AWS SOLUTION ARCHITECT

Client service report

Author
Niall CREECH

October 8, 2019

Contents

1 Solution analysis

The current implementation cannot be accessed from external networks. Here we will go through the current setup, look at the issues that are preventing successful access by customers, and provide minimum changes to return the site to working health.

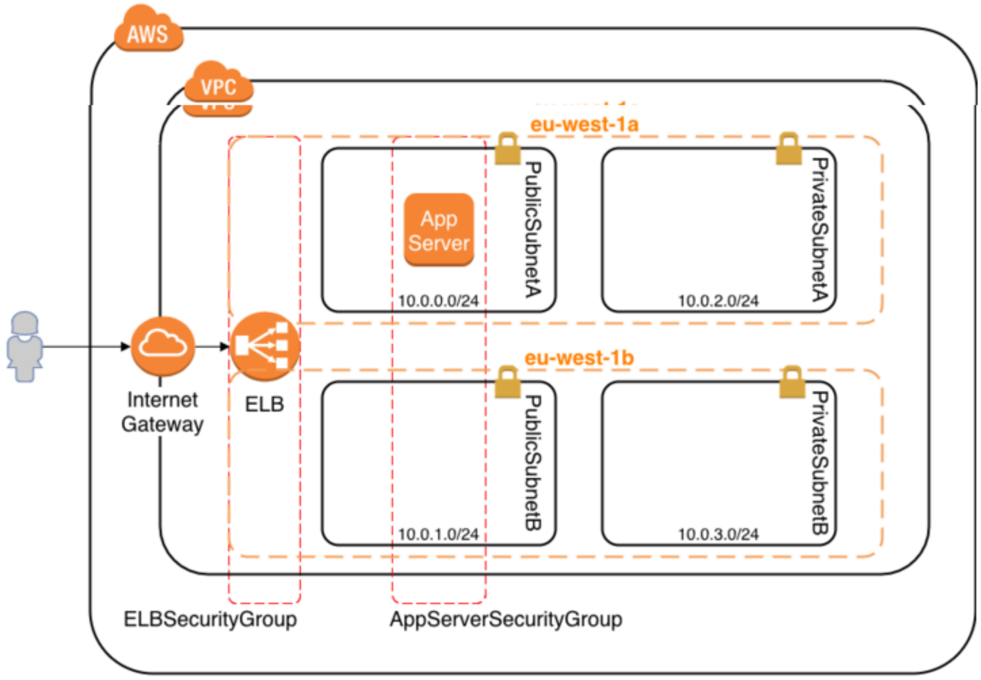


Figure 1: Architecture diagram of the current site

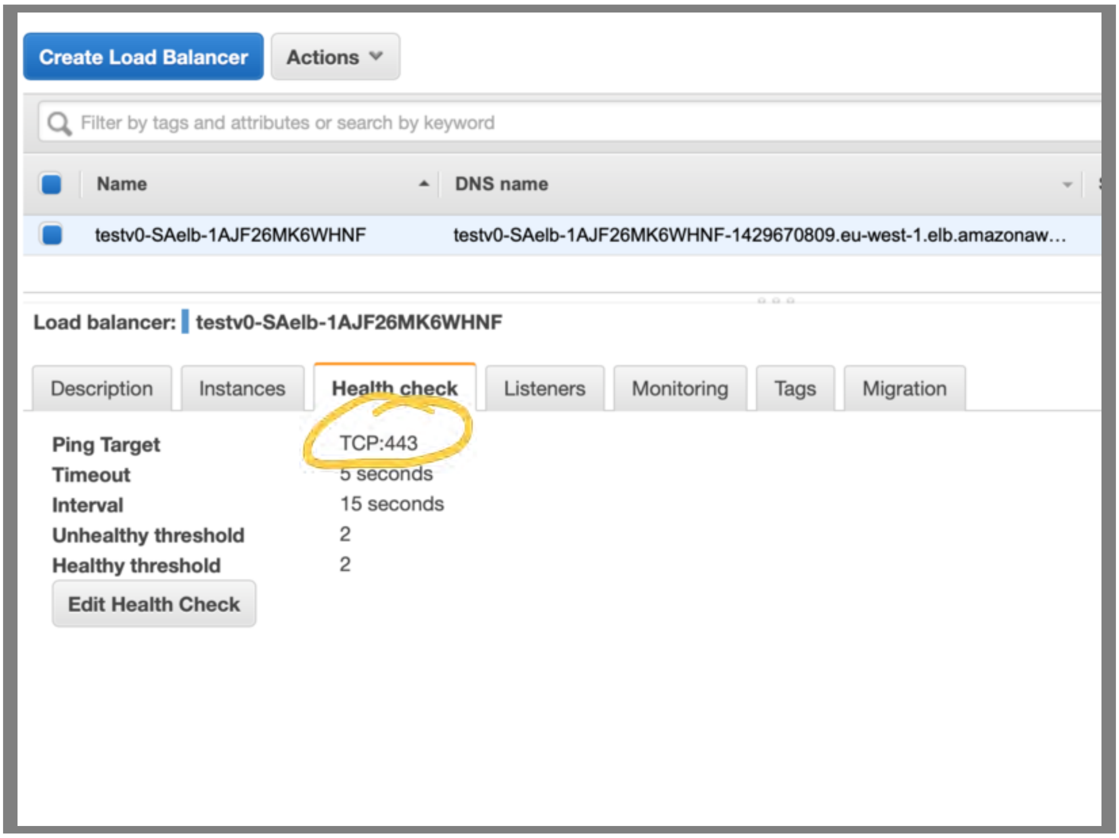
Summary of the setup of the initial solution (Figure ??)

