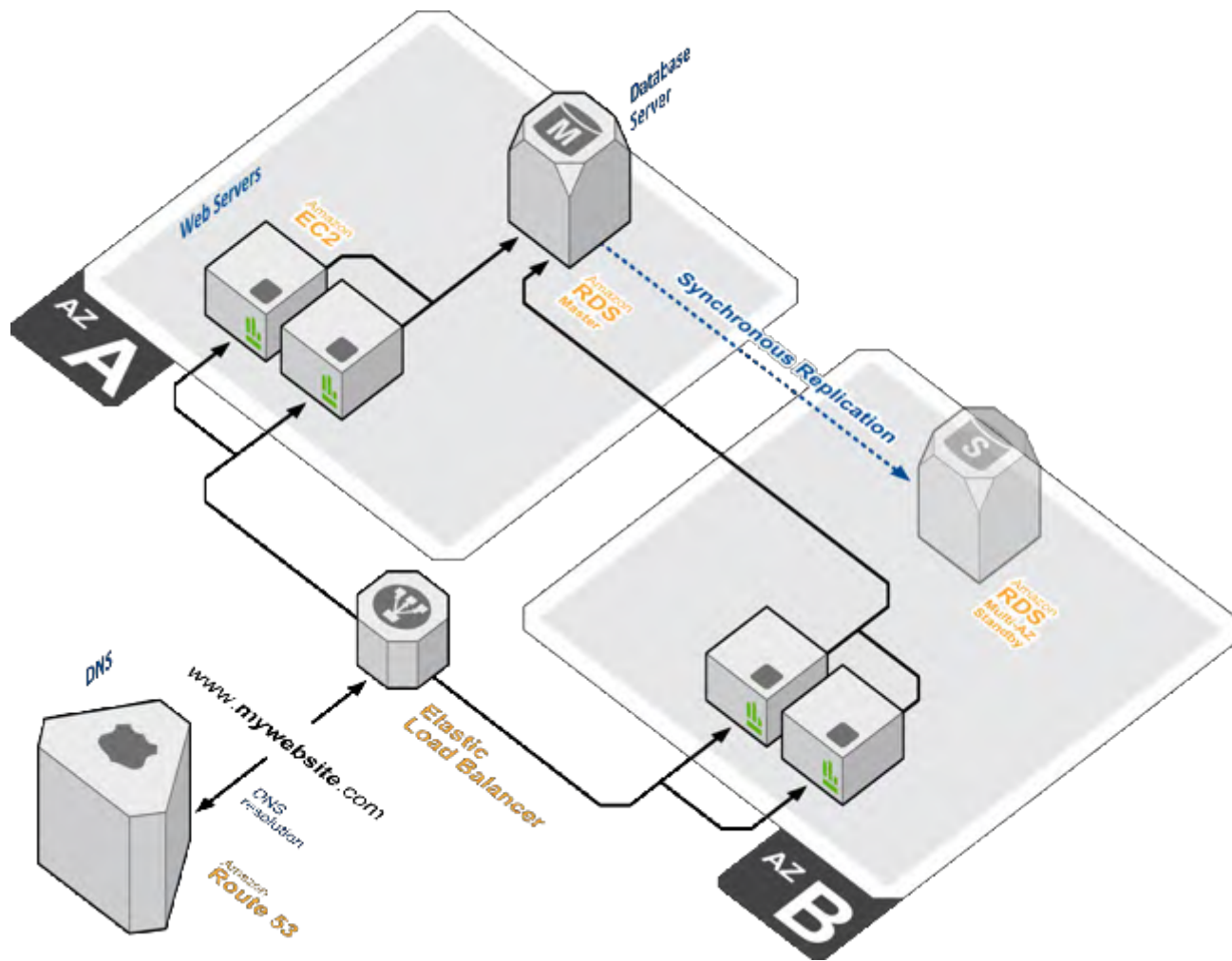
	Amazon Relational Database Service (Ireland)	Service is operating normally
--	--	-------------------------------

