

Compute Cluster

Harnessing Infinite Computing Potential with
ConverSight.





Table of Contents

1. Introduction	02
2. Cluster Building with ConverSight	02
2.1 Cluster Classification	04
2.2 Worker Type	05
2.3 Instance Type	06
2.4 Nodes	06
2.5 Subscription level of Cluster	07
2.6 Environment	09
2.7 Cluster Culling	10
2.8 Cluster Initiation with Athena	10
3. Key features of Compute Cluster	12
4. Conclusion	13

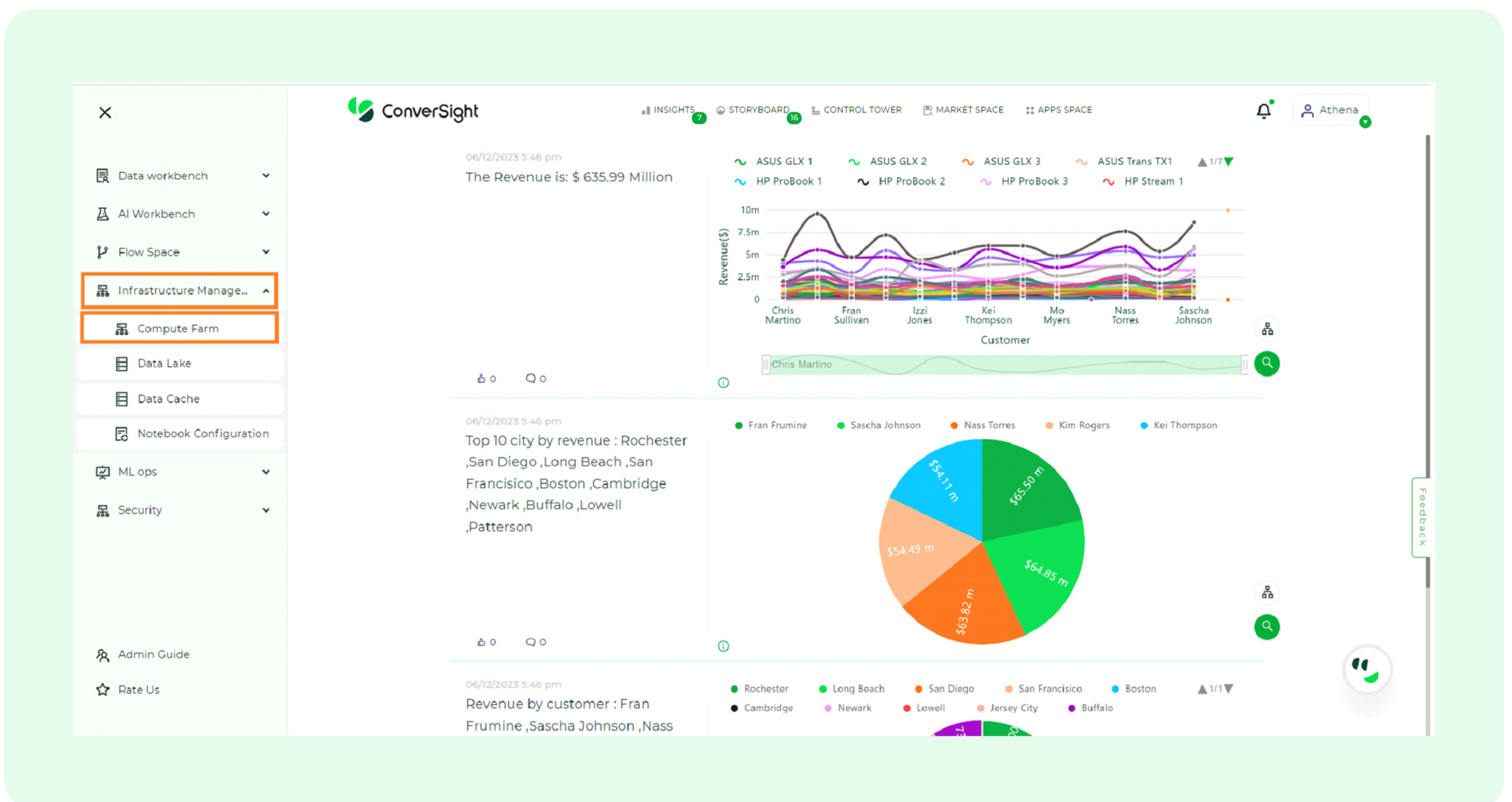


1. Introduction

In the data-driven landscape of today's business world, the significance of computational power cannot be underestimated. It serves as a driving force for innovation, efficiency and competitiveness. Recognizing the vital role of computational capabilities, the ConverSight platform provides a key feature – the ability for users to create Compute Clusters under the Compute Farm. These Clusters are essential assets that fuel research and innovation by harnessing interconnected nodes for parallel processing. This dramatically reduces computation time for complex tasks and vast datasets, enabling businesses to make swift, data-driven decisions. Clusters are designed to collaborate as a unified system, sharing resources and distributing computational tasks efficiently among nodes. This optimized utilization of computational power results in enhanced performance and increased efficiency for resource-intensive operations. With the ConverSight platform, users have the flexibility to create Clusters tailored to their