Avesha's technologies are the **next generation innovation** abstracting the network and infrastructure for application needs.

Avesha

KubeTally

Tracking Chargeback Amounts per Team in Kubernetes Environments

In Kubernetes, the ability to accurately attribute resource consumption to the appropriate business unit or team is essential for operational efficiency and cost management. The existing toolset available to tally costs per team (like Kubecost) lacks the capability to effectively isolate, total and attribute these costs across a business unit's multi-cluster deployments. **Kube Tally** is the only offering in the market, that has advanced multi-cluster capabilities to accurately aggregate and attribute chargeback amounts for multi-cluster Kubernetes environments.

The Challenge of Chargeback in Kubernetes

Kubernetes environments, especially those that extend across multiple clusters, present challenges in accurately tracking and attributing resource usage. Business units or application teams within an organization often share resources, complicating the process of accurately charging back the costs incurred by each. The difficulty in isolating and tracking usage for namespaces per business unit further complicates this process, leading to inefficiencies and disputes over resource consumption charges.

Chargeback per Team Example

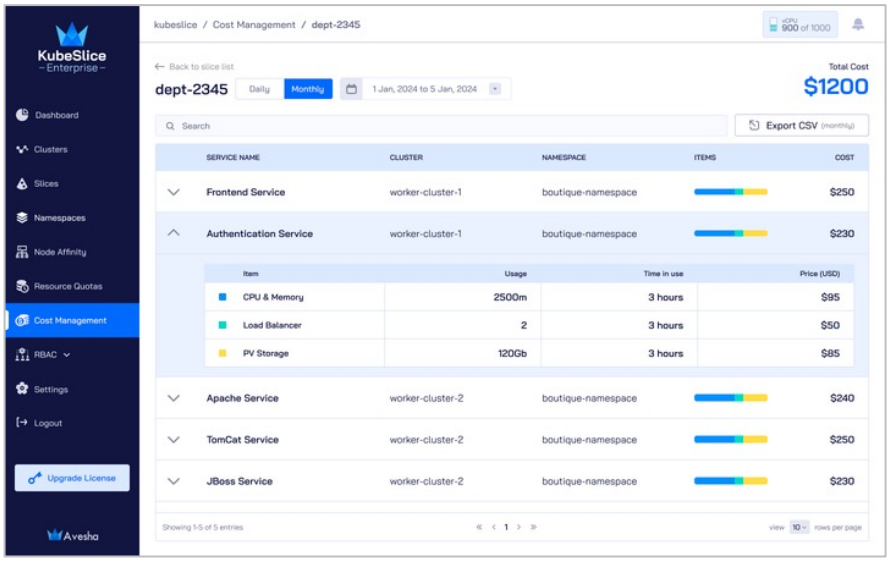
Consider a scenario involving two teams: the "Boutique Front End Team" and the "Boutique Back End Team." Both teams share two clusters: "Boutique1" and "Boutique2."

The platform engineering team wants to calculate charges for "Boutique Front End Team" separately from "Boutique Back End Team" although they share the two clusters. Typically a report would look like this:

Team	Cluster 1	Cluster 2	Total
Boutique Front End Team	Cpu = $500 Memory = $600 LoadBalancer = $150 Storage = $200	Cpu = $400 Memory = $600 LoadBalancer = $100 Storage = $200	$2750
Boutique Front End Team			

How Kube Tally aggregates chargeback per team

The Platform team can set up two Slices, one for each team. Kube Tally calculates the costs related to PV, Load Balancer, CPU, and Memory for the set of namespaces within each Slice. Thanks to the complete isolation of the two Slices (and thus their namespaces), Kube Tally can accurately extract individual resource amounts for every namespace per Team, across the entire multi-cluster deployment. The key features of the underlying KubeSlice technology that facilitates multi-cluster chargeback include:



Namespace Sameness Across Clusters: KubeSlice ensures namespace sameness across multiple clusters, allowing the same namespace, in different clusters, to be added to the Slice with ease. This makes onboarding namespaces and applications easy across multi-cluster

Resource Utilization Dashboard: The KubeSlice dashboard shows details on resource utilization across a multi-cluster deployment. It shows resource metrics such as the number of pods, CPU, memory, and ephemeral storage for each Team, over a period of time.

Manage K8s Resources using UI: KubeSlice's "Manage Resources" feature allows teams to define resource limits per Slice, ensuring “requests” and “limits” for every namespace or at an aggregated Team level. Setting these limits using K8s native commands is cumbersome. The KubeSlice UI thus makes these parameters easily accessible using the UI. These limits, when set responsibly will have a direct impact on enabling teams to not exceed their cloud costs.

Node Utilization Insights: The node utilization feature in KubeSlice provides a detailed view of how resources are used at the node level for multi-cluster environments, aiding in optimizing resource allocations within a cluster.

Conclusion

Kube Tally is a streamlined approach to managing Kubernetes resources and implementing effective chargeback mechanisms. Its ability to consolidate usage data per namespace across multiple clusters, coupled with in-depth insights into resource and node utilization, empowers organizations to accurately allocate costs to the correct business units. This fosters fairness and transparency in chargeback practices, significantly enhancing operational efficiency in the management of Kubernetes environments.