Web层的多可用区部署

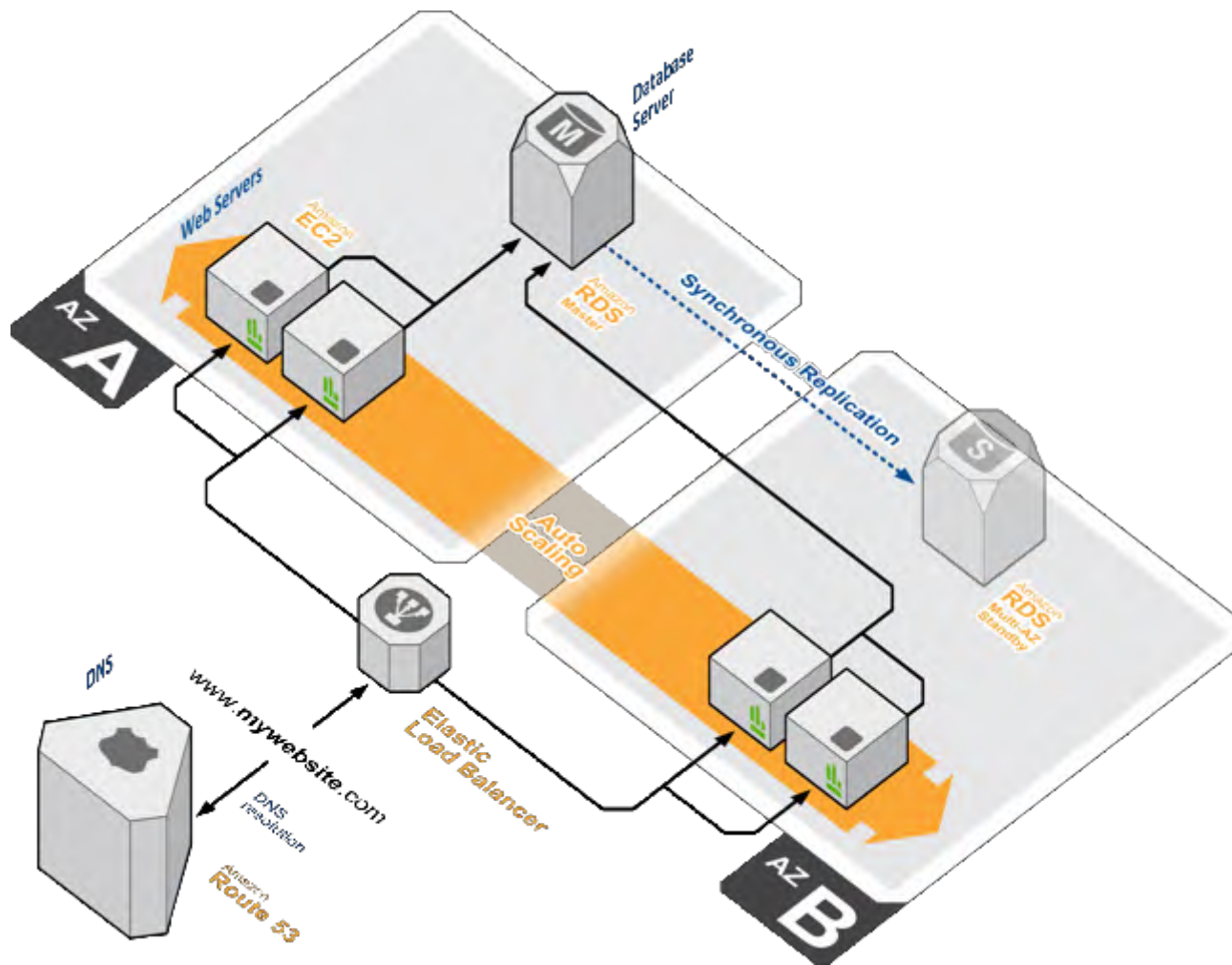


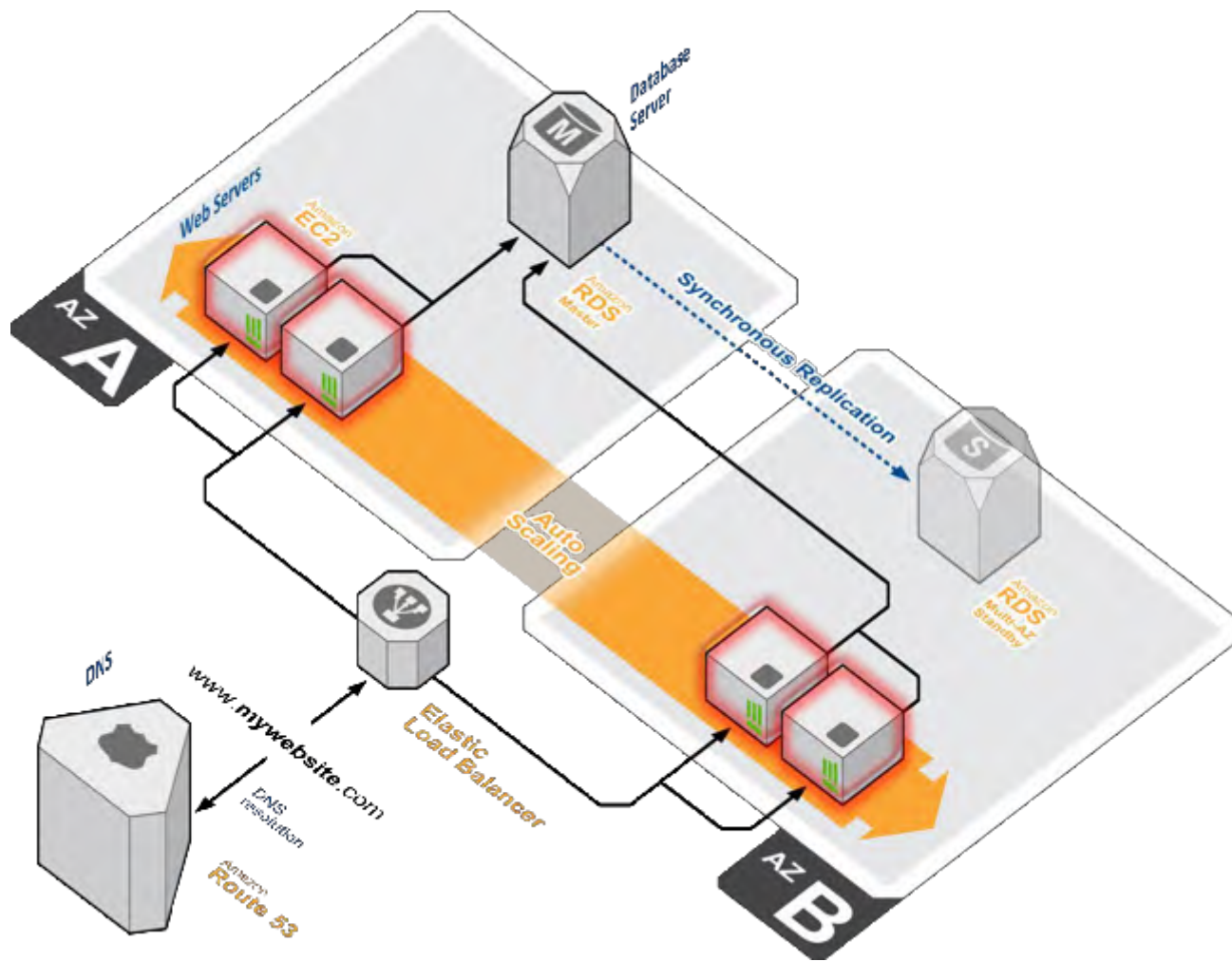


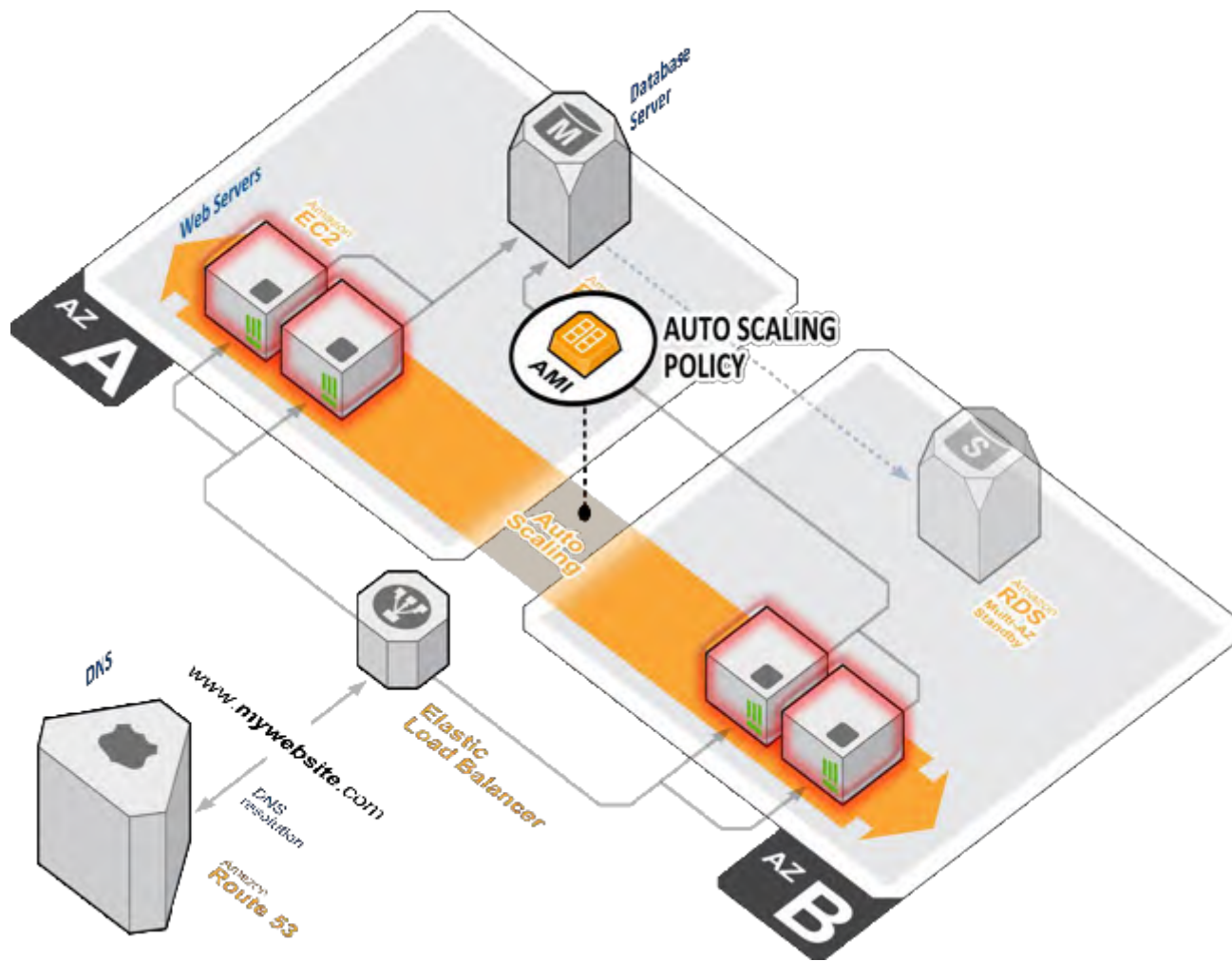
设计原则3：自动扩展设计

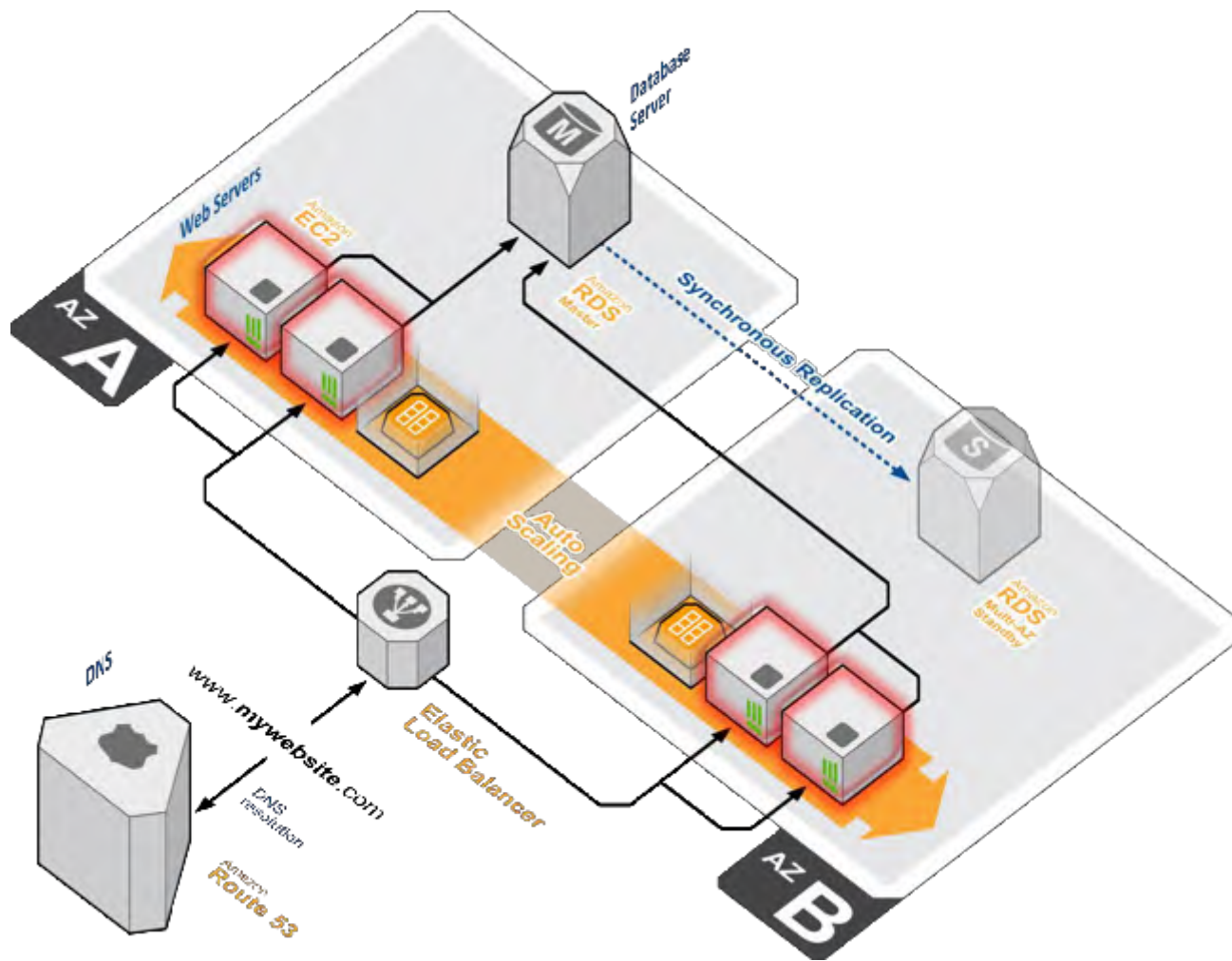


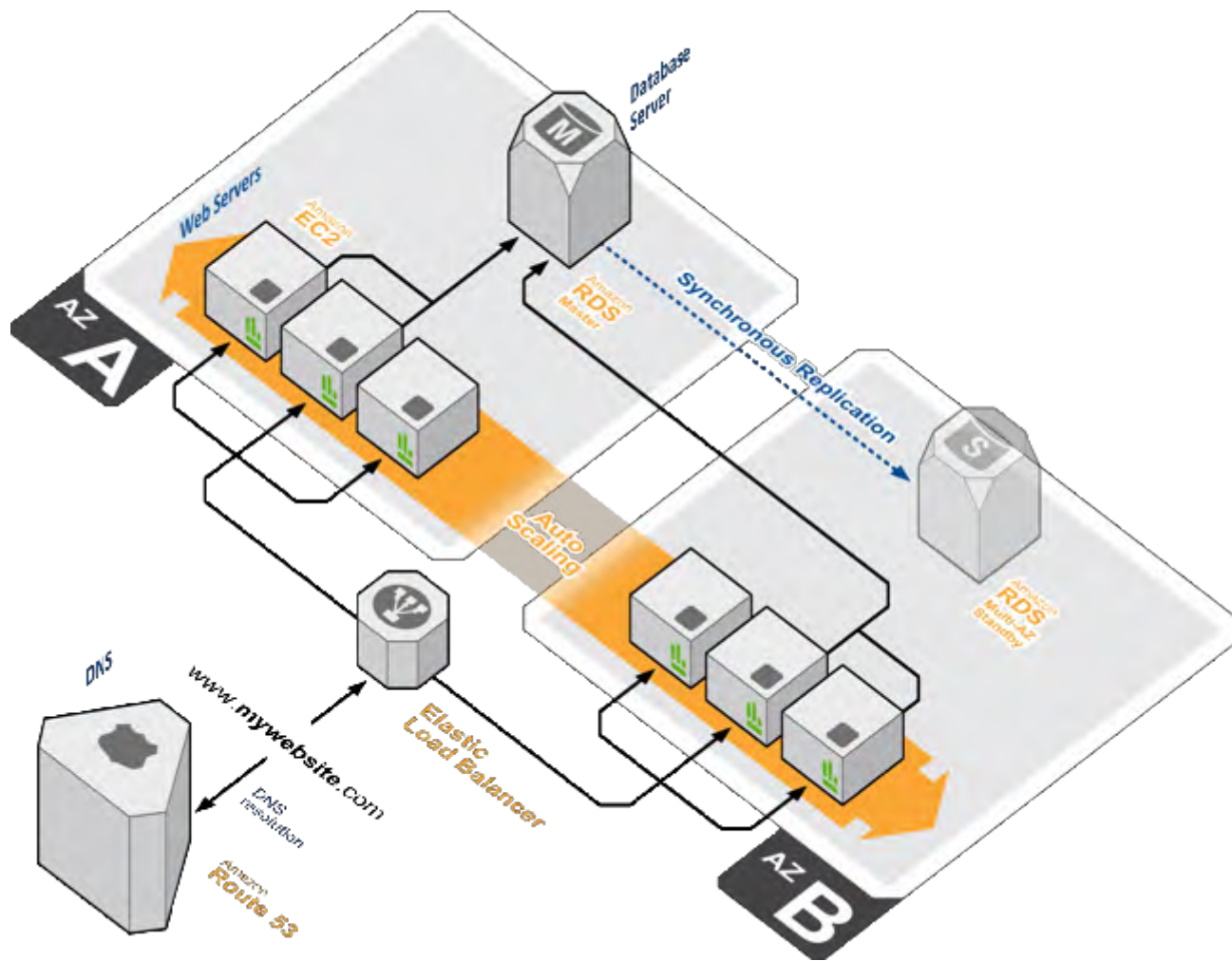
设置Auto Scaling组

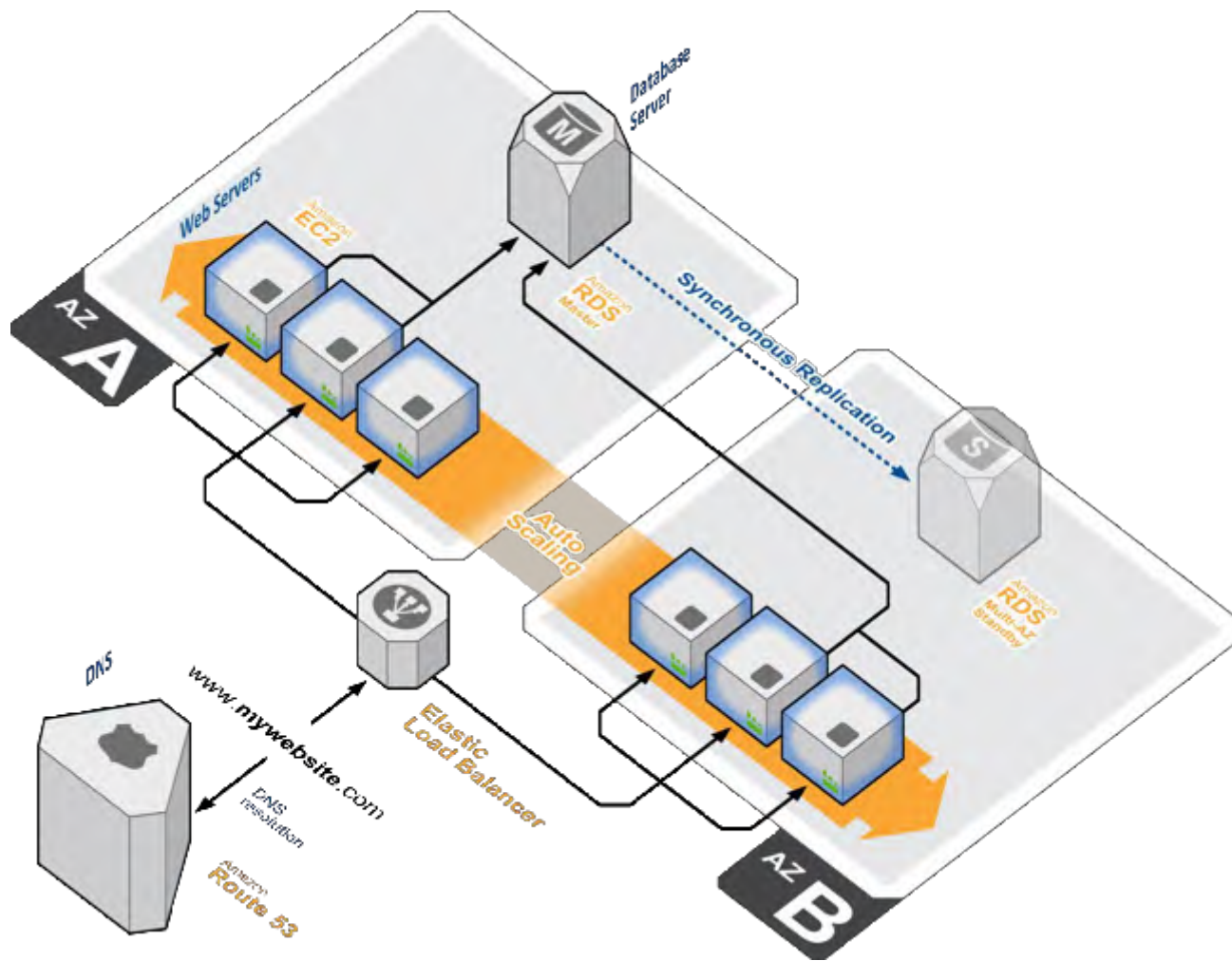


