



Triển khai Multi-Region ECS Cluster với Blue-Green Deployment

SINH VIÊN: ĐẶNG VĂN KIÊN

MSSV :2180600712

MENTOR :VĂN HOÀNG KHA

THỜI GIAN: 01/07/2025 - 10/08/2025

Executive Summary

Dự án: Thiết kế và triển khai hệ thống đa vùng sử dụng Amazon ECS Fargate nhằm đảm bảo khả năng phục hồi thảm họa (Disaster Recovery), tăng độ sẵn sàng dịch vụ (High Availability), và giảm thiểu tối đa thời gian gián đoạn hệ thống (Zero-downtime Deployment).

- Xây dựng hệ thống ESC Fargate đa vùng (Singapore, Tokyo)
- Áp dụng Blue-Green Deployment đảm bảo Zero-downtime
- Hỗ trợ tự động chuyển vùng khi có sự cố
- Giám sát và cảnh cáo bằng Cloudwatch+SNS

Chi Phí Ước Tính:

Tổng chi phí hoạt động hàng tháng khoảng **39 USD**, giúp tiết kiệm hơn **61%** so với giải pháp EC2 tự quản.

Lợi ích kinh doanh:

- ✓ **Triển khai không gián đoạn:** Giảm downtime khi cập nhật phiên bản ứng dụng mới.
- ✓ **Khả năng phục hồi cao:** Tự động chuyển vùng nếu một khu vực gặp sự cố.
- ✓ **Tiết kiệm chi phí và công sức vận hành:** Không cần quản lý máy chủ vật lý hoặc EC2 lâu dài.
- ✓ **Chuẩn bị cho mở rộng toàn cầu:** Hạ tầng có thể nhân rộng dễ dàng ở các khu vực địa lý khác.

Problem Statement

Bối cảnh hiện tại:

- Nhiều hệ thống chỉ triển khai trên **một vùng AWS**, dẫn đến nguy cơ trở thành **single point of failure**.
- **Thiếu cơ chế kiểm soát triển khai (Blue-Green)** khiến việc cập nhật phiên bản dễ gây lỗi sản phẩm.
- **Không có cảnh báo chủ động hoặc tự động chuyển đổi**, gây trì hoãn trong xử lý sự cố.

Hậu quả thực tế:

- Downtime khiến **người dùng mất niềm tin**, ảnh hưởng đến **doanh thu và hình ảnh thương hiệu**.
- Theo **Atlassian**, mỗi phút downtime có thể gây thiệt hại lên đến **\$5.600**.
- **RingCentral** ước tính các công ty trong danh sách Fortune 500 có thể mất đến **\$46 triệu mỗi năm** do gián đoạn dịch vụ.

Problem Statement

Xu hướng thị trường:

Gartner (TechRepublic) dự báo: "Đến năm 2025, 85% doanh nghiệp sẽ ưu tiên chiến lược Cloud-first."

Mục tiêu đặt ra:

- ✓ Thiết kế hệ thống đa vùng giúp loại bỏ điểm chết hạ tầng.
- ✓ Đảm bảo triển khai nhanh chóng, không gián đoạn và có khả năng tự động phục hồi.

Calculating the cost of downtime

Understanding the financial impact of major incidents

In March 2015, a 12-hour Apple store outage cost the company [\\$25 million](#).

In August 2016, a five-hour power outage in an operation center caused 2,000 cancelled flights and an estimated loss of [\\$150 million](#) for Delta Airlines.

In March 2019, a 14-hour outage cost facebook [an estimated \\$90 million](#).

And those are the big guys. The industry leaders. The ones with fat operating margins and millions in the bank. They can weather a one-day financial storm. The truth is that while smaller companies may face smaller losses during a major incident, those smaller numbers can have an even bigger effect on their bottom line.

In fact, [one study of 101 startups](#) found that 29% of those that fail do so because they run out of cash. If startups are already at risk, it's hard to imagine most could weather a major incident without going under.

The moral of the story: downtime is a big deal. Anyone who says otherwise hasn't been paying attention. Incidents are not only potentially toxic to customer trust and loyalty. They're also the financial grim reaper.