(1) Networking is a VPC with CIDR 10.0.0.0/16 with 2 public subnets 10.0.0.0/24 and 10.0.1.0/24 in eu-west-1b and eu-west-1b respectively, as well as 2 private subnets 10.0.2.0/24, 10.0.3.0/24 in eu-west-1b and eu-west-1b.
(2) VPC has an internet gateway and a classic load balancer (ELB)
(3) Routing is setup up to allow all traffic between subnets and out to external addresses.
(4) There are 2 security groups, one for the single EC2 application instance, and one for the ELB.
(5) DNS is enabled and the instance has a pubic IP address

1.1 Troubleshooting

ELB healthcheck on wrong instance port

Summary: The EC2 instance runs a user-data script that installs and starts `Httpd` on port 80. The Elastic load balancer (ELB) correctly forwards ports from from 80 to instance port 80. However, the load balancers healthcheck looks to verify instance health on port 443 through a tcp connection. See `https://docs.aws.amazon.com/elasticloadbalancing/latest/classic/elb-troubleshooting.html`.



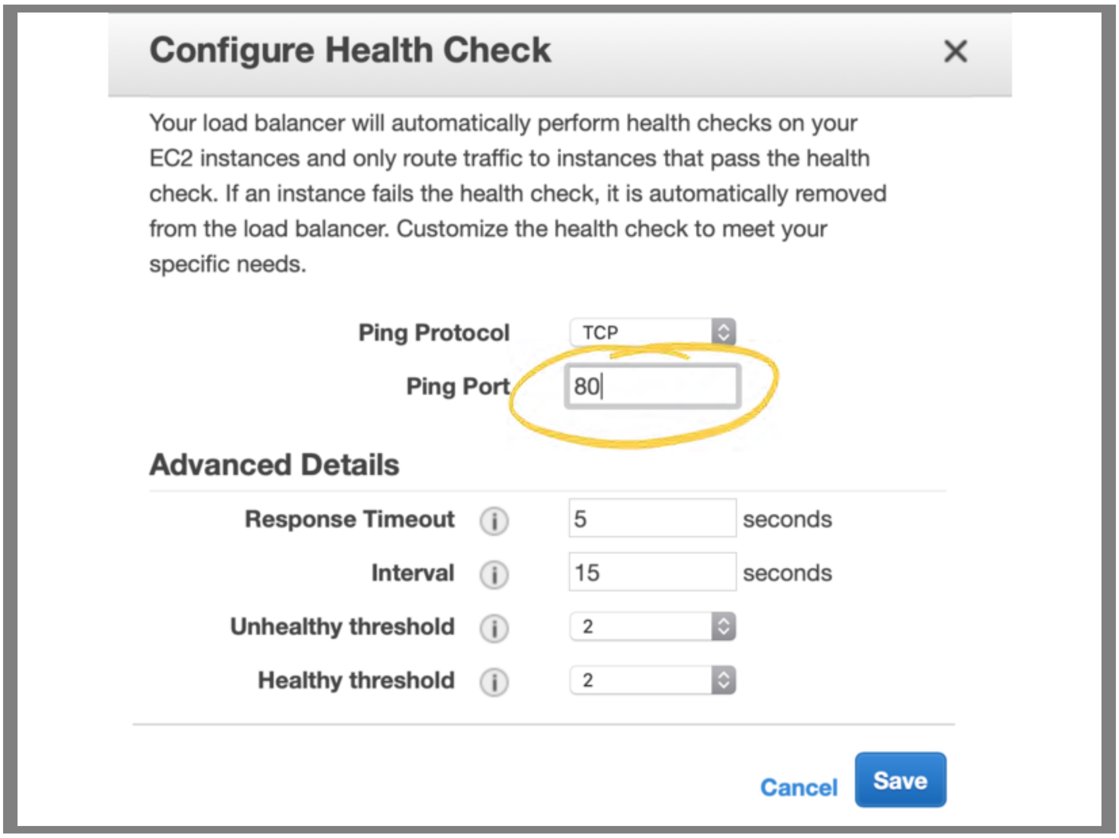
Name	DNS name
testv0-SAelb-1AJF26MK6WHNF	testv0-SAelb-1AJF26MK6WHNF-1429670809.eu-west-1.elb.amazonaw…

Load balancer: testv0-SAelb-1AJF26MK6WHNF

Ping Target	TCP:443
Timeout	5 seconds
Interval	15 seconds
Unhealthy threshold	2
Healthy threshold	2

Figure 2: In the AWS console, navigate to Services → EC2 → Load Balancers and select the 'Health check' tab.

Resolution:



Configure Health Check

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. Customize the health check to meet your specific needs.