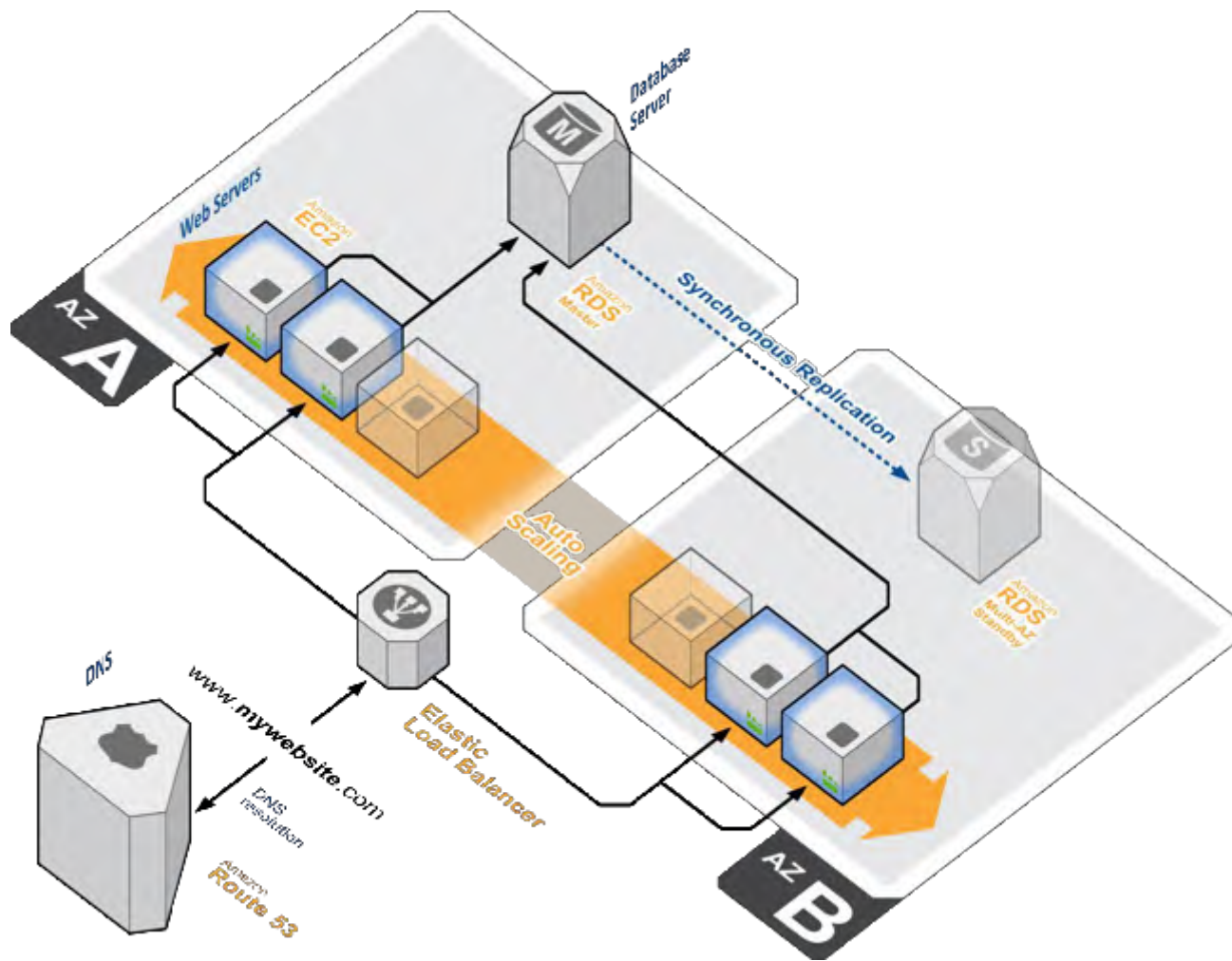


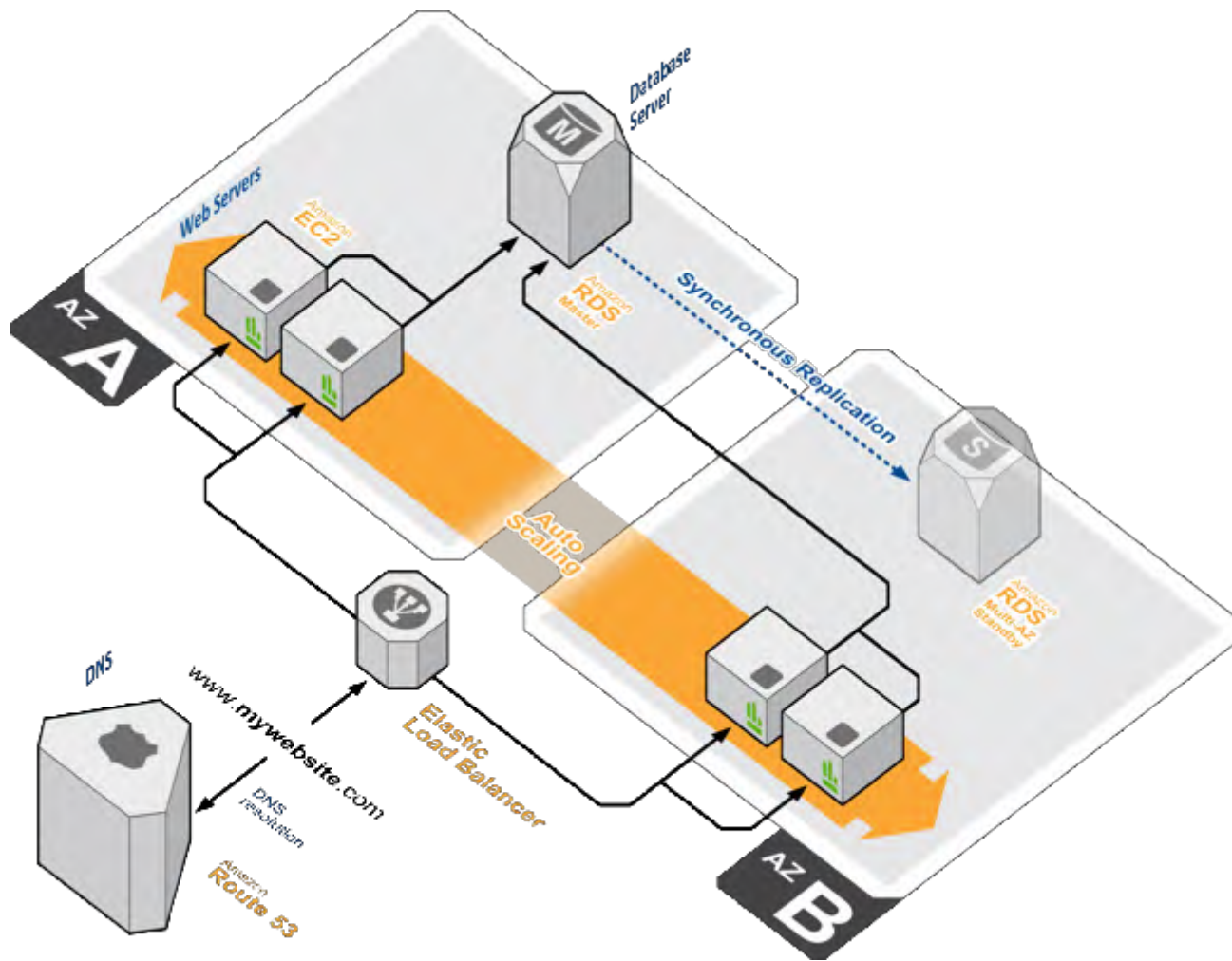






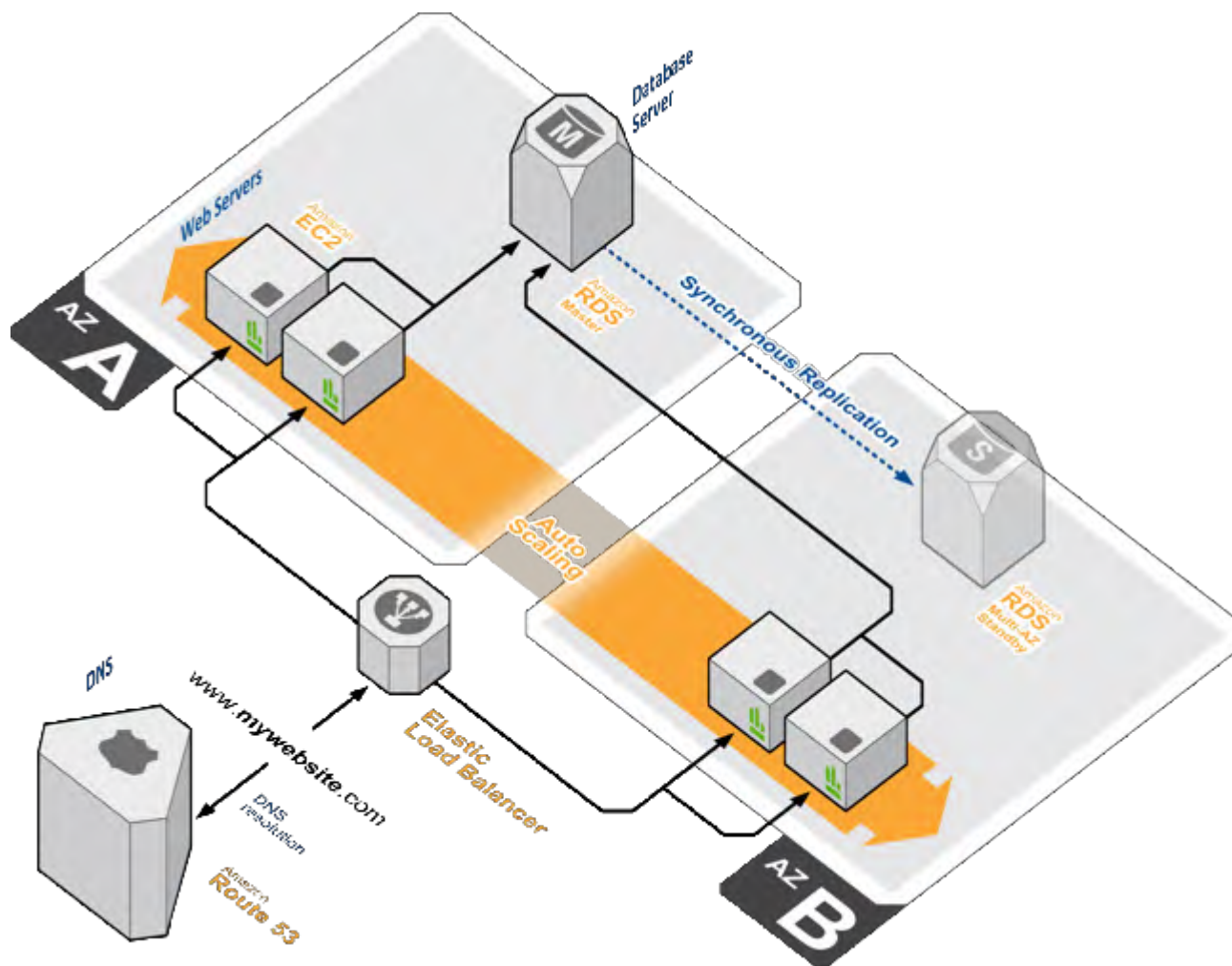


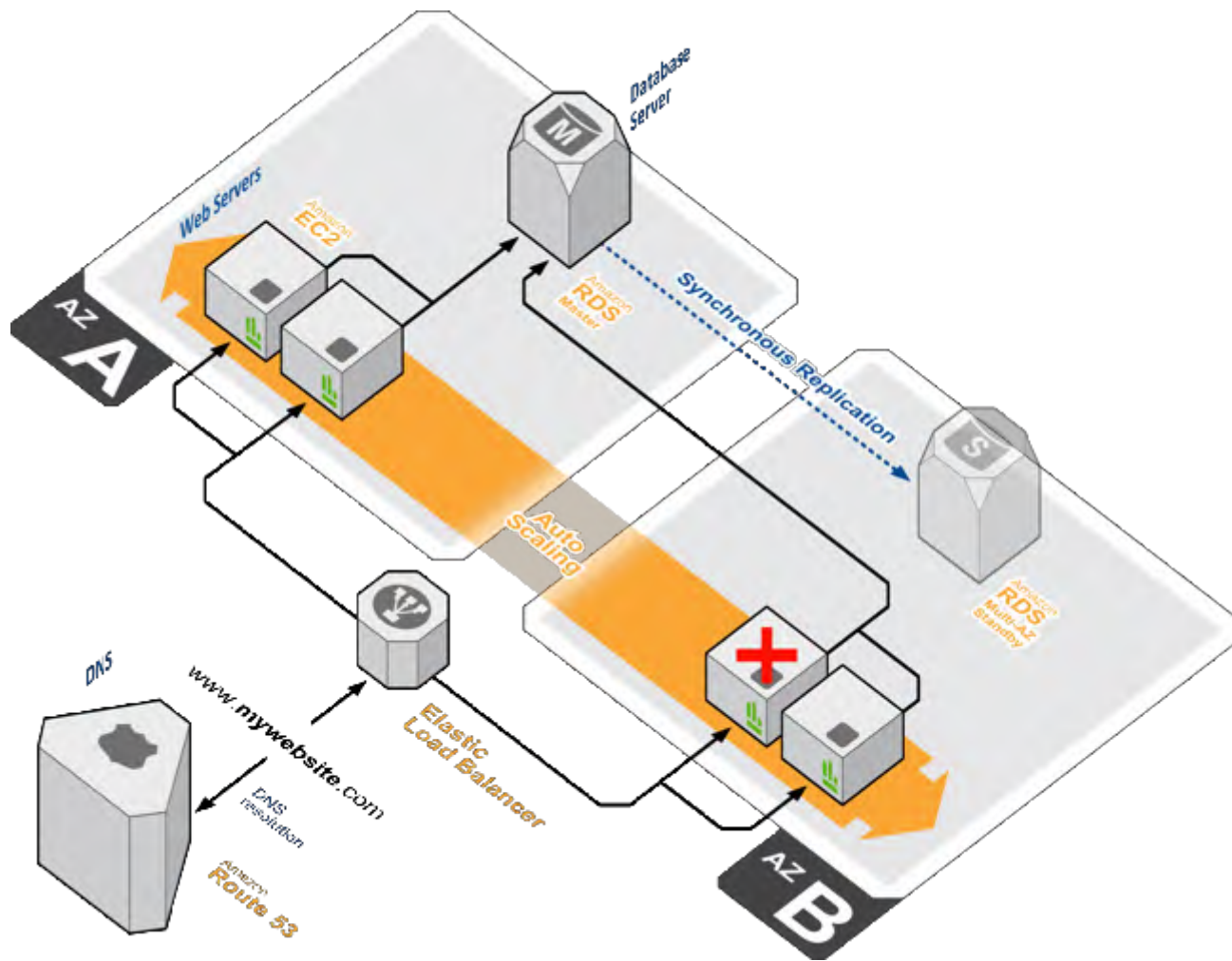


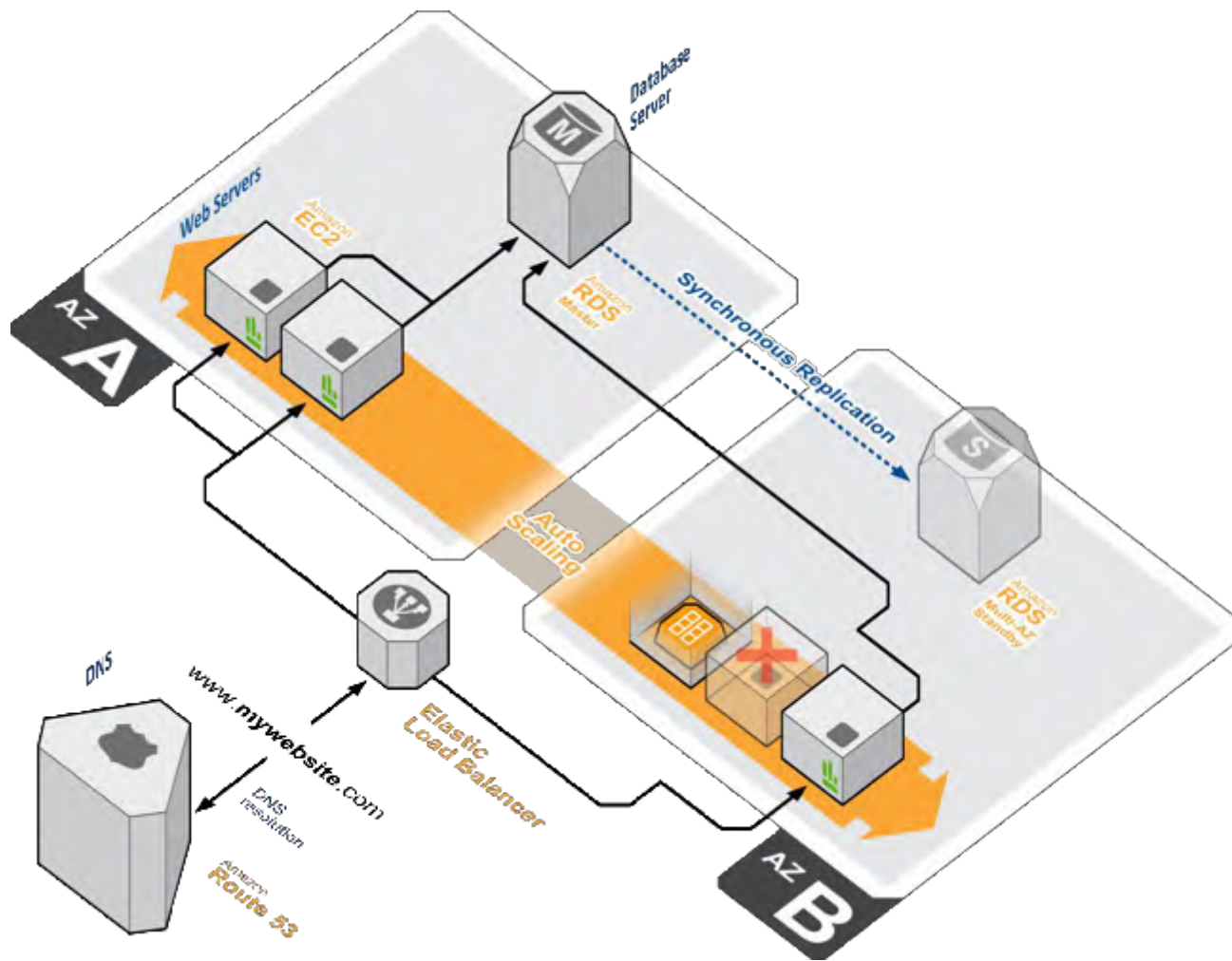


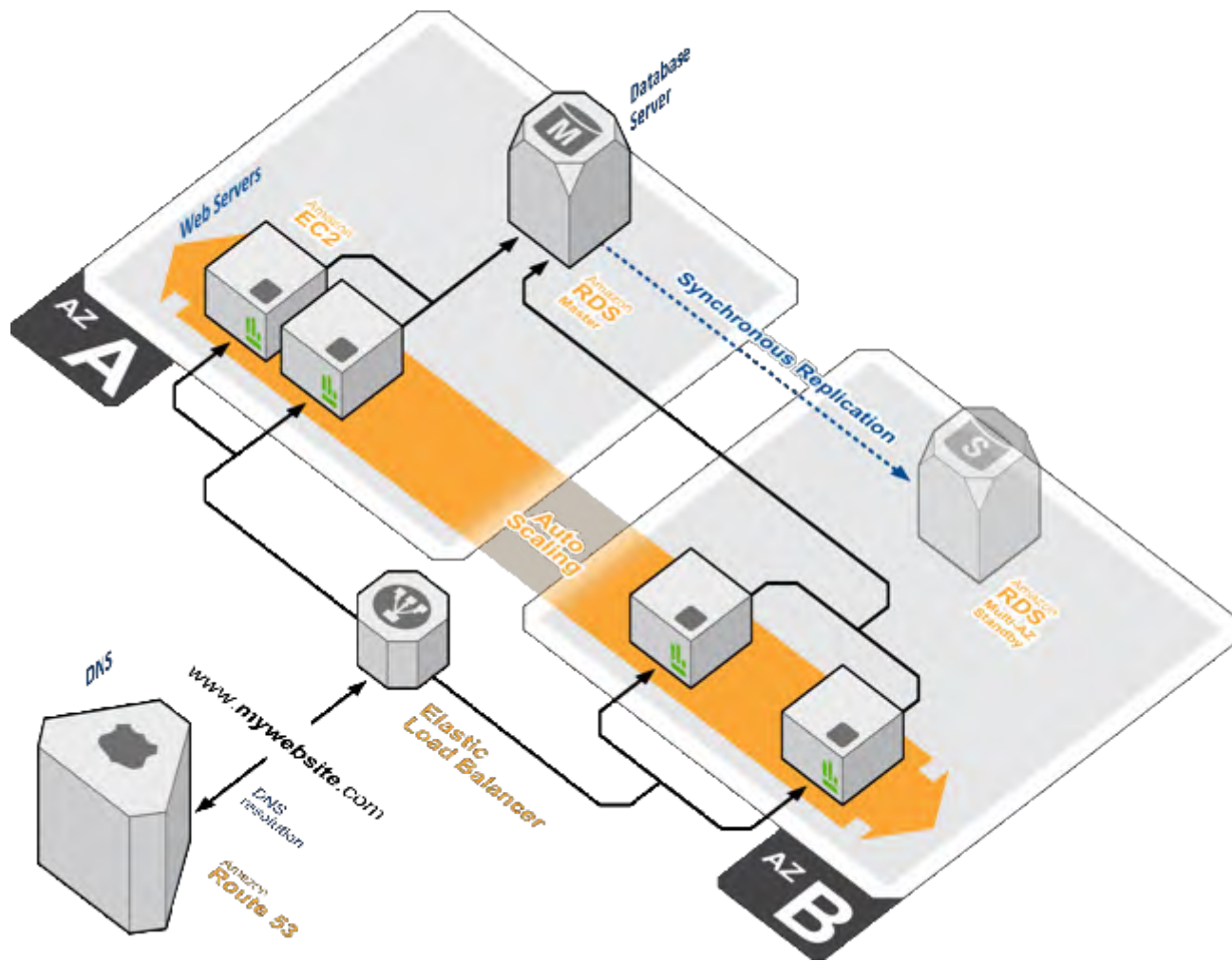
设计原则4：自我修复设计