specific processing and analytical needs. Whether it's high-performance computing for data-intensive operations or low-level analytics for efficient data processing, ConverSight offers a wide range of customization options. This ensures that each Cluster is perfectly aligned with the unique demands of the business, empowering organizations to achieve their goals with precision and speed.

2. Cluster Building with ConverSight

To create a cluster within the ConverSight platform, navigate to the configuration panel and choose the **'Compute Farm'** option located in the **'Infrastructure Management'** menu.



On selecting the **'Compute Farm'** option, you will be directed to the Compute Farm page. Click on the **'Create'** button.

The screenshot shows the ConverSight dashboard with a table of subscriptions. The 'Create' button is highlighted in a red box. The table has the following data:

SUBSCRIPTION NAME	TYPE	INSTANCE TYPE	MINIMUM NODE	MAXIMUM NODE	MAXIMUM CLUSTERS	PENDING CLUSTERS	RUNNING CLUSTERS	SUBSCRIBED AT	UPDATED AT	ACTION
Sample	Standard Compute Cluster	OnDemand	1	2	2	0	0	Feb-08-2024 05:31:22 pm	Feb-13-2024 04:17:37 pm	

In Create New, provide the cluster details starting with the cluster's name.

The screenshot shows the 'Create New' modal open over the subscription table. The modal contains the following fields:

- * Name:
- * subscribedResourceTypeName:
- * Instance Type:
- * Minimum Node:
- * Maximum Node:
- * Maximum Clusters:
- * labels:
- * Subscription Level:

