



# Launch automated fleet operations at scale with Rafay

Simplify the management of your Kubernetes  
environment on AWS

In collaboration with





# Contents

|                                                                                                      |   |
|------------------------------------------------------------------------------------------------------|---|
| The rise of containers and container management.....                                                 | 3 |
| The challenges of fleet operations .....                                                             | 4 |
| Rafay Automated Fleet Operations: Simplified, streamlined<br>management of Amazon EKS clusters ..... | 5 |
| Automated Fleet Operations use cases.....                                                            | 6 |
| Use case 1: Streamlining cluster upgrades.....                                                       | 6 |
| Use case 2: Deployment safety .....                                                                  | 7 |
| Use case 3: Governance at scale .....                                                                | 7 |
| Use case 4: Faster security response.....                                                            | 8 |
| Use case 5: AI/HPC performance consistency.....                                                      | 8 |
| Conclusion.....                                                                                      | 9 |



## The rise of containers and container management

Rapid deployment, scalability, consistency, and support for DevOps automation are just some of the ways containers help modern engineering teams and organizations meet today's fast-paced software development and delivery cycles. Gartner<sup>1</sup> predicts that by 2028:

**95+%**

of global organizations will be running containerized applications in production

**25%**

of all enterprise applications will run in containers

This rapid expansion of container use is expected to drive further investment in Kubernetes clusters. According to a Cloud Native Computing Foundation survey<sup>2</sup>, 96 percent of organizations are either using or evaluating Kubernetes as their container orchestration platform. While Kubernetes makes containers easier to deploy, manage, and scale, organizations often look to managed services solutions to help them extract the greatest value from Kubernetes as their orchestration platform.

### Fully managed Kubernetes with Amazon EKS

Many organizations choose to run Kubernetes on [Amazon Elastic Kubernetes Service](#) (Amazon EKS). Amazon EKS is a managed Kubernetes service for orchestrating and managing containerized applications on Amazon Web Services (AWS) and on-premises. Amazon EKS makes it easy for organizations to run Kubernetes without installing and operating their own control plane and worker nodes. It provides a managed, auto-scale Kubernetes control plane and integration to AWS networking and security services.

Organizations typically manage sets of Amazon EKS clusters as a **fleet**. Depending on how the organization is structured and operates, the fleet could consist of groups of Amazon EKS clusters for development, test, and production. Or it might include Amazon EKS clusters running on AWS alongside Kubernetes deployments on-premises, or at edge locations such as factories, stores, or distribution centers. Or, clusters could be grouped because of the need to apply similar operations, monitoring, and governance. **Fleet operations** is the process for managing, monitoring, and governing a heterogeneous fleet of Amazon EKS clusters.

<sup>1</sup>Gartner. [A CTO's Guide to Navigating the Cloud-Native Container Ecosystem](#). March 31, 2023.

<sup>2</sup>Cloud Native Computing Foundation. [Annual Survey 2021](#). February 10, 2022.

# The challenges of fleet operations

When done well, fleet operations ensures uniformity, scalability, and high availability of Amazon EKS clusters. However, despite their numerous advantages, current fleet-based operation practices have their own challenges, even in relatively small environments:



**Limited tooling:** One of the most pressing issues in fleet operations is the lack of robust, dedicated tooling. While Kubernetes itself provides tools for managing individual clusters, many organizations find it difficult to extend these tools to manage fleets in a streamlined and automated manner.



**Complexity in multi-cluster management:** Managing more than one Kubernetes cluster can be challenging due to configuration, policy, and security differences across clusters and environments. This complexity can often lead to inconsistencies and operational challenges, impacting the efficiency and reliability of applications.



**Security and compliance concerns:** As organizations scale up their Kubernetes deployments, ensuring security and compliance across all clusters becomes increasingly difficult. Centralized security policy enforcement across diverse clusters and environments can be particularly complex at scale.



**Manual pre-op and post-op checks:** Another crucial gap concerns the end-to-end workflow process of complex operations. This generally involves a set of prerequisites, such as safety checks or workload offloading, followed by a set of core operations on Kubernetes resources, followed finally by a set of post operations to either validate or bring workloads back to production. The lack of an automated end-to-end workflow hampers an organization's ability to build a full picture of its operations and can have a significant impact if not done correctly.



**Resource orchestration:** Fleet-based operations that require broad resources, such as Kubernetes upgrades, should be staged across the environment to maintain stability and availability for developers and business functions. Manual staging introduces inefficiencies and slows the overall process.