85% of organizations will be "cloud-first" by 2025, says Gartner

Published November 12, 2021

Written by
[Brandon Vigliarolo](#)

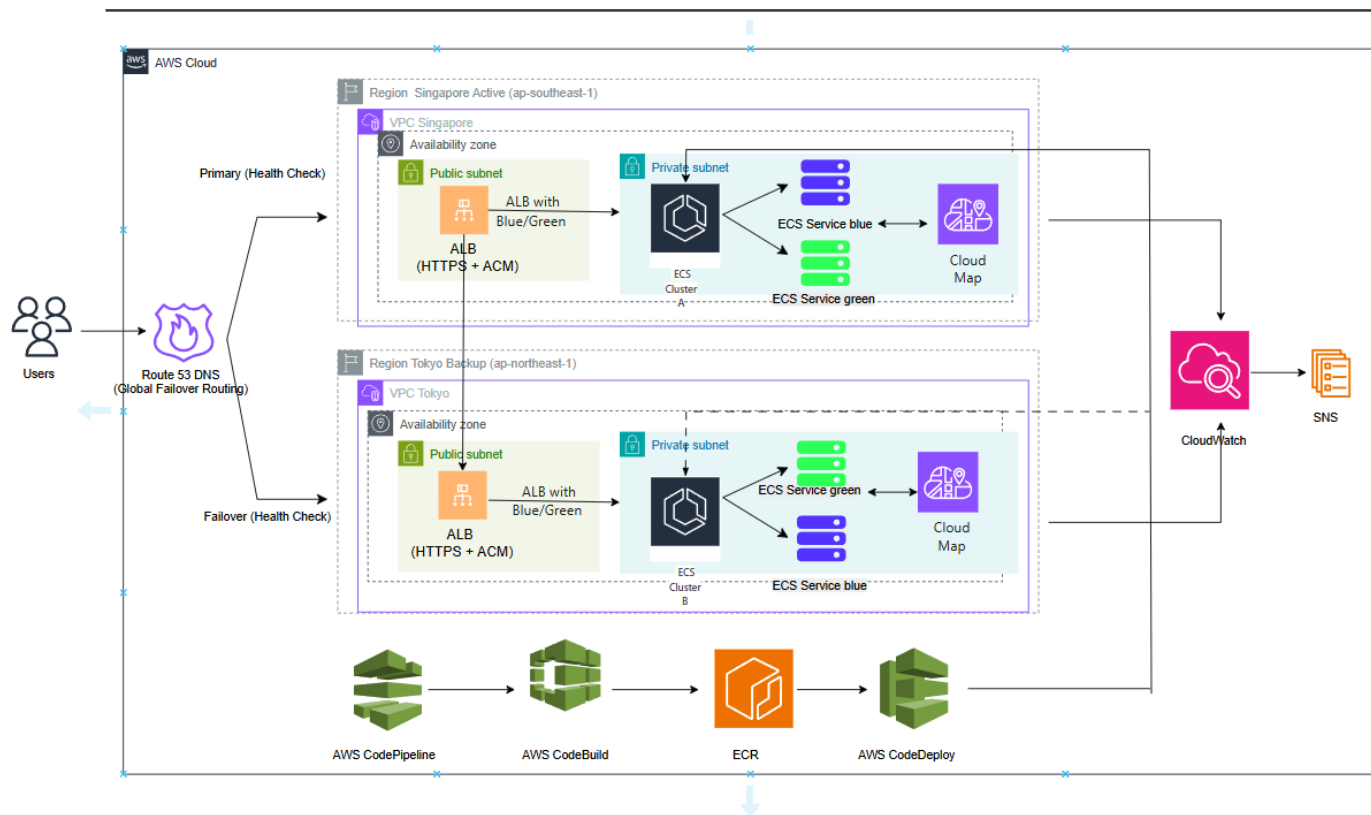
Digital transformation pressure, combined with pandemic-driven changes, will move us into a world where anything that isn't cloud is considered legacy, one Gartner analyst said.



Image: Shutterstock/Blackboard

The cloud, says Gartner, is going to be the centerpiece of new digital experiences. For businesses that means it's time to adopt a cloud-first approach, or risk languishing in an on-premise data center.

