



# Production-Ready Kubernetes, CI/CD, and Hybrid Connectivity on AWS

## Executive Summary

HiCap is an AI-driven technology company building high-performance platforms that rely on scalable, containerized workloads and advanced data services. Following the successful completion of its AWS Landing Zone and security foundation (SOW 1), HiCap moved into the next phase of its cloud migration journey: deploying a production-grade Kubernetes platform on AWS and enabling secure hybrid connectivity with its existing Azure environment. To achieve this, HiCap partnered again with binbash, an AWS Advanced Consulting Partner, to design and implement a production-ready Amazon EKS environment, GitOps-based CI/CD pipelines, and a Site-to-Site VPN with AWS Transit Gateway, ensuring seamless connectivity between AWS and Azure while maintaining AWS Well-Architected best practices.

## Customer Challenge

After establishing a secure AWS foundation, HiCap faced new challenges as it prepared to run production workloads on AWS:

- The need for a robust Kubernetes orchestration platform capable of supporting production traffic with high availability and scalability.
- Hybrid networking requirements, as key components and entry points were still running in Microsoft Azure.
- A lack of standardized CI/CD pipelines and GitOps workflows to consistently build, deploy, and operate containerized applications.
- Secure access to Amazon Bedrock from Kubernetes workloads without exposing traffic to the public internet.
- Ensuring the new platform remained operationally manageable, secure, and aligned with AWS best practices.



HiCap is an AI capacity management platform that helps teams run top foundation models at lower cost and higher throughput. It provides rate-limit-free inference, instant model switching across OpenAI, Anthropic and Gemini, and reserved GPU capacity for faster performance. With drop-in SDK compatibility, clear cost visibility and strong security, HiCap simplifies scaling AI.



binbash has helped accelerate our compliance goals by building our secure AWS foundation and infrastructure automation. The team has also helped accelerate our application deployments by integrating with our teams. Thanks to binbash, we've reduced deployment times, fewer compliance issues,

**JAVIER CEVALLOS**  
CTO & CO-FOUNDER



# binbash®

HiCap engaged binbash to extend the Landing Zone into a full production platform, bridging AWS and Azure while enabling modern DevOps and platform engineering workflows.

## **Solution**

binbash designed and delivered a production-grade cloud-native platform on AWS, fully automated using Infrastructure as Code (Terraform) and accelerated by the binbash Leverage™ framework.

## **Key AWS Services and Technologies**

- Amazon EKS (Prod) with managed node groups and production-grade networking.
- AWS Transit Gateway to centralize routing and enable scalable hybrid connectivity.
- AWS Site-to-Site VPN with static routing between AWS and Azure.
- Amazon ECR as the centralized container registry.
- GitHub Actions for image build and push pipelines.
- ArgoCD and Argo Image Updater to implement GitOps-based deployments.
- AWS Bedrock VPC Endpoints for private, secure GenAI access.
- Amazon Route 53, ExternalDNS, Cert-Manager, and NGINX Ingress Controller for traffic management and TLS automation.

## **Architecture Overview:**

A hybrid AWS–Azure architecture with AWS Transit Gateway as the networking hub, a production Amazon EKS cluster for container workloads, GitOps-based CI/CD pipelines, and private Amazon Bedrock access through VPC endpoints.



# binbash®

## Key Components of the Solution

- **Hybrid Networking & Connectivity**
  - Dedicated Networking account to centralize routing and governance.
  - AWS Transit Gateway connecting Prod, Dev, Shared, and Networking VPCs.
  - Site-to-Site VPN enabling private connectivity between AWS and Azure environments.
  - Route 53 inbound/outbound resolvers to support cross-cloud DNS resolution.
- **Production Amazon EKS Platform**
  - Separate EKS clusters for Production, deployed in isolated AWS accounts.
  - Managed node groups with on-demand instances for production stability.
  - Kubernetes isolation via namespaces, RBAC, and network policies.
- **Cluster Services & Platform Tooling**
  - AWS Load Balancer Controller and NGINX Ingress.
  - ExternalDNS integrated with Route 53 (public and private zones).
  - Cert-Manager with Let's Encrypt for automated TLS.
  - Cluster Autoscaler for elastic scaling.
  - External Secrets integrated with AWS Secrets Manager.
- **CI/CD & GitOps**
  - GitHub Actions pipelines to build and push container images to Amazon ECR.



- ArgoCD for declarative, GitOps-based deployments.
- Argo Image Updater to automate image version updates.
- Clear separation between build and deploy responsibilities.
- **GenAI Enablement**
  - Private Amazon Bedrock VPC endpoints for secure inference from EKS workloads.
  - Consideration of cross-region inference to handle traffic bursts and improve availability.

## Results

The second phase of the engagement enabled HiCap to:

- Run production containerized workloads on AWS with high availability and scalability.
- Seamlessly connect AWS and Azure environments, enabling a phased migration strategy.
- Standardize application delivery through CI/CD and GitOps workflows.
- Secure GenAI access to Amazon Bedrock without public internet exposure.
- Reduce operational complexity through repeatable, automated infrastructure patterns.

## Key Milestones

- **Hybrid Networking Foundation**
  - Created a dedicated networking account.
  - Deployed AWS Transit Gateway and Site-to-Site VPN to Azure.
- **Production EKS Deployment**
  - Provisioned production Amazon EKS cluster and networking.
  - Installed core Kubernetes and platform components.
- **CI/CD Enablement**
  - Implemented GitHub Actions pipelines to build and push images to Amazon ECR.
  - Established GitOps deployment workflows using ArgoCD.
- **GenAI Private Connectivity**
  - Deployed Amazon Bedrock VPC endpoints.
  - Enabled private inference access from Kubernetes workloads.
- **Knowledge Transfer & Enablement**
  - Delivered demos, documentation, and best practices.
  - Enabled HiCap's team to independently operate and extend the platform.



# binbash®

## Metrics

- Time to production platform: ~2 months end-to-end.
- Infrastructure automation: ~100% of EKS and networking resources deployed via Terraform.
- Deployment consistency: 100% of production deployments managed via GitOps.
- Hybrid connectivity: Private, low-latency AWS–Azure communication established.
- Security posture: No public exposure of Kubernetes APIs or Bedrock endpoints.

## Conclusion

With the successful completion of SOW 2, HiCap transitioned from a secure cloud foundation to a fully operational, production-grade Kubernetes platform on AWS. By combining Amazon EKS, hybrid networking, GitOps-based CI/CD, and private GenAI access, binbash enabled HiCap to scale its platform while maintaining security, reliability, and operational excellence.

Built on the AWS Well-Architected Framework and accelerated by binbash Leverage™, this second phase positioned HiCap to confidently run production workloads, evolve its hybrid architecture, and continue its cloud-native transformation with AWS at the core.