通过Auto Scaling + ELB健康检查









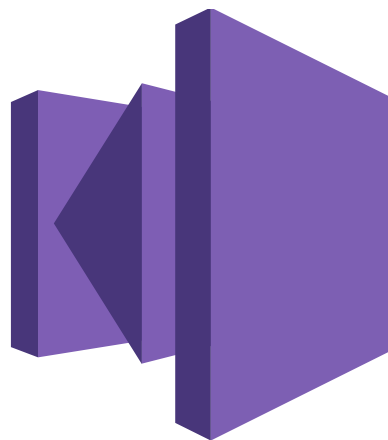
设计原则5：松耦合设计

关于耦合度

耦合度与灵活性相反

耦合度越小，扩展性越好，容错能力越大

消息队列服务



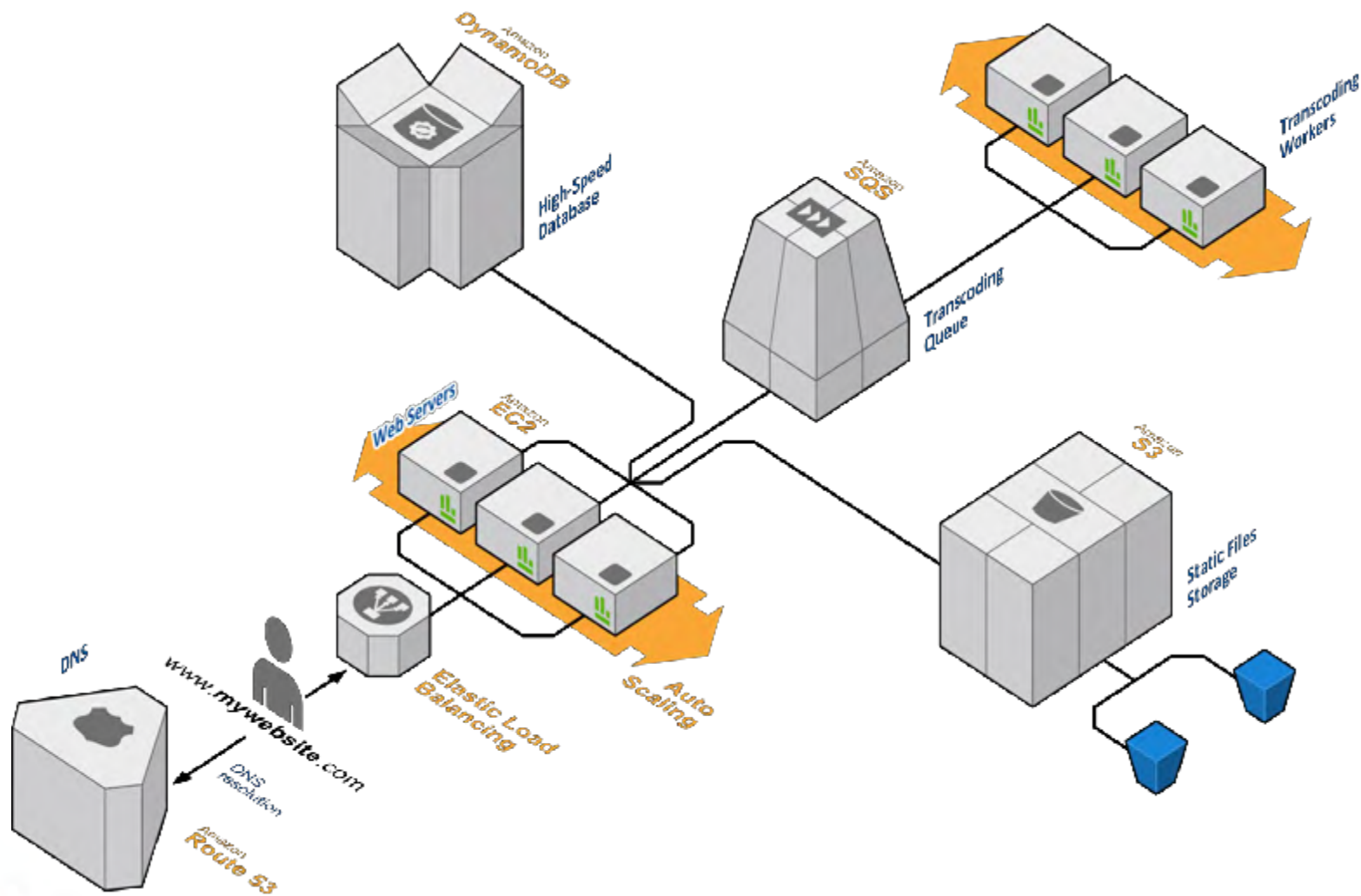
AMAZON SQS
SIMPLE QUEUE SERVICE

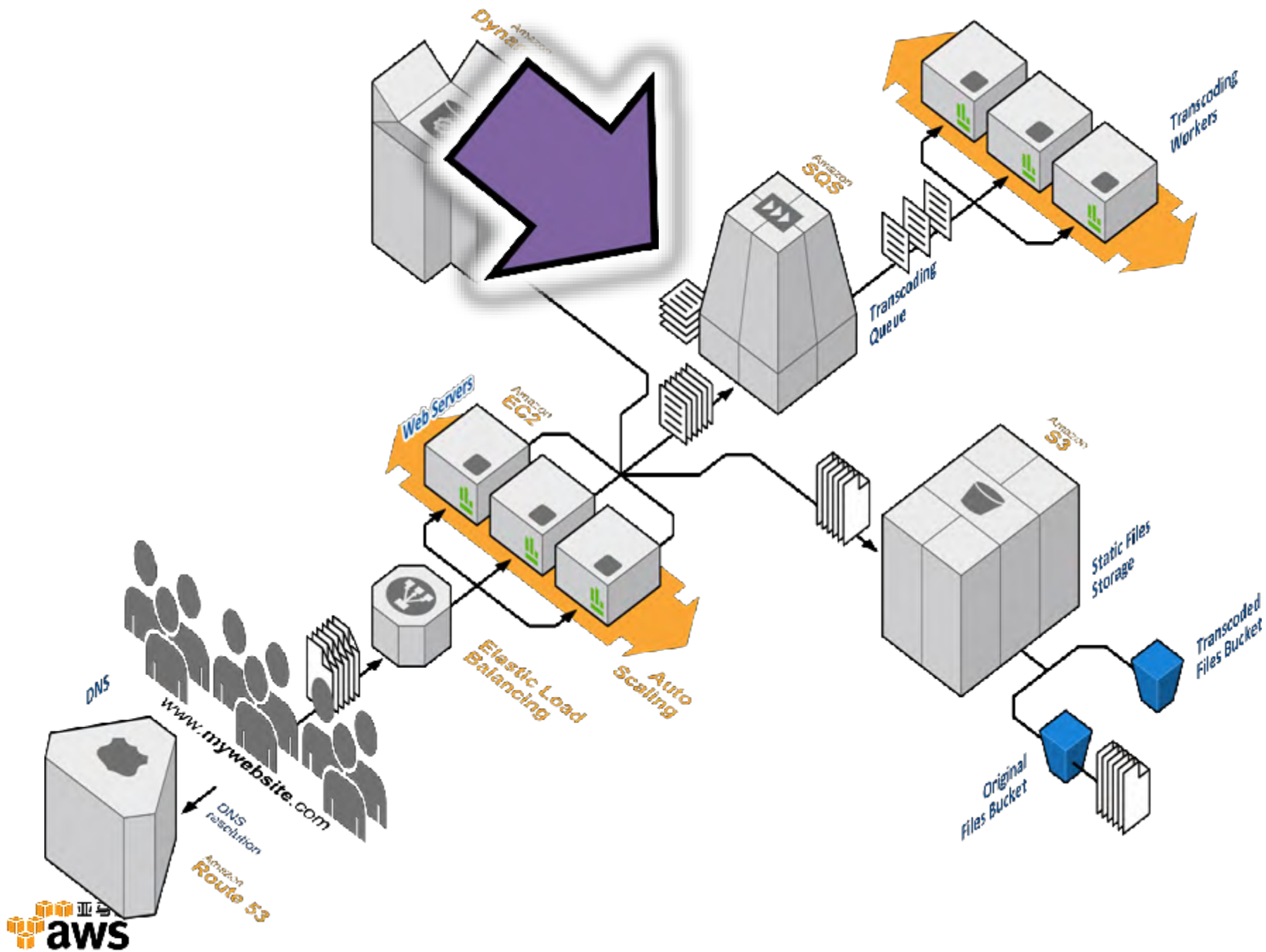
媒体数据处理应用场景

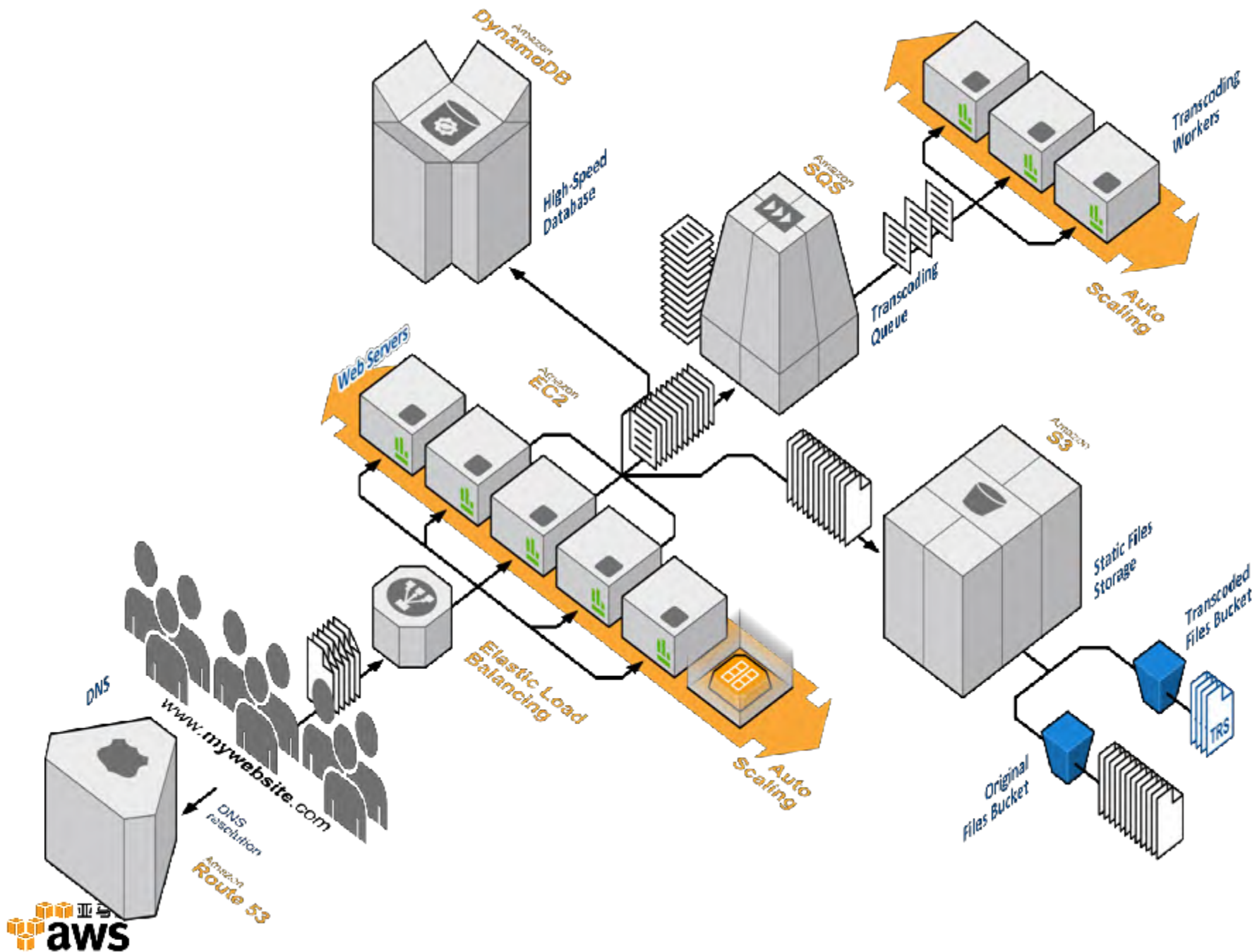