Solution Architecture



- ✓ **ECS Fargate** làm nền tảng triển khai ứng dụng dạng container.
- ✓ **ECS Cluster** ở 2 vùng (Blue & Green)
- ✓ **ALB** định tuyến theo path
- ✓ **Route 53 Health Check + Failover DNS**
- ✓ **CloudFormation** quản lý hạ tầng
- ✓ **CloudWatch** giám sát, **SNS** gửi alert

Solution Architecture

● User → Route 53 DNS (Failover)

- Người dùng truy cập domain ứng dụng (VD: app.example.com)
- Route 53 định tuyến đến ALB Singapore hoặc Tokyo dựa trên kết

quả health check

● Route 53 → Application Load Balancer (HTTPS + ACM)

- ALB là cổng vào ứng dụng, xử lý kết nối HTTPS và phân phối request đến ECS Service

● ALB → ECS Service (Fargate)

- ECS Task chạy trong môi trường không máy chủ (Fargate)
- ALB định tuyến đến Blue hoặc Green version tùy theo giai đoạn triển khai

● ECS → Cloud Map

- ECS dùng Cloud Map để phát hiện các service nội bộ trong Cluster

● CloudWatch → SNS

- CloudWatch theo dõi metric hệ thống như CPU, error rate, health check
- Gửi cảnh báo qua SNS đến Email hoặc các kênh khác khi phát hiện sự cố

● CodePipeline → CodeBuild → ECR → CodeDeploy → ECS

- Mỗi lần cập nhật code, pipeline tự động build image, đẩy lên ECR
- CodeDeploy triển khai bản mới vào ECS Green để kiểm thử
- Nếu thành công, ALB sẽ chuyển routing sang Green (Blue-Green Deployment)

🔒 IAM Role → ECS Task

- ECS Task được gán IAM Role để truy cập an toàn vào các dịch vụ AWS như ECR, CloudWatch, SNS mà không cần hardcoded credentials

Solution Architecture

Yếu tố bảo mật và hiệu năng:

- ✓ ECS Task chạy trong **Private Subnet**, chỉ ALB trong Public Subnet mới được phép truy cập.
- ✓ Kết nối HTTPS thông qua **AWS Certificate Manager (ACM)**.
- ✓ Tự động mở rộng (Auto Scaling) dựa trên CPU/MEM.

So sánh giải pháp:

Tiêu chí	EC2 tự quản lý	ECS Fargate
Quản lý hạ tầng	Phải tự cấu hình	Serverless, không cần quản lý
Độ ổn định triển khai	Dễ lỗi khi update	Hỗ trợ Blue-Green, rollback nhanh
Khả năng mở rộng	Thủ công	Auto Scaling tích hợp
Chi phí	~\$100/tháng	~\$39/tháng

Technical Implementation

TUẦN	NỘI DUNG CHÍNH
1	Docker image + ECR
2	ECS tại Singapore
3	ECS Tokyo + Blue-Green
4	Route 53 + test failover
5	Monitoring + báo cáo

Timeline & Milestones

Tuần	Công việc chính	Kết quả mong đợi
1	Build + Docker hóa + Push ECR	Docker Image sẵn sàng trên ECR
2	ECS Cluster + ALB Singapore	Triển khai Blue version hoạt động
3	Tokyo Cluster + Blue-Green Service	Có thể chuyển đổi giữa 2 phiên bản
4	Cấu hình Route 53 + test failover	Tự động chuyển sang Tokyo khi lỗi
5	Cảnh báo CloudWatch + Demo	Hệ thống giám sát hoàn thiện, sẵn sàng demo

Timeline & Milestones

Tuần	Mốc triển khai	Kết quả
Tuần 1(01–07/07)	Nghiên cứu kiến trúc ECS Multi-Region + Blue-Green Chuẩn bị môi trường AWS (VPC, IAM, ECR)	Hoàn thành sơ đồ kiến trúc và môi trường AWS sẵn sàng
Tuần 2(08–14/07)	Dockerize app mẫu, Push image lên ECR Tạo ECS Cluster Singapore + ALB	ECS Singapore hoạt động, app Blue triển khai thành công
Tuần 3(15–21/07)	Tạo Tokyo Cluster + thiết lập Blue-Green Thử nghiệm triển khai và chuyển đổi routing ALB	Đang tạo region tokyo
Tuần 4(22–28/07)	Cấu hình Route 53 Health Check + Failover DNS Giả lập lỗi và kiểm thử chuyển vùng tự động	Đang thực hiện
Tuần 5(29–04/08)	Thiết lập CloudWatch + SNS cảnh báo Viết báo cáo, demo nội bộ	Đang thực hiện
Tuần 6(05–10/08)	Final proposal, kiểm thử lần cuối, gửi báo cáo + chuẩn bị demo trình bày	--