Ping Protocol: TCP

Ping Port: 80

Advanced Details

Response Timeout: 5 seconds

Interval: 15 seconds

Unhealthy threshold: 2

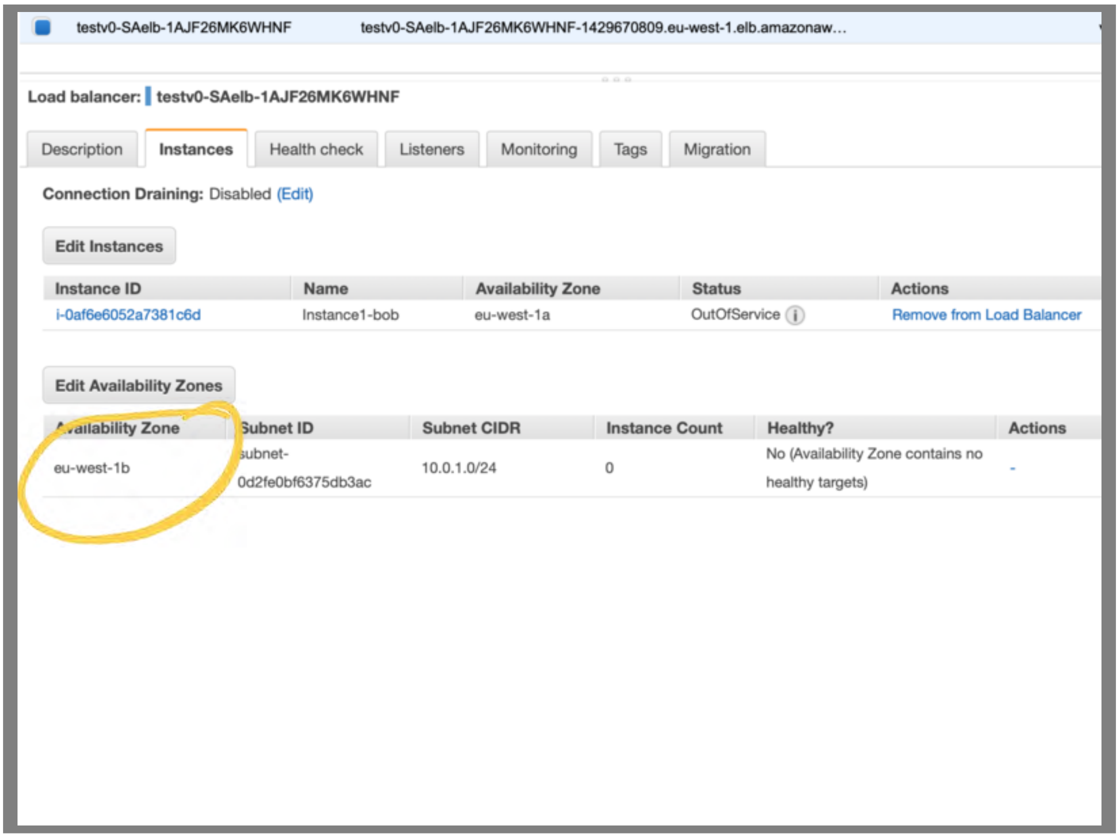
Healthy threshold: 2

Figure 3: Change the port from 443 to 80. This will leave the load balancer checking the instance health by connecting through tcp on the servers default HTTP port.

Load balancer cannot route to subnet

Summary: The load balancer can only route to instances in `PublicSubnetB`. However, the instance running the application is in `PublicSubnetA` so the ELB cannot pass traffic to it.

Resolution: By adding `PublicSubnetA` to the load balancers available subnets the instance can have traffic routed through to it.



testv0-SAelb-1AJF26MK6WHNF testv0-SAelb-1AJF26MK6WHNF-1429670809.eu-west-1.elb.amazonaw…

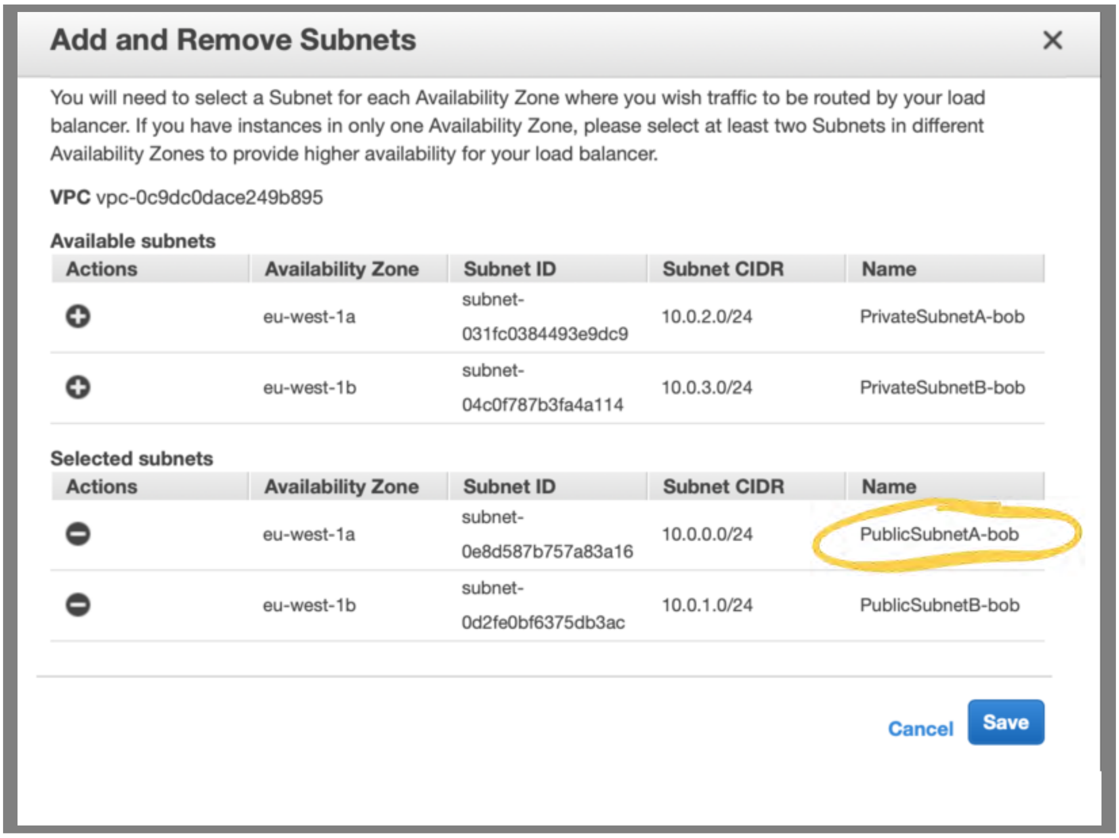
Load balancer: testv0-SAelb-1AJF26MK6WHNF

