

>> CASE STUDY

Augury runs Kubernetes and AI production workloads on Spot VMs and saves 80%

With Flexera Ocean, Augury cut costs while maintaining high availability for Azure Kubernetes Service (AKS) and AI workloads running on Azure Spot VMs

>> INTRODUCTION

Augury strategically adopted cloud infrastructure with container architecture; they chose Microsoft Azure as their primary cloud provider and leverage Azure Kubernetes Service (AKS). This move positioned them at the forefront of cloud innovation, ensuring continued operational efficiency and scalability.

The challenge

Augury's AKS workloads were spread across three Azure accounts: development and AI, staging, and production. The staging and production environments each had a single cluster with several thousand pods, while the development and AI account maintained four clusters, with two dedicated to AI workloads.

Initially, Augury's DevOps team handled this complex infrastructure internally. However, facing the need to reduce operational complexity, streamline cloud costs, and improve gross margins, Guy Carmy, Augury's DevOps lead, sought an optimization solution for their AKS infrastructure. "The need to cut our cloud costs became more pressing over the years," said Carmy.

**>> AT A GLANCE**

Industry: Technology

Location: New York, NY

Product: Flexera Ocean

>> ABOUT AUGURY

Augury, a leader in machine health and process health solutions, uses purpose-built AI technology, trained by industry experts and the world's largest data library, to help manufacturing and industrial companies eliminate production downtime, improve process efficiency, maximize yield and reduce waste and emissions.

Prior to onboarding Ocean, Augury had been using Azure Spot Virtual Machines (VMs) via Terraform. However, Azure's autoscaler proved too rigid, limiting node pools to a single virtual machine (VM) size, and yielding minimal cost savings.

Meeting the rigorous standards for partnership

Flexera Ocean emerged as the ideal choice for Augury, but first, Ocean was rigorously evaluated to meet Augury's criteria for third-party solutions:

1. **Tangible and measurable value:** Ocean's ability to automatically optimize cloud usage and report cost savings aligned perfectly with Augury's goals
2. **Integration with existing workflows:** Augury values smooth integration with their existing tooling, including Terraform and Helm charts. Ocean's compatibility with these tools ensured a smooth transition for Augury's DevOps team
3. **Compliance with CISO-defined security standards:** Augury prioritizes data security and requires any solution, including Ocean, to adhere to their CISO-defined security standards
4. **Adherence to corporate procurement guidelines:** Augury insists on partnering with reputable enterprise-grade vendors. Ocean's established reputation made it an ideal choice

Ocean's enterprise-grade SaaS offering easily met these requirements, securing its position as Augury's preferred solution.

The solution

Following a successful Proof of Concept (POC), Augury migrated their staging and their development and AI environments to Virtual Node Groups (VNGs) managed by Ocean. Subsequently, production clusters were also migrated, achieving nearly 100% Spot VM coverage.

Ocean's intelligent automation dynamically scales infrastructure based on real-time analysis and workload demands of AKS clusters, continuously optimizing performance, efficiency, and cost.

With Ocean, Augury achieved significant cost savings and simplified their AKS infrastructure through key features:

- **Bin packing:** Condensing workloads onto fewer, more utilized nodes reduced complexity and improved Augury's cost efficiency
- **Rightsizing:** Following Ocean's recommendations, node sizes were adjusted to actual workload demands, allowing Augury to operate more efficiently on smaller, cost-effective nodes
- **Heterogenous VNGs:** Ocean's ability to support diverse machine types within a single node pool minimized the number of pools Augury needed to manage

“We quickly migrated our production to Ocean-run spot VMs,” said Carmy. “Ocean handles them efficiently, ensuring our production availability remains unaffected.”

Guy Carmy
DevOps Lead
Augury

The results

Over a 12-month period, Augury saved a staggering **80% compared to equivalent on-demand costs**. Further, Ocean's reliability has enabled Augury to optimize its internal developer platform. All environments autonomously spun up by Augury's developers are now hosted on Ocean-managed AKS clusters, benefiting from efficient computing and resource-effective autoscaling.

"Ocean's VNG flexibility delivered near-instant cost and time savings," said Carmy. "We drastically simplified our AKS infrastructure—from requiring unique Azure node pools for each workload type to consolidating multiple workloads with varying needs to run on the same Ocean VNG. Thanks to Ocean's VNGs accommodating heterogeneity, managing our infrastructure became a 'set it and forget it' process. Upon implementing Ocean, we set up two VNGs that have hosted all our workloads ever since."

Ocean ensured **uninterrupted operations**, supported by its **seamless integration with AKS** that enhanced control over Kubernetes operations. By automating key processes and providing actionable insights, Ocean helped Augury to achieve greater operational efficiency and enabled their DevOps team to prioritize pioneering digital innovations for competitive advantage.

“

Thanks to Ocean's VNGs accommodating heterogeneity, managing our infrastructure became a 'set it and forget it' process. Upon implementing Ocean, we set up two VNGs that have hosted all our workloads ever since.”

Guy Carmy
DevOps Lead
Augury



Looking ahead

Augury is interested in the development of Ocean's automatic rightsizing for AKS, a groundbreaking feature set to transform its cloud management strategy. This unique capability will facilitate continuous rightsizing in the background through automation policies, ensuring optimal resource allocation without manual intervention. [Note: Ocean [automatic rightsizing](#) was released in June 2025.]

“It would be deeply disappointing if we lost Ocean. It plays a crucial role in cost management without compromising performance or scalability. Without it, optimizing costs would be a lot more challenging, and the results might not measure.”

Guy Carmy
DevOps Lead
Augury

About Flexera

Flexera helps organizations understand and maximize the value of their technology, including the rising costs and risks introduced by AI, saving billions of dollars in wasted spend. Our Flexera One platform connects the dots between what technology you have, how it is used, what it costs, and where it creates risk, helping teams take control of the increasingly complex IT estate across cloud, SaaS and on-premises. We are leading the way to unify IT asset management, FinOps and SaaS management with high fidelity data from Technopedia, our proprietary reference library of technology asset data, and intelligent automation fueled by AI. That's why thousands of global organizations rely on the Flexera One platform and Technopedia. Learn more at flexera.com

» NEXT STEPS

Run a Kubernetes workload analysis to uncover your savings opportunities with Flexera Ocean

[Get your savings analysis →](#)