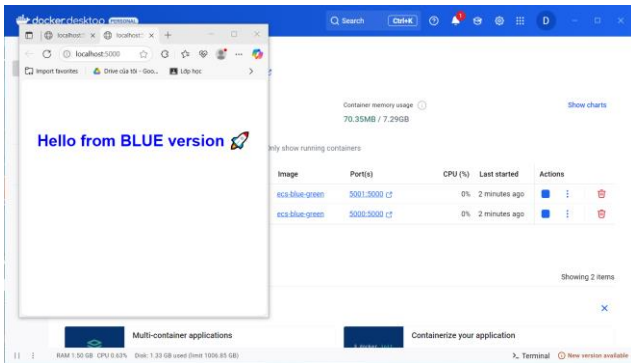
PROOF

```

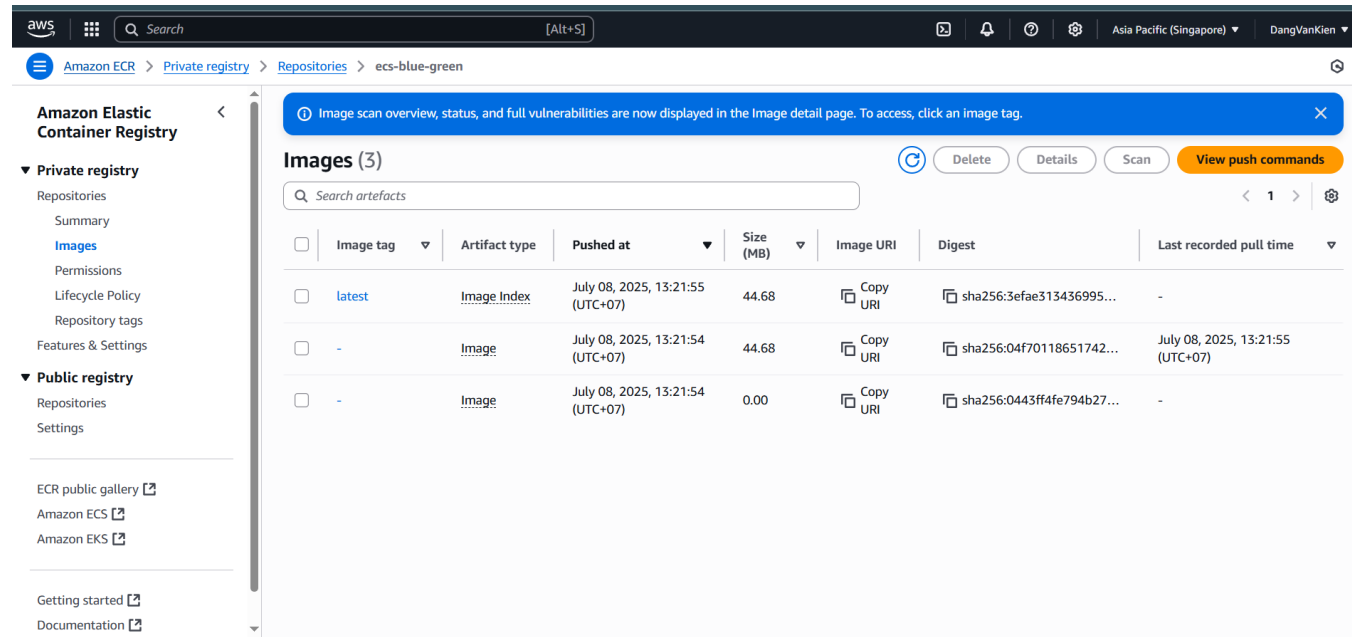
1 // index.js
2 const express = require('express');
3 const app = express();
4 const port = 3000;
5
6 const version = process.env.VERSION || "UNKNOWN";
7
8 app.get('/', (req, res) => {
9   const color = version === "blue" ? "blue" : "green";
10   res.send(
11     `Hello from ${version}! I'm ${color}!`
12   );
13 });
14
15 app.get('/health', (req, res) => {
16   res.status(200).send('OK');
17 });
18
19 app.listen(port, () => {
20   console.log(`App running on port ${port} - version: ${version}`);
21 });