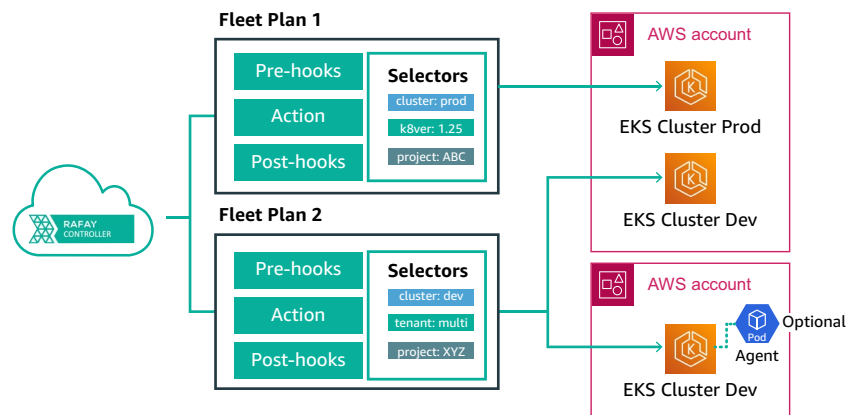
Rafay Systems' Cloud Automation Platform, a solution for platform teams that wish to build automated self-service cloud infrastructure workflows, helps address the challenges of managing a fleet of Kubernetes clusters spanning various environments, data centers, and accounts with [Automated Fleet Operations](#). It's a solution that solves, at scale, the challenges associated with multi-cluster management.

# Rafay Automated Fleet Operations: Simplified, streamlined management of Amazon EKS clusters

Building on its expertise as a leading provider of automations for Kubernetes management and operations, Rafay developed Automated Fleet Operations to simplify the way organizations manage Kubernetes cluster environments. Automated Fleet Operations eliminates the error-prone, manual processes of traditional Kubernetes cluster fleet operations by allowing platform and operations teams to implement repeatable, automated workflows that can be applied to multiple clusters. This allows them to increase productivity and improve the reliability of lifecycle management operations performed on Kubernetes clusters fleet-wide.

Rafay Automated Fleet Operations consists of four components:

- **Fleet plan:** A collection of operations that can be used to apply comprehensive tasks such as upgrades, backups, patching, and scaling in/out.
- **Operations:** A set of pre-hooks, an action, and a set of post-hooks. Hooks are executed as a pod by either cluster type runner or agent type runner.
- **Fleet:** A collection of targeted Amazon EKS clusters for the operations, grouped by project and label.
- **Agent:** A container that runs on a selected cluster to execute the hooks.



## Benefits of Automated Fleet Operations

With Rafay Automated Fleet Operations, users can address the complexities and challenges described previously and gain valuable benefits, including:



Enhanced visibility through the centralized management of clusters



Improved productivity and operational efficiency



Better compliance with organizational governance and security policies



Increased consistency of cluster operations applied across the fleet



Higher availability of key resources for developers

# Automated Fleet Operations

## use cases

Fleet operations plays an important role when it comes to streamlining, managing, monitoring, and governing common operational workflows on Amazon EKS. The following use cases show the benefit of using Rafay Automated Fleet Operations for multi-cluster fleet management.



### Use case 1: Streamlining cluster upgrades

At least three times a year—or sometimes even more frequently—new versions of Amazon EKS are released in line with AWS's commitment to remain current with the pace of Kubernetes development. When this happens, older versions of Amazon EKS fall out of support. Whether your organization has hundreds of clusters or just a few, they all need to be upgraded.

However, cluster upgrades aren't simple, and there's often a list of things IT operations teams must check before executing an upgrade. Because clusters have live workloads running on them, there are certain dependencies that must be considered. Kubernetes also deprecates APIs at regular intervals, impacting the clusters or workloads that use them—something else that must be accounted for in the upgrade plan. It's no wonder that platform and operations teams often must plan for weeks before executing an Amazon EKS upgrade across their environment.

With Automated Fleet Operations, teams can create customizable, highly tuned upgrade plans that make it possible to upgrade in days rather than weeks with zero downtime and no impact to applications or users. Automated Fleet Operations automates pre-op checks for things like deprecated APIs or other dependencies, making sure your cluster is ready to upgrade. When you're ready, you can use Automated Fleet Operations to orchestrate the automation of upgrades for an entire cluster fleet or just certain portions. Once the upgrade is finished, Automated Fleet Operations performs a post-check, validating that the upgrade was successful or that your cluster resources are back online before you allow traffic to them.



## Use case 2: Deployment safety

Large enterprises often have central IT teams that are responsible for the underlying infrastructure of Amazon EKS clusters, but that lack control over the applications or workloads running on those clusters.

When cluster upgrades or changes must be made, that's the responsibility of the central IT team. But first, that team has to work with each individual application or workload team on every issue that is identified as a pre-check and then solve that issue before the upgrade can happen.

