Kubernetes cost allocation: finding which team spent what

26 October 2026 · 5 min read · blogs.iictp.org/kubernetes-cost-allocation

A Kubernetes cluster is a shared pool, so the cloud bill does not say which team or product caused it. Allocation means attributing the cluster's cost back to its users in a way people accept.

TL;DR The short version

- A Kubernetes cluster is a shared pool, so its bill must be attributed back to the teams and products that use it.
- Label every workload with team, service and environment, enforce it at admission, and attribute cost by resource requests.
- Show each team its own cost and flag the largest gaps between requested and actually used resources.

Read online, listen to it, or change the text size: https://blogs.iictp.org/kubernetes-cost-allocation

The basics

- Label every workload with team, service and environment, and enforce it at admission.
- Attribute cost by the resources each pod requests, and compare it to what it actually uses.
- Share the cost of common services, such as the control plane and ingress, by an agreed rule.

The pitfalls

- Unlabelled workloads pile up in an "unallocated" bucket.
- Requests set far above real use waste capacity for everyone.
- Idle capacity has to belong to someone.

Use the numbers

1. Show each team its own cost.
2. Flag the largest gaps between requested and used resources.
3. Set requests from measured use.
4. Review after each change.

Related credential: Cloud Architecture Practitioner. Explore the exam facts and syllabus.

Sources and further reading

1. Kubernetes documentation. Resource management for pods and containers. kubernetes.io/docs.

Published by IICTP, an independent technical credential body, on 26 October 2026. The current version of this post is at https://blogs.iictp.org/kubernetes-cost-allocation. More guides: blogs.iictp.org.