```

Code web test bằng Node.js và Express

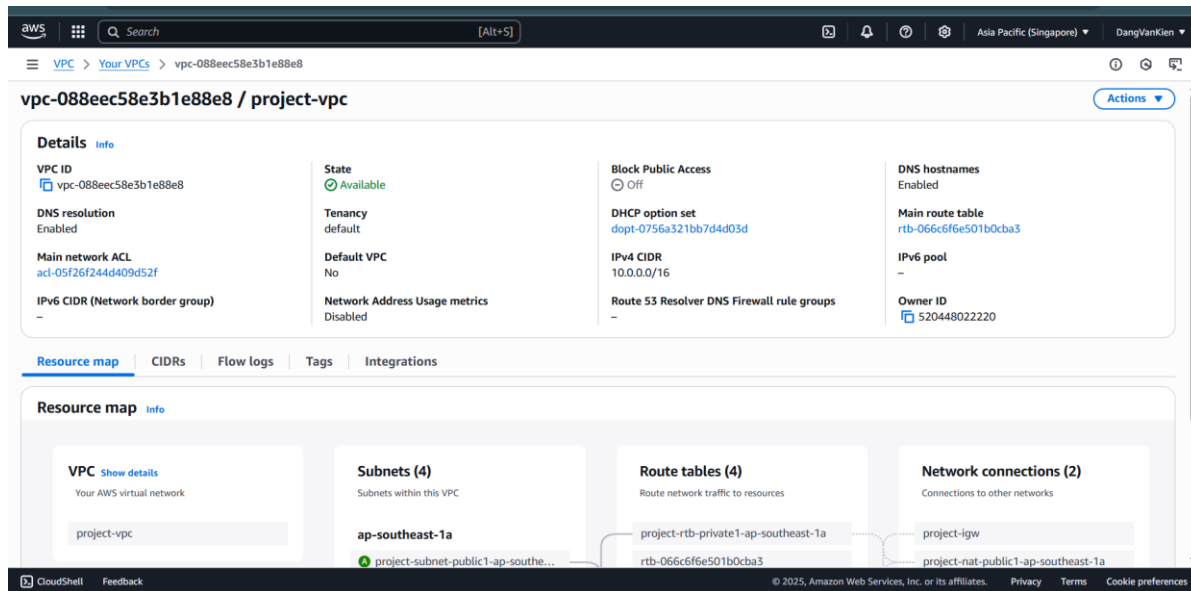


Đóng gói code vào Docker chạy thử cục bộ

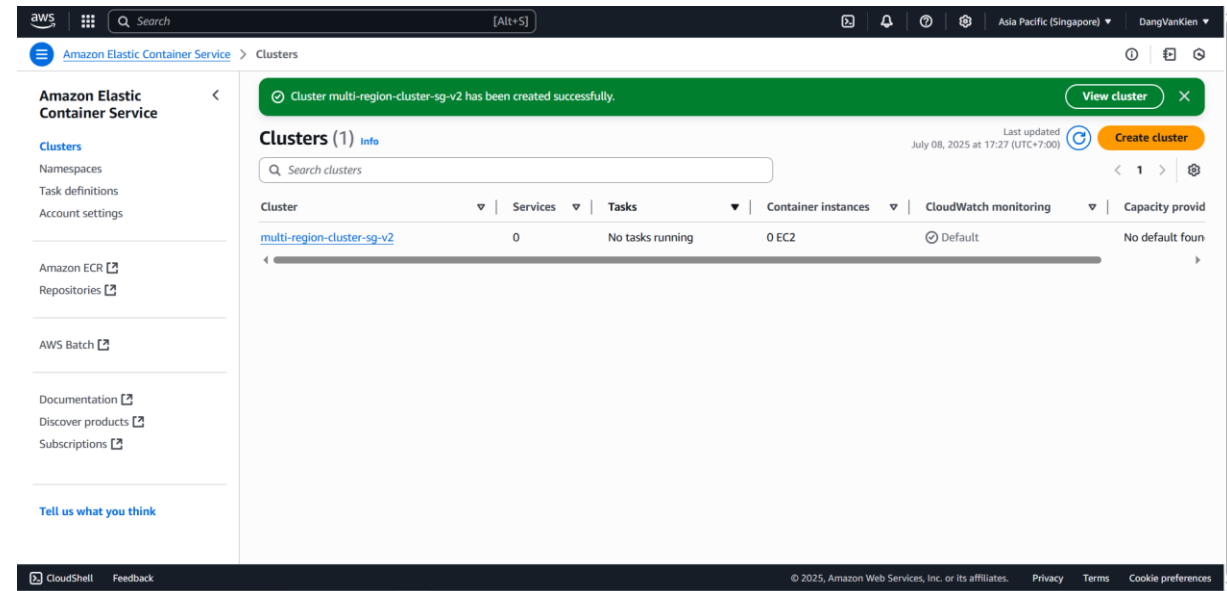


Push image lên ECR thành công

PROOF

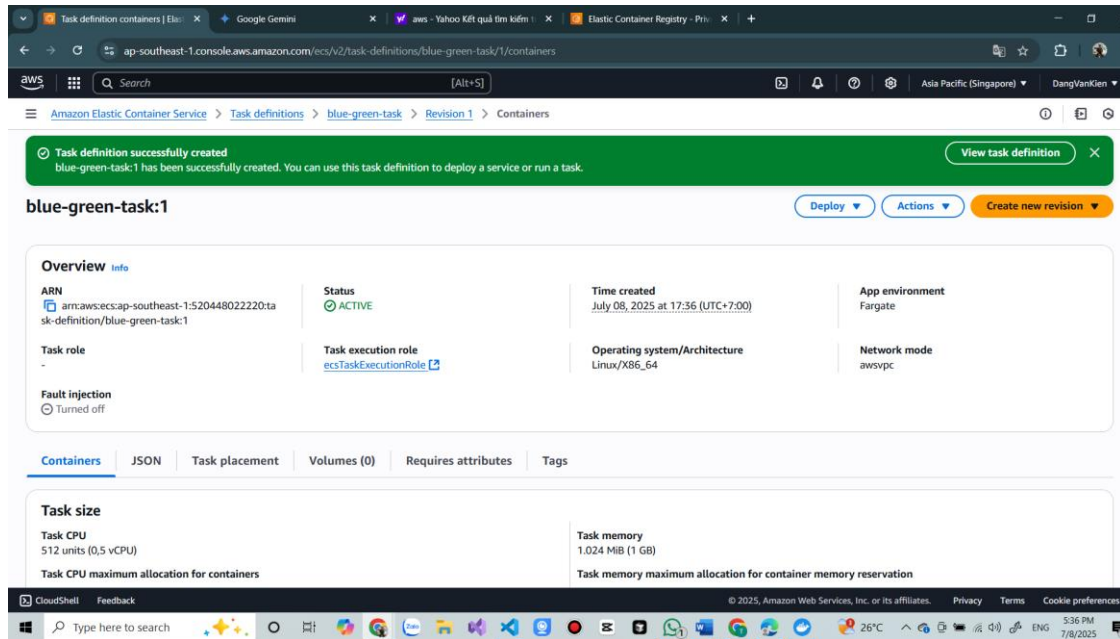