With Automated Fleet Operations, you can perform an upgrade "dry run" with multiple sets of validations. The dry run creates a holistic report that identifies any issues that must be resolved before an upgrade can happen. Based on that report, you can tell the application and workload teams what they need to fix. By the time you're ready to perform the upgrade enterprise-wide, all issues that could affect deployment safety will have been resolved, paving the way to a successful upgrade.



## Use case 3: Governance at scale

Organizations often have stringent policies in place that demand thorough checks on specific configurations of Amazon EKS. For instance, many mandates may strictly prohibit any form of public access to the cluster. When you're operating at a large scale, ensuring that every component and configuration adheres to these stipulations can be a daunting task. Relying on a fragmented, manual process to enforce these policies can lead to oversight, errors, and inconsistencies, making it an untenable solution for a growing enterprise.

Automated Fleet Operations is designed to take the hassle and risk out of governance by offering automated configuration checks and robust policy enforcement across the entire organization. Instead of dedicating precious resources to perform these repetitive tasks, which are often prone to human error, Automated Fleet Operations ensures that configurations and policies are consistently upheld to the highest standard.

In instances where configurations do not align with organizational policies, there might be a need for corrective action or remediation. Manual interventions can be both slow and error-prone, especially when they must be executed on a vast scale. Automated Fleet Operations can identify any issues and, if needed, initiate and execute remediations across the fleet, ensuring that the organization's infrastructure remains compliant and secure at all times. This not only saves time but also significantly reduces the potential for errors, making governance more streamlined and reliable.



## Use case 4: Faster security response

Security patch application is essential to rectify vulnerabilities and maintain the integrity of environments and softwares. However, their effectiveness is contingent upon their actual implementation. When dealing with applications scattered across numerous Amazon EKS clusters, it can be daunting to swiftly and efficiently disseminate patches that address urgent and critical security loopholes.

Automated Fleet Operations is designed to manage the intricacies of patch deployment from end to end, ensuring that your applications remain protected at all times. Its functionalities include:

- **Pre-op code execution:** Before the actual deployment begins, Automated Fleet Operations initiates a pre-op code. This process intelligently targets the appropriate deployment areas across multiple clusters, ensuring that no node is left vulnerable.
- **Actual patch deployment:** This is the heart of the operation. The system ensures that the patches are not just deployed but also integrated seamlessly, without causing disruptions or inconsistencies.
- **Post-op code validation:** Post-deployment, it is crucial to validate that the patches have been correctly implemented and that there are no lingering vulnerabilities. Automated Fleet Operations has a built-in post-op code execution phase that rigorously checks for the deployment's success.

With Automated Fleet Operations, what traditionally could take days, fraught with potential errors and oversights, is now accomplished in only minutes. This swift turnaround is not just about speed but also about securing applications and data promptly, reducing the window of potential exploitation.



## Use case 5: AI/HPC performance consistency

Many organizations are looking to take advantage of the computational power of graphical processing units (GPUs) running on Amazon EKS for use cases such as training artificial intelligence (AI) models and enabling high performance computing (HPC).

Before deploying an application or upgrade to GPU-based clusters or nodes, engineering teams typically run a pre-hook to calculate current GPU performance. If benchmarks are within limits, the team executes the upgrade and then runs a post-check to ensure GPU performance remains within the previously established benchmarks.

Automated Fleet Operations automates this process, streamlining the entire performance validation framework and ensuring AI and HPC workloads can run as expected.

# Conclusion

Rafay Automated Fleet Operations simplifies the management of Amazon EKS clusters, improving visibility, driving operational efficiency, and making it easier to achieve compliance with organizational governance policies—even in relatively small environments. Automated Fleet Operations is available now for customers using Rafay to manage their Amazon EKS fleets.

**Try Automated Fleet Operations today.**

Find Rafay in the AWS Marketplace ›

## About Rafay Systems

Every modernizing enterprise has complicated cloud infrastructure, and they're all building automation to hide that complexity from users. But they shouldn't have to do it all.

Rafay's Cloud Automation Platform provides a solution for platform teams that wish to build automated self-service cloud infrastructure workflows that bridge the complexity gap, so your business can focus on rapid innovation rather than cloud management. Rafay's solution allows platform teams to enable anyone who depends on rapid access to cloud infrastructure to move faster safely, which is why Gartner has recognized Rafay as a Cool Vendor in Container Management in 2023.

With our solution, platform teams at MoneyGram, GuardantHealth, Verizon, and many other companies deliver autonomy to cloud users (developers, data scientists, researchers, and more) while maintaining control and efficiency over cloud operations.

For more information, please visit [www.rafay.co](https://www.rafay.co).

In collaboration with