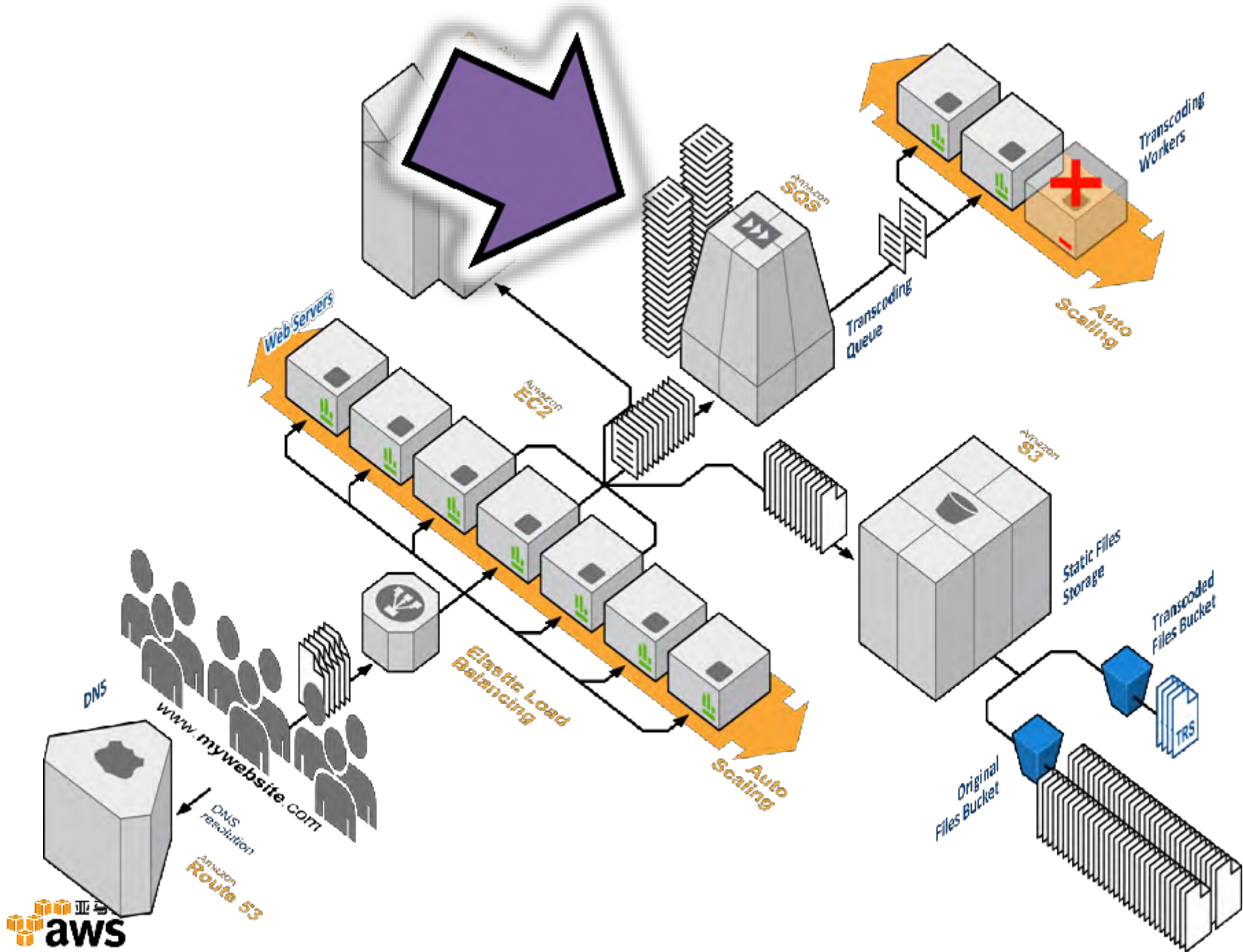


基于SQS的设计











Navigation

- > Dashboard
- > Alarms
 - > All states
 - > ALARM
 - > INSUFFICIENT DATA
 - > OK
 - > Billing Alarms
- > Metrics
 - > All metrics
 - > Billing
 - > DynamoDB
 - > EBS
 - > EC2
 - > ELB
 - > ElastiCache
 - > ElasticMapReduce
 - > RDS
 - > SNS
 - > SQS
 - > StorageGateway

Metrics

Viewing: All Metrics ▾ AWS/SQS Search 1 to 8 of 8 Metrics Help

SQS: Queue Metrics

	QueueName	MetricName
<input type="checkbox"/>	riab-transcoding	ApproximateNumberOfMessagesDelayed
<input type="checkbox"/>	riab-transcoding	ApproximateNumberOfMessagesNotVisible
<input checked="" type="checkbox"/>	riab-transcoding	ApproximateNumberOfMessagesVisible
<input type="checkbox"/>	riab-transcoding	NumberOfEmptyReceives
<input type="checkbox"/>	riab-transcoding	NumberOfMessagesDeleted