Cấu hình VCP cho Singapore thành công

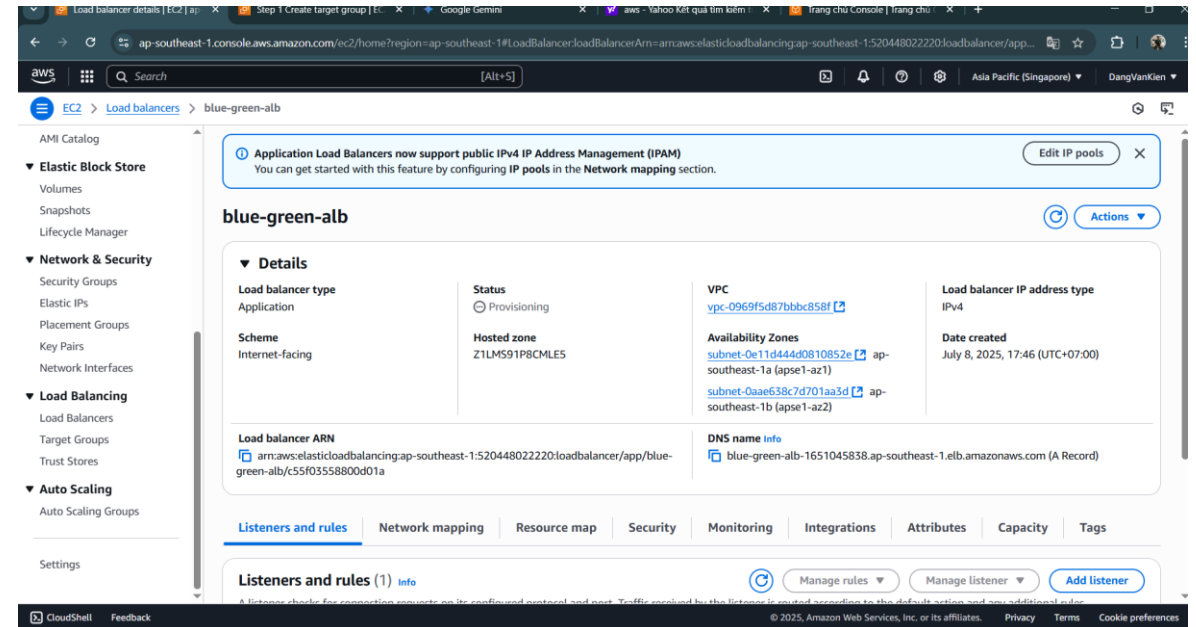


Cấu hình ECS Cluster cho Singapore thành công

PROOF

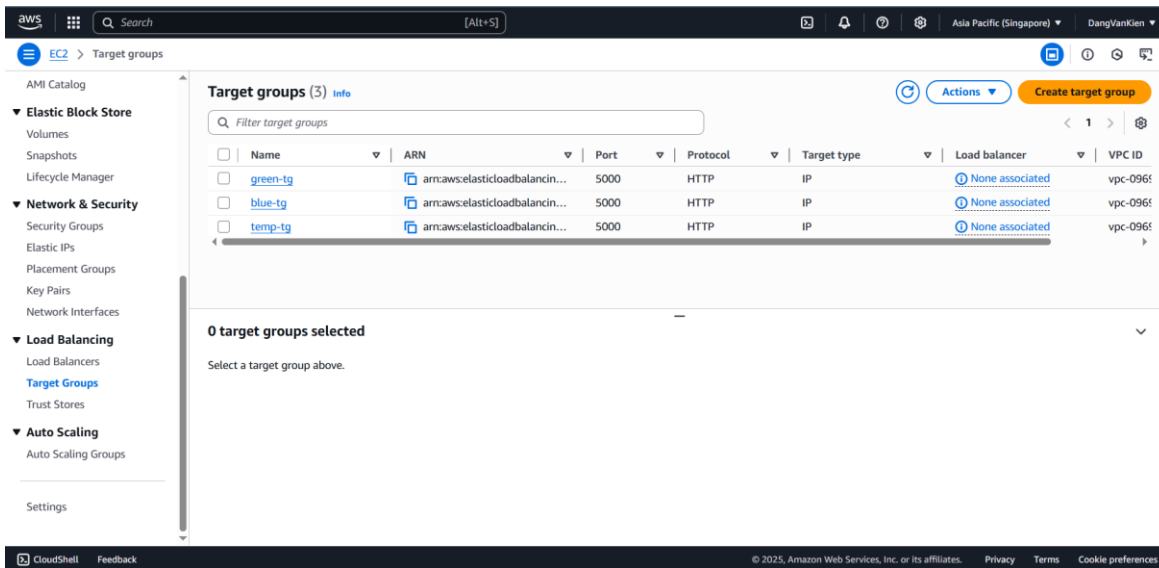


Cấu hình task definitions

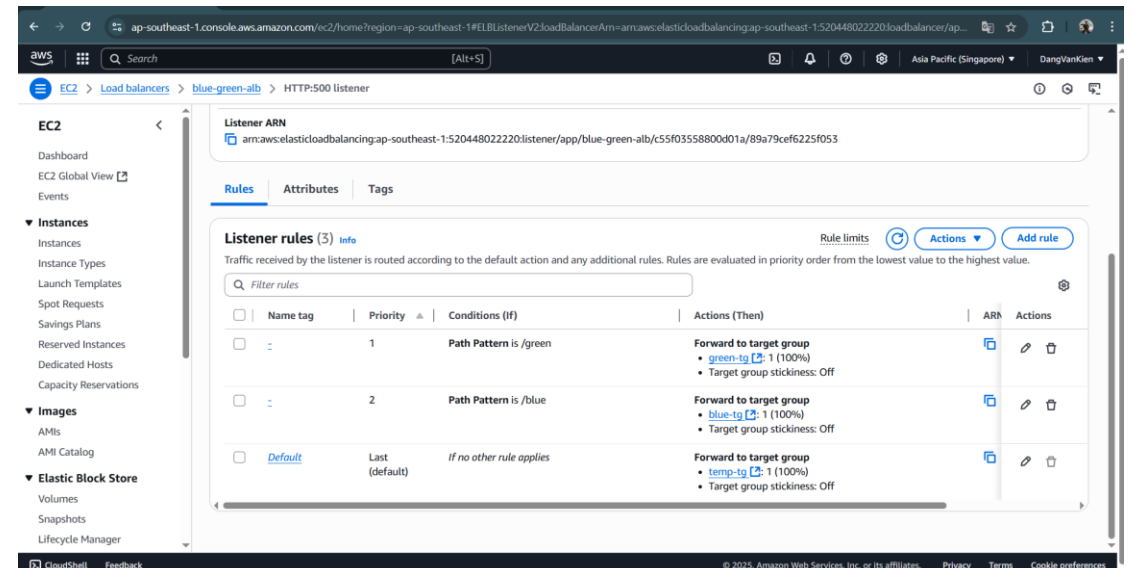


Cấu hình ALB cho Singapore thành công

PROOF



Cấu hình target Group cho blue và green



Add Rule và check ALB hoạt động thành công

➤ Hiện tại tiếp tục chuẩn bị và thực hiện các cấu hình cho Tokyo và các dịch vụ aws khác theo tiến trình