Connection Draining: Disabled (Edit)

Instance ID	Name	Availability Zone	Status	Actions
i-0af6e6052a7381c6d	Instance1-bob	eu-west-1a	OutOfService ⓘ	Remove from Load Balancer

Availability Zone	Subnet ID	Subnet CIDR	Instance Count	Healthy?	Actions
eu-west-1b	subnet-0d2fe0bf6375db3ac	10.0.1.0/24	0	No (Availability Zone contains no healthy targets)	-

Figure 4: In the Load Balancers configuration page, select the load balancer and then the 'Instances' tab. Note that only the `eu-west-1b` zone is available to the load balancer, which has no instances



Add and Remove Subnets

You will need to select a Subnet for each Availability Zone where you wish traffic to be routed by your load balancer. If you have instances in only one Availability Zone, please select at least two Subnets in different Availability Zones to provide higher availability for your load balancer.

VPC vpc-0c9dc0dace249b895

Available subnets

Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
⊕	eu-west-1a	subnet-031fc0384493e9dc9	10.0.2.0/24	PrivateSubnetA-bob
⊕	eu-west-1b	subnet-04c0f787b3fa4a114	10.0.3.0/24	PrivateSubnetB-bob

Selected subnets

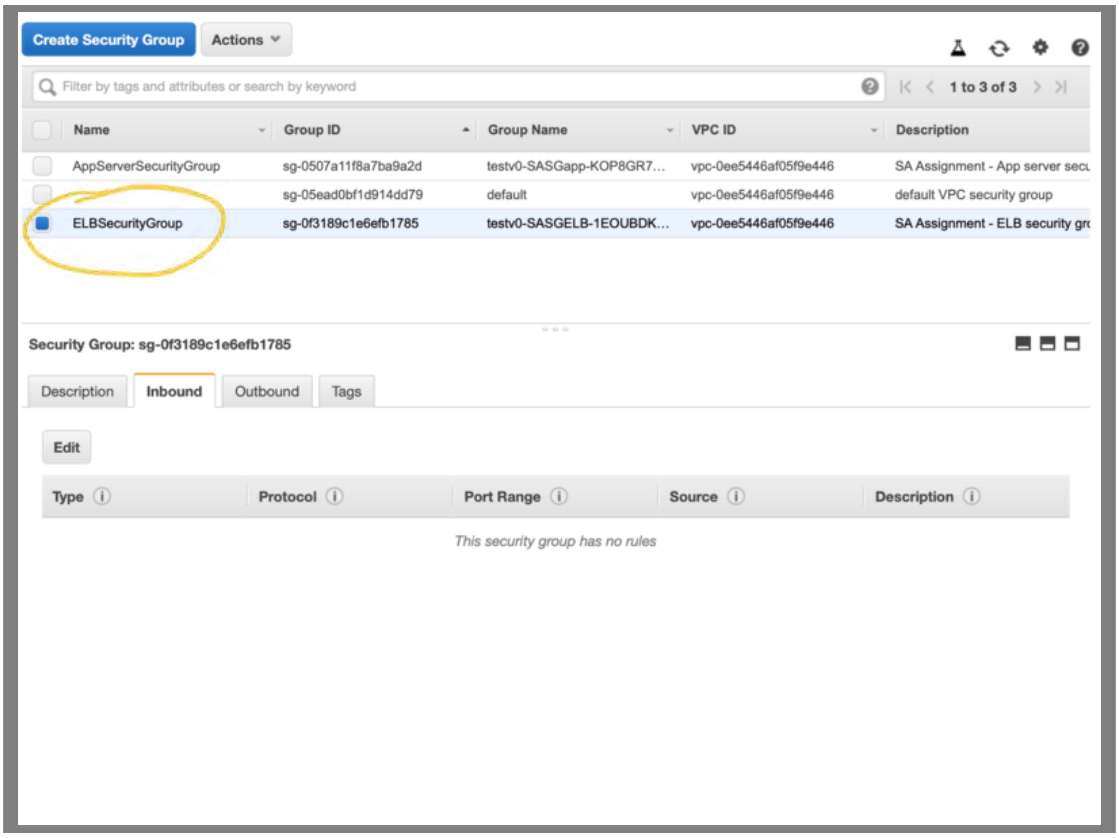
Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
⊖	eu-west-1a	subnet-0e8d587b757a83a16	10.0.0.0/24	PublicSubnetA-bob
⊖	eu-west-1b	subnet-0d2fe0bf6375db3ac	10.0.1.0/24	PublicSubnetB-bob

Figure 5: Add `PublicSubnetA` to the load balancers available subnets. This means that the load balancer can now direct traffic to the instance in `eu-west-1a`

The site cannot be accessed from external addresses

Summary: There are no inbound security group rules for the group `ELBSecurityGroup`, which means that the ELB will not route any traffic to the instance.

Resolution: As we want the site to be accessible by everyone, we allow all addresses, by using the CIDR range 0.0.0.0/0, to access the load balancer on port 80.