ApproximateNumberOfMessagesVisible (Cour... edit

Statistic: Maximum ▾

Period: 5 Minutes ▾

Time Range

Zoom: 1h | 3h | 6h | 12h | 24h | 1W | 2W

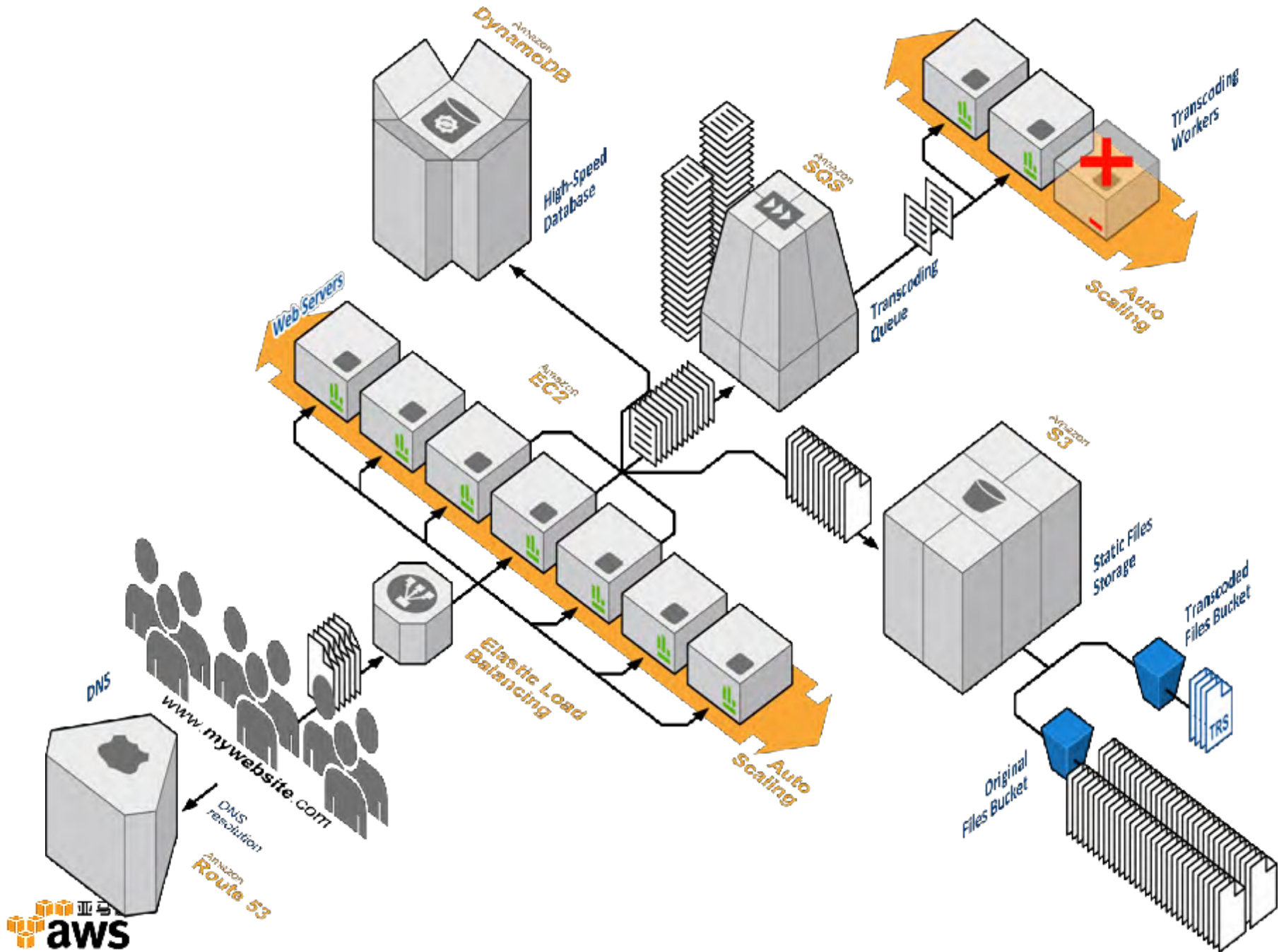
Relative Absolute

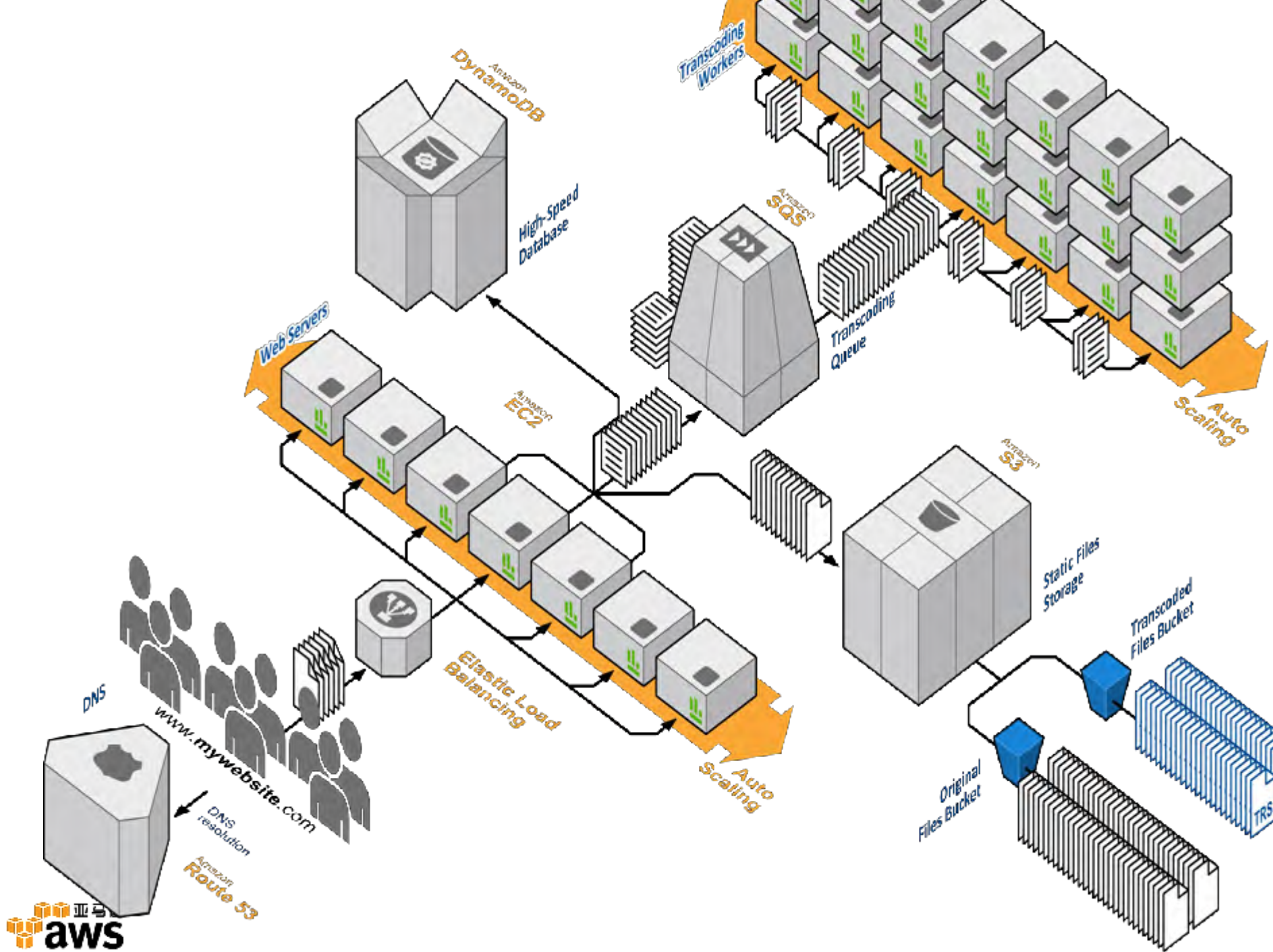
From: 6 hours ago ▾

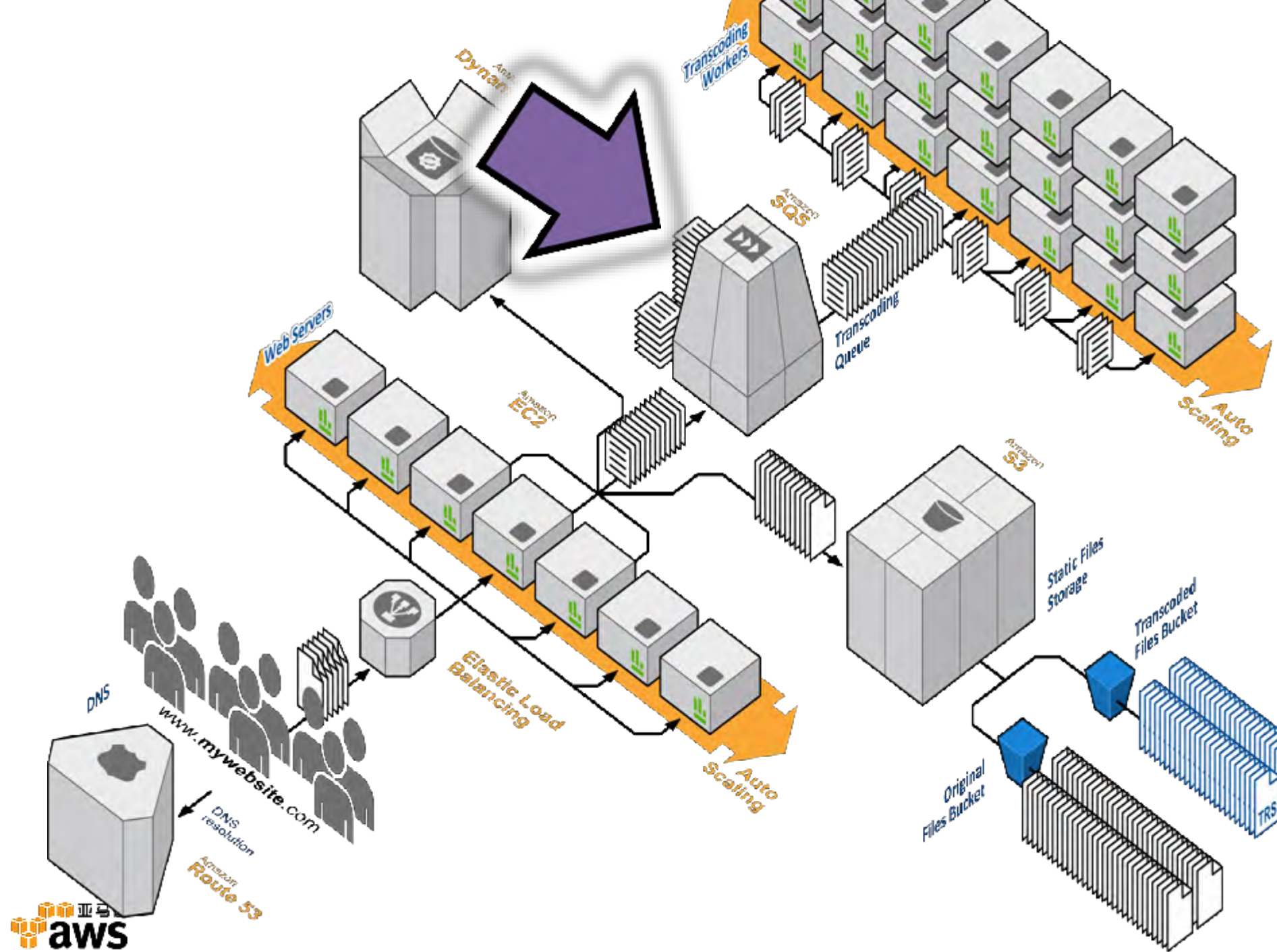
To: 0 hours ago ▾

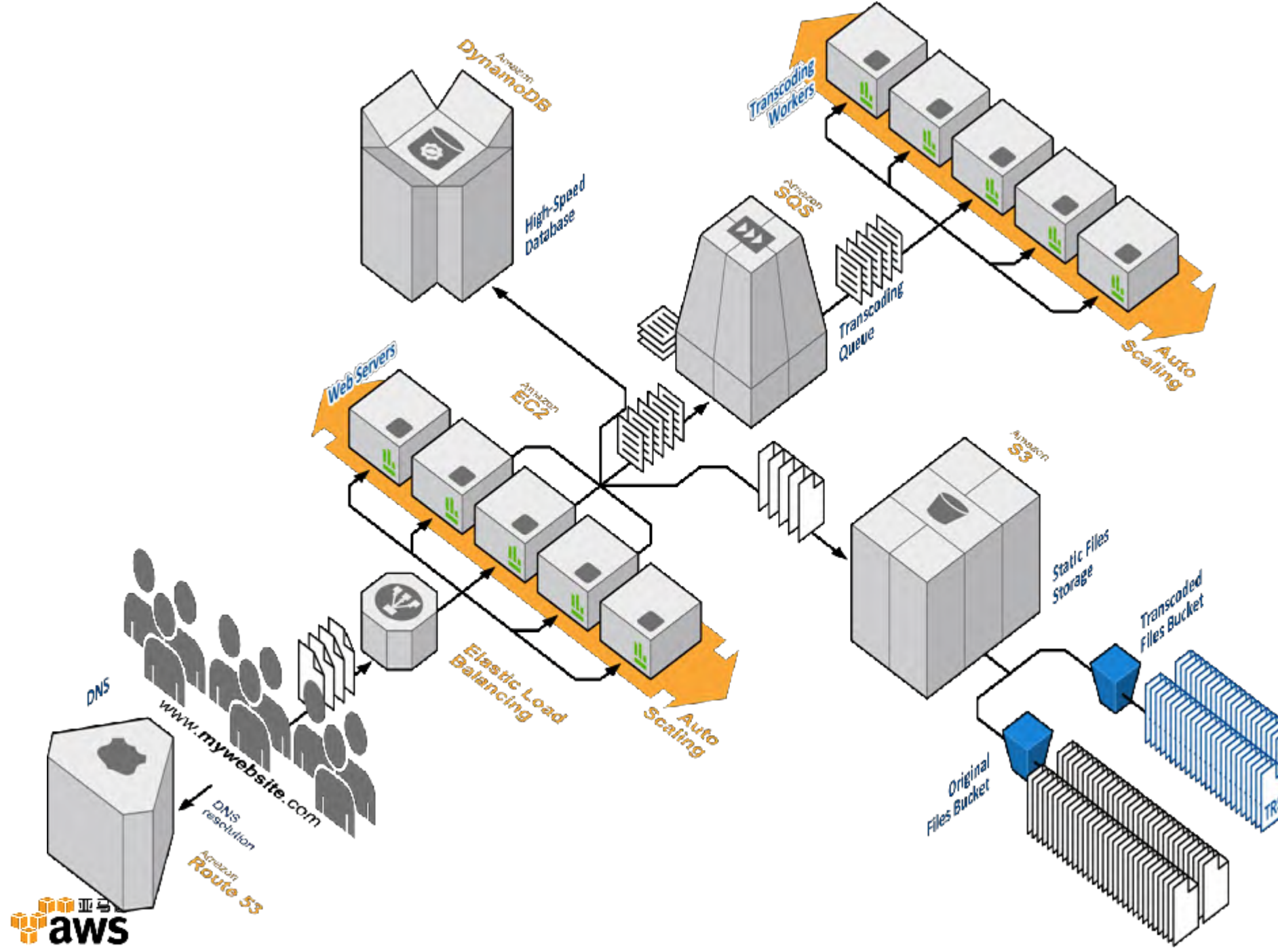


■ ApproximateNumberOfMessagesVisible ⓘ









测试你的高可用应用



总结

- 高可用应用设计原则



- AWS简化了HA应用的设计和部署！

参考资料

- AWS官网
 - <http://aws.amazon.com>
- AWS参考架构
 - <http://aws.amazon.com/architecture/>
- AWS白皮书
 - <http://aws.amazon.com/whitepapers/>
- AWS英文博客
 - <http://aws.typepad.com>
- 在线动手实验
 - <http://run.qwiklab.com/>