Budget Estimation

Ước tính chi phí hàng tháng (theo AWS Calculator):

Dịch vụ	Ước tính	Ghi chú
ECS Fargate	\$20	2 cluster nhỏ hoạt động liên tục
ALB x2	\$15	Tính theo giờ sử dụng và lưu lượng
Route 53	\$1	1 hosted zone + Health Check
CloudWatch	\$3	Log cơ bản và cảnh báo giám sát
ECR	\$0	Miễn phí nếu dưới 500MB
Tổng cộng	~\$39	Có thể tối ưu thêm bằng cách tắt sau giờ làm

Phân tích ROI:

- ✓ Tiết kiệm **61% chi phí** so với giải pháp EC2 tự quản lý.
- ✓ Không cần đội DevOps duy trì máy chủ 24/7.
- ✓ Có thể mở rộng linh hoạt theo nhu cầu sử dụng thực tế.

“Chi phí ước tính được tính toán dựa trên mô hình demo gồm 2 ECS task nhỏ chạy liên tục trên 2 vùng, phù hợp với môi trường thử nghiệm trong giai đoạn thực tập. Khi triển khai thật, chi phí có thể tăng lên từ 180–250 USD/tháng tùy mức độ tải và mở rộng.”

Risk Assessment

Phân tích rủi ro và phương án giảm thiểu:

Nguy cơ	Mức độ	Tác động	Phương án ứng phó
ECS Task fail bất ngờ	Cao	Gián đoạn dịch vụ	CloudWatch Alarm + SNS Alert
Route 53 failover trễ	TB	Downtime vài phút	TTL ngắn (10s), kiểm thử nhiều lần
Free Tier AWS hết hạn , hết dung lượng Free	TB	Tăng chi phí bất ngờ	Theo dõi usage, tắt tài nguyên rảnh
Credential bị lộ	Cao	Mất quyền truy cập AWS	MFA + Vault lưu trữ an toàn

Expected Outcomes

Chỉ số thành công:

- ✓ Zero –Downtime(<5s chuyển)
- ✓ ECS Uptime >99.95%
- ✓ Cảnh báo CloudWatch < 30s khi có sự cố
- ✓ Route 53 failover <1 phút khi vùng chính lỗi
- ✓ Có thể mở rộng CI/CD + backup tự động
- ✓ Là nền tảng DevOps nâng cao

Lợi ích lâu dài:

- Làm nền tảng cho CI/CD tự động
- Tái sử dụng kiến trúc này cho các hệ thống khách hàng
- Sẵn sàng mở rộng sang đa vùng toàn cầu và tích hợp High Availability

References & Phụ lục

- Atlassian/Gartner: <https://www.atlassian.com/incident-management/kpis/cost-of-downtime>
- TechRepublic/Gartner: <https://www.techrepublic.com/article/85-of-organizations-will-be-cloud-first-by-2025-says-gartner/>
- AWS Pricing Calculator: <https://aws.amazon.com/vi/aws-cost-management/aws-pricing-calculator/>
- AWS ECS Multi-Region Blog: <https://aws.amazon.com/vi/blogs/devops/amazon-ecs-multi-region-deployment-with-amazon-codecatalyst/>
- RingCentral: <https://www.ringcentral.com/us/en/blog/how-much-does-an-unplanned-it-outage-cost/>

Em xin chân thành cảm ơn anh/chị trong **team JCJ** đã nhiệt tình hỗ trợ và tạo điều kiện để em có thể hoàn thiện Proposal thực tập này.

Em gửi Proposal trước thời hạn để thuận tiện cho việc xác nhận và đóng dấu. Rất mong anh/chị xem xét và hỗ trợ giúp em hoàn tất thủ tục thực tập theo quy định.

Em rất mong sẽ tiếp tục nhận được sự hướng dẫn, góp ý từ anh/chị trong suốt quá trình thực tập sắp tới.

- TRÂN TRỌNG CẢM ƠN ANH/CHỊ VÀ TEAM JCJ!