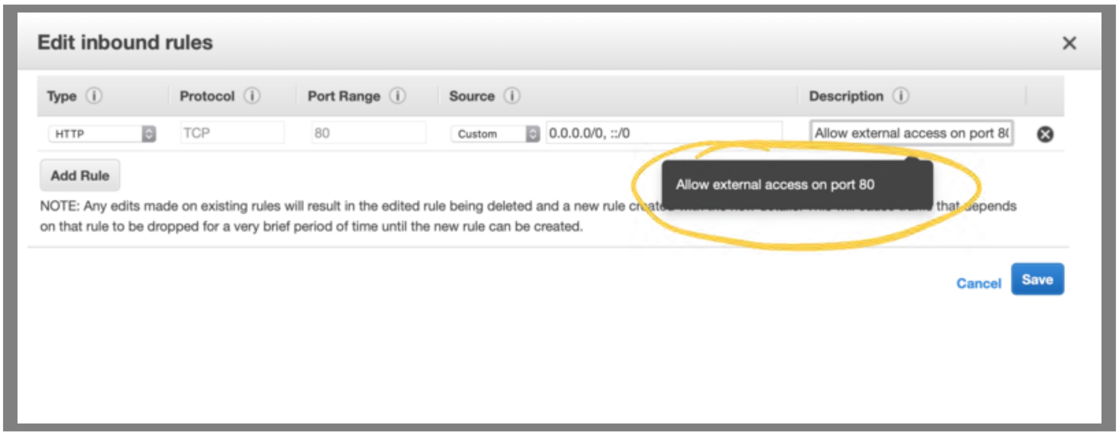
	Name	Group ID	Group Name	VPC ID	Description
□	AppServerSecurityGroup	sg-0507a11f8a7ba9a2d	testv0-SASGapp-KOP8GR7...	vpc-0ee5446af05f9e446	SA Assignment - App server secu
□		sg-05ead0bf1d914dd79	default	vpc-0ee5446af05f9e446	default VPC security group
☑	ELBSecurityGroup	sg-0f3189c1e6efb1785	testv0-SASGELB-1EOUBDK...	vpc-0ee5446af05f9e446	SA Assignment - ELB security gr

Security Group: sg-0f3189c1e6efb1785

Type	Protocol	Port Range	Source	Description

This security group has no rules

Figure 6: Navigate to Services → EC2 → Security Groups. Select the `ELBSecurityGroup` group and the 'Inbound' tab. By default, no inbound rules are set.


Edit inbound rules

Type	Protocol	Port Range	Source		Description
HTTP	TCP	80	Custom	0.0.0.0/0, ::/0	Allow external access on port 8(

Allow external access on port 80

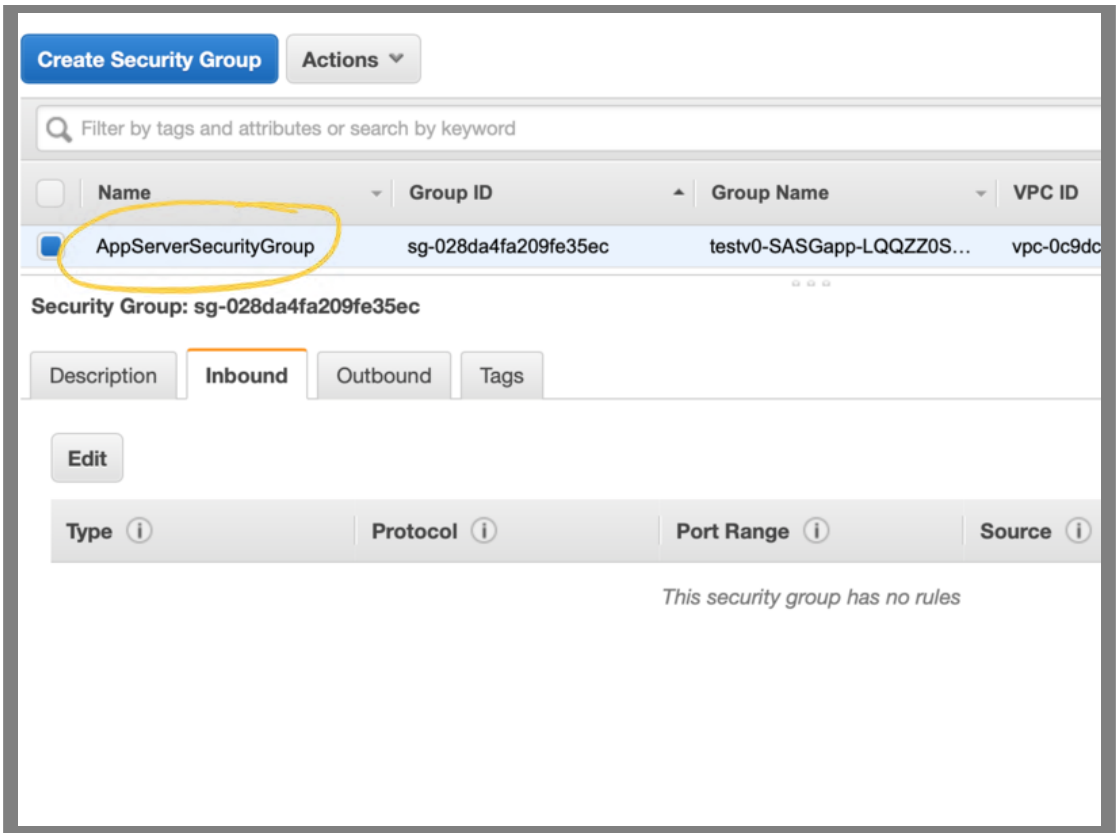
NOTE: Any edits made on existing rules will result in the edited rule being deleted and a new rule creat [illegible] that depends on that rule to be dropped for a very brief period of time until the new rule can be created.

Figure 7: Edit the group and add an inbound rule to allow http access from everywhere, 0.0.0.0/0. Users will now be able to reach the load balancer from the internet and other external networks.

Instances cannot be accessed from the load balancer

Summary: The current security group for the EC2 instances does not contain any inbound rules. This means that traffic from the load balancer is blocked

Resolution: We can change the security group to allow traffic from the ELB to the instance by adding a rule to allow TCP traffic to port 80. This also allows the ELBs TCP-based healthcheck to pass



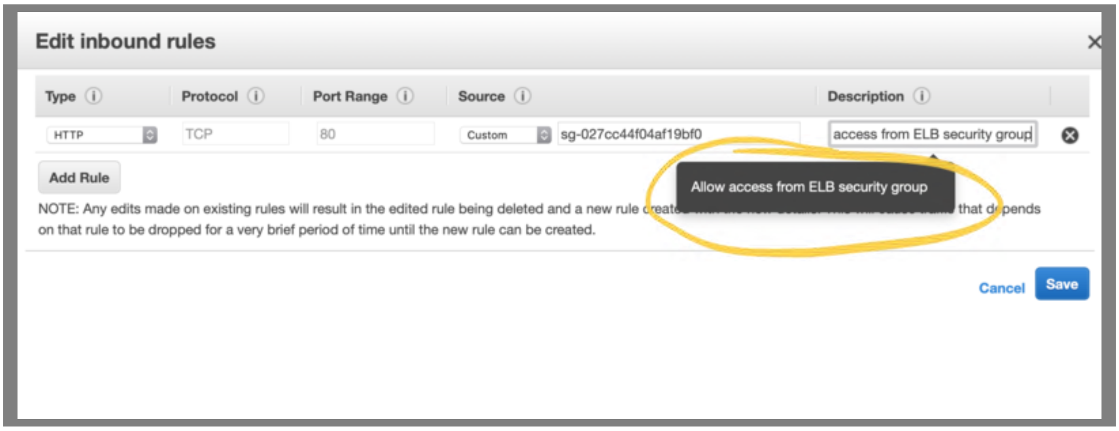
Name	Group ID	Group Name	VPC ID
AppServerSecurityGroup	sg-028da4fa209fe35ec	testv0-SASGapp-LQQZZ0S...	vpc-0c9dc

Security Group: sg-028da4fa209fe35ec

Type	Protocol	Port Range	Source

This security group has no rules

Figure 8: Still in the 'Security Groups' section, select the AppServerSecurityGroup group and the 'Inbound' tab. As we can see, there are no rules set.



Edit inbound rules

Type	Protocol	Port Range	Source		Description
HTTP	TCP	80	Custom	sg-027cc44f04af19bf0	access from ELB security group

Allow access from ELB security group

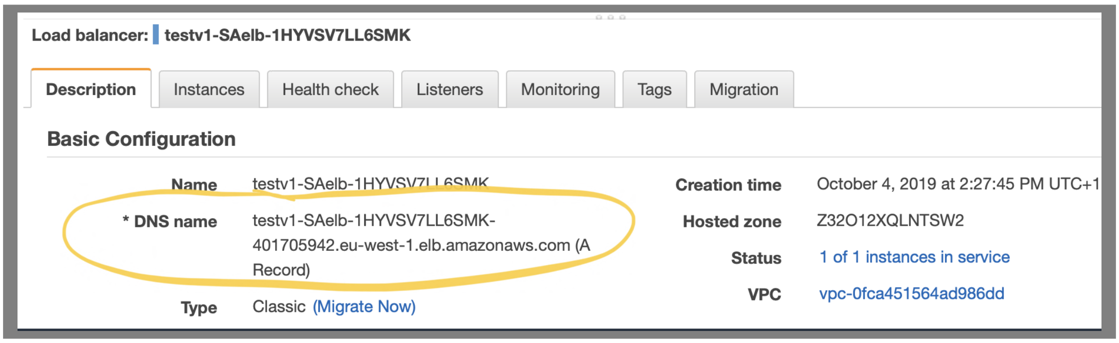
NOTE: Any edits made on existing rules will result in the edited rule being deleted and a new rule creat[illegible] that depends on that rule to be dropped for a very brief period of time until the new rule can be created.

Figure 9: Edit the group and add an inbound rule to allow http access from the ELBSecurityGroup. The security group Id needed for the 'Source' value can be found by typing 'sg' then choosing from one of the values in the popup menu.

1.2 Success!

The DNS name for the site can be found on the 'Description' tab of the 'Load balancers' configuration page when your load balancer is selected.

```
http://<load-balancer-dns-name>/demo.html
```

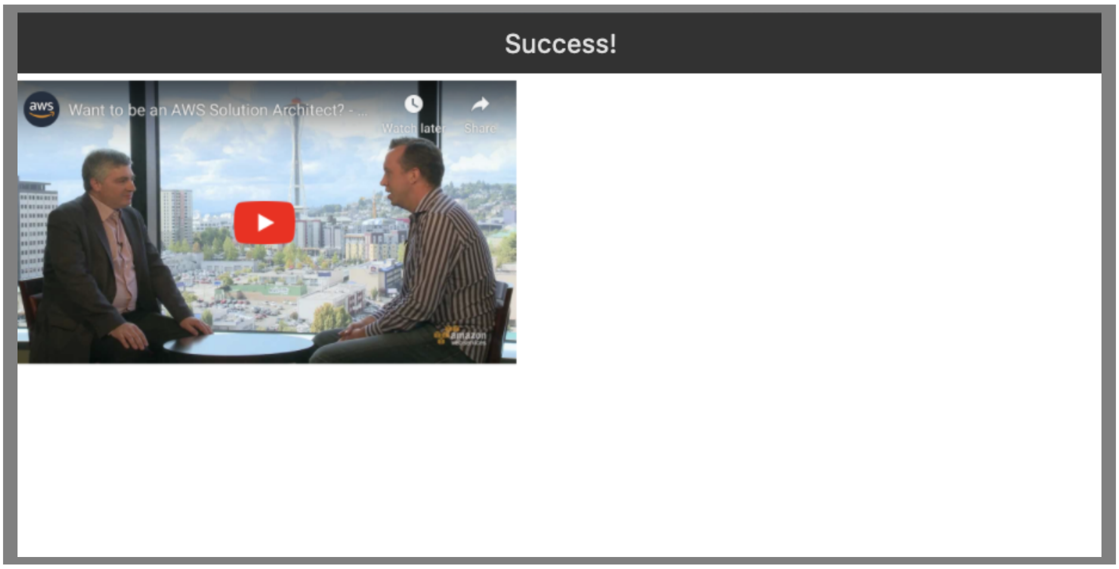


Load balancer: testv1-SAelb-1HYVSV7LL6SMK

Basic Configuration

Name	testv1-SAelb-1HYVSV7LL6SMK	Creation time	October 4, 2019 at 2:27:45 PM UTC+1
* DNS name	testv1-SAelb-1HYVSV7LL6SMK-401705942.eu-west-1.elb.amazonaws.com (A Record)	Hosted zone	Z32O12XQLNTSW2
		Status	1 of 1 instances in service
Type	Classic (Migrate Now)	VPC	vpc-0fca451564ad986dd

Figure 10: The DNS name of the load balancer can be found in the load balancers description tab



Success!

Want to be an AWS Solution Architect? - ...

Figure 11: The instances are now reachable through the load-balancer to customers on the internet

2 Solution Enhancement

There are a number of small-scale improvements we can make to the initial solution that will bring benefits in reliability, security, and observability. These improvements are small and iterative but have a large impact on the production performance.

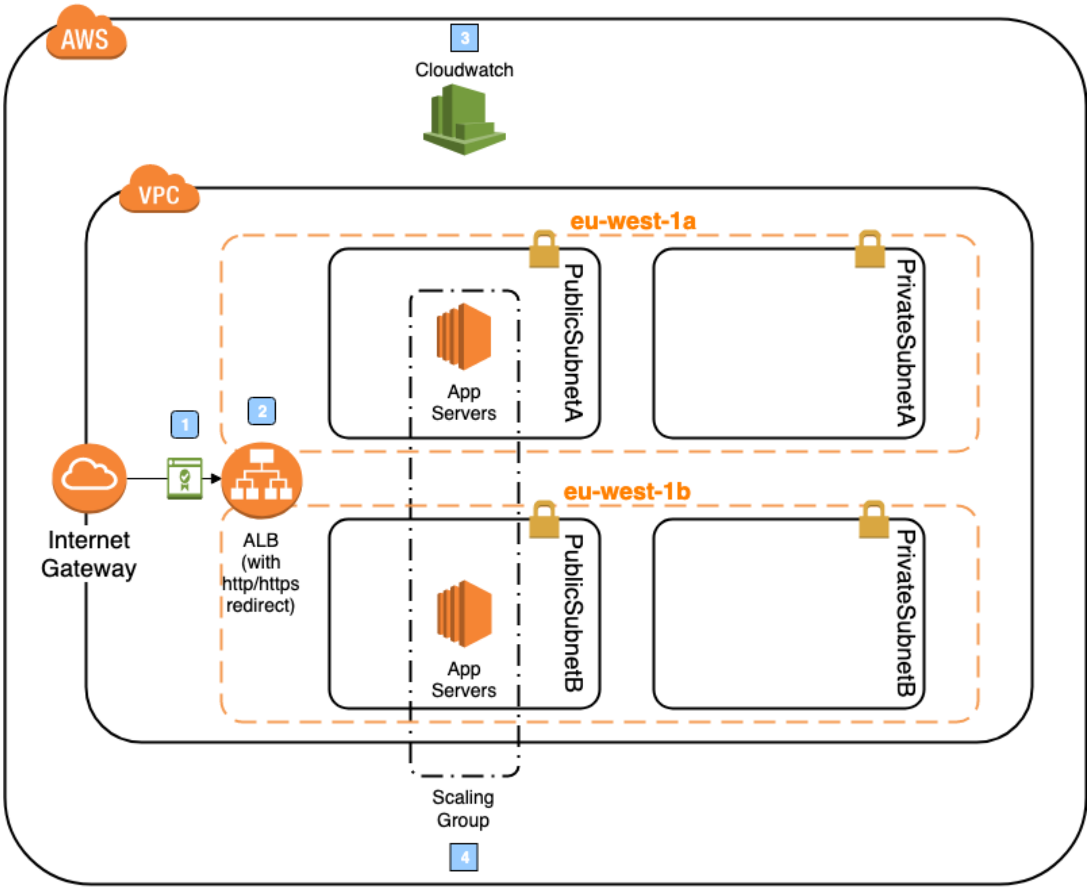


Figure 12: Architecture diagram of an improved solution.

Changes and enhancements to the initial solution (Figure ??)

(1) Addition of ACM certificate to load balancer to allow HTTPS to be enabled

(2) Moving to Application Load Balancer(ALB) for HTTP->HTTPS redirection and layer 7 routing and WAF filtering

(3) Enable detailed instance and load balancer metrics, shipping to Cloudwatch with dashboard. Log shipping and OS-level metrics through collectd and cloudwatch agent installation on AMI

(4) Instances placed in Auto-scaling group (ASG) for automatic management and healing of instances using load balancer HTTP:80 healthcheck

2.1 Operational

The ability to run and monitor systems to deliver business value and to continually improve supporting processes and procedures

2.1.1 Improve monitoring and logging

Monitoring and logging on the existing instance can be increased by enabling the cloudwatch agent on the application instances

2.2 Security

The ability to protect information, systems, and assets while delivering business value through risk assessments and mitigation strategies

2.2.1 Harden EC2 instances

2.2.2 Enable SSM session manager

2.2.3 Strengthen NACL to prevent cross-subnet traffic

2.2.4 Only allow NACL to pass traffic to private subnets from public subnets

The private NACL allows all external traffic to pass. We can increase security by allowing only access from public subnets, or additionally from private endpoints, NAT gateways etc.

2.3 Reliability

The ability of a system to recover from infrastructure or service disruptions, dynamically acquire computing resources to meet demand, and mitigate disruptions such as misconfigurations or transient network issues

Currently there is a single instance with no auto-recovery

2.3.1 Replicate instances into multiple AZs

2.3.2 Create an auto-scaling group for application instances

2.4 Performance

The ability to use computing resources efficiently to meet system requirements, and to maintain that efficiency as demand changes and technologies evolve

2.4.1 Change instance type

The current solution uses bursting t2.micro instances with CPU credit limits. This is very cost efficient, but performance is not suitable for reasonable traffic levels and exhausting CPU credits will cause degradation of availability.

2.5 Cost

The ability to run systems to deliver business value at the lowest price point

3 Future expansion

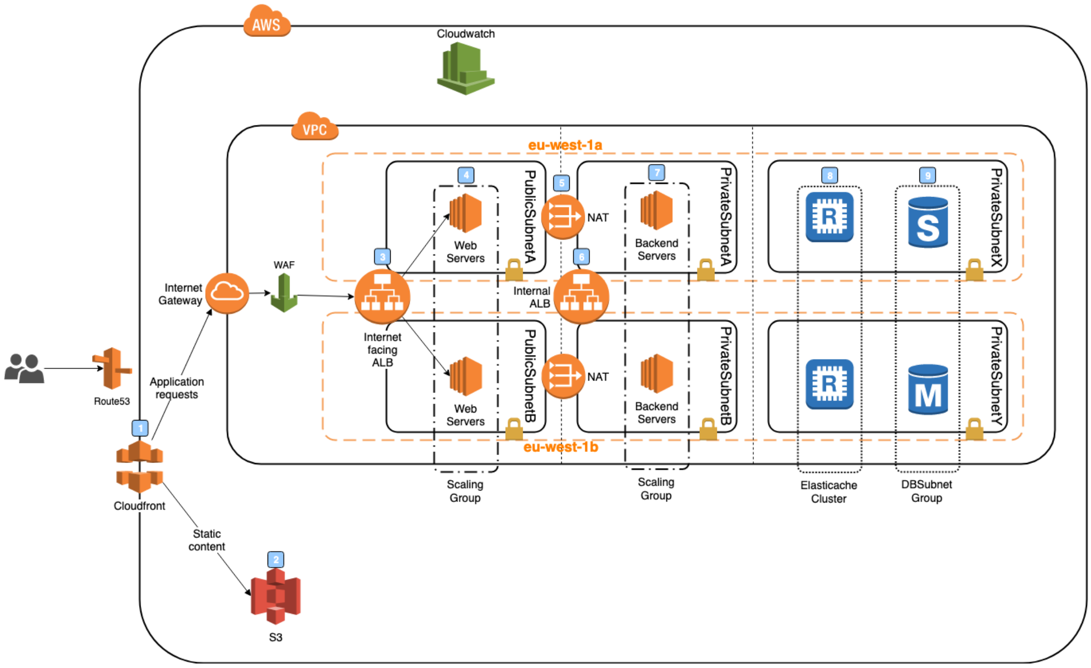


Figure 13: Large scalable production solution

Building on the initial solution but with the use of a number of additional AWS services, the site can scale globally and meet the performance demands of a high capacity business

(1) Users get DNS resolution from Route53 entries
(2) Cloudfront provides cached content from edge locations around the world.
(3) A secured S3 bucket provides a highly scalable source of static content
(4) Active content is sourced through a public application load balancer (ALB).
(5) WAF is configured with rules in front of the ALB
(6) Applications requiring public internet addresses are installed on EC2 instances inside an auto-scaling group (ASG) in public subnets distributed across more than one availability zone.
(7) The NAT gateway service provides address translation to allow outbound internet access for the private subnets
(8) Applications that do not require public internet addresses are installed on EC2 instances inside an ASG in private subnets distributed across more than one availability zone.
(9) Elasticache (redis) provides in-memory caching for the database backend.
(10) RDS in multi-AZ mode gives resilient failover from primary to secondary, as well as < 5 min point in time recovery.

References

[1] Amazon Web Services. Amazon capacity planning for sap on aws. https://d1.awsstatic.com/enterprise-marketing/SAP/capacity-planning-for-sap-on-aws.pdf, 2019.