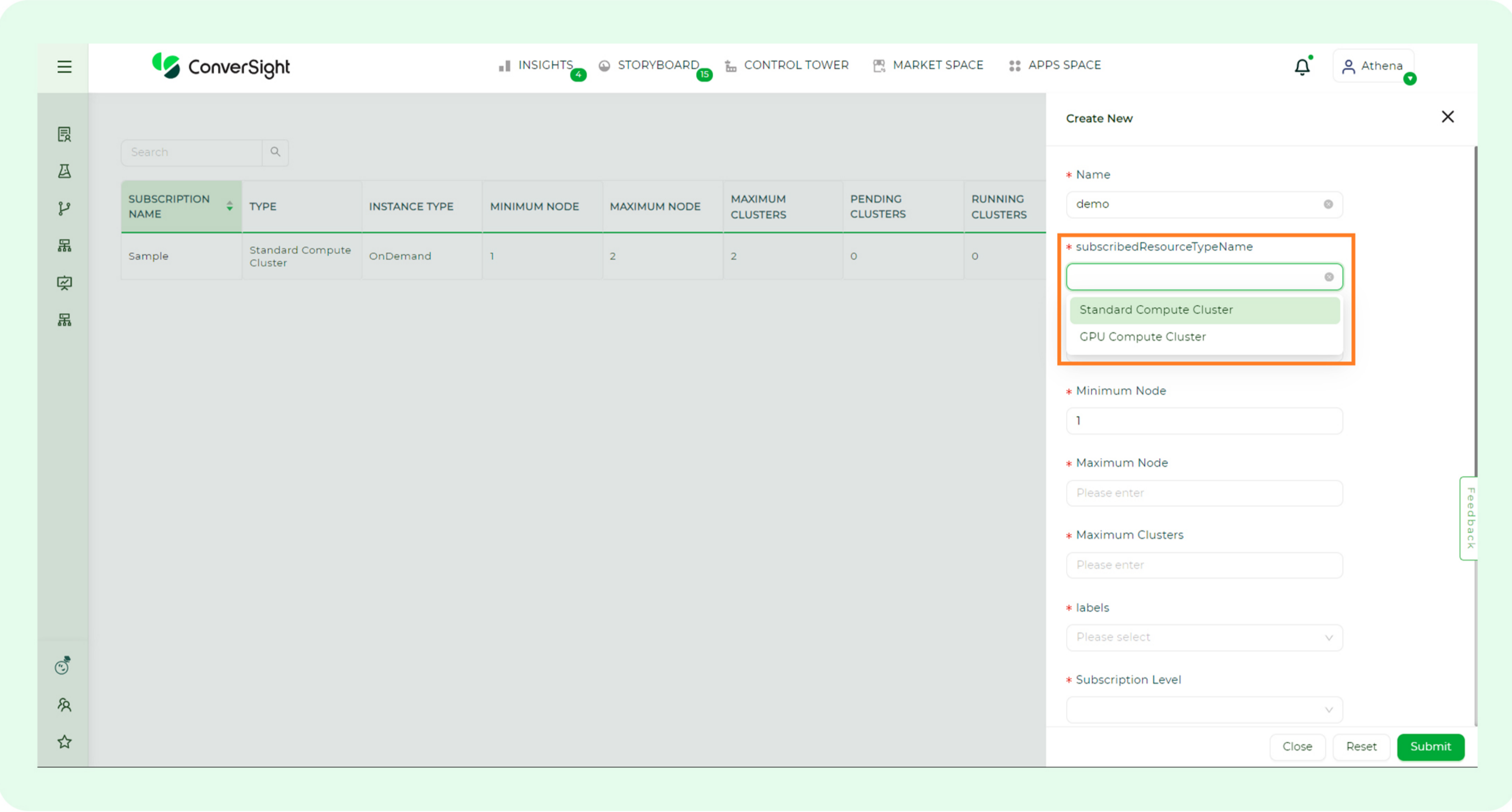
Buttons: Close, Reset, Submit



2.1 Cluster Classification

Within the ConverSight platform, there are two distinct types of clusters available: the Standard Compute Cluster and the GPU Compute Cluster. These cluster types offer different capabilities and functionalities to suit various requirements.

Additionally, ConverSight allows for further customization of the cluster type based on specific needs. This flexibility ensures that users can tailor the cluster configuration to best align with their computational demands, enabling optimal performance and resource allocation within the platform.



Standard Compute Cluster

The **Standard Compute Cluster** is a widely used cluster type within the ConverSight platform. The performance of the cluster lies in the underlying hardware utilized by each cluster type. The Standard Compute Cluster primarily relies on CPU (Central Processing Unit) compute resources to carry out its operations.

GPU Compute Cluster

The **GPU Compute Cluster** harnesses the power of GPU (Graphics Processing Unit) compute resources, which excel in parallel processing and handling intricate graphics-related calculations. GPUs are specifically designed to handle substantial volumes of data simultaneously, making them ideal for computationally intensive tasks that require significant high processing capabilities.



2.2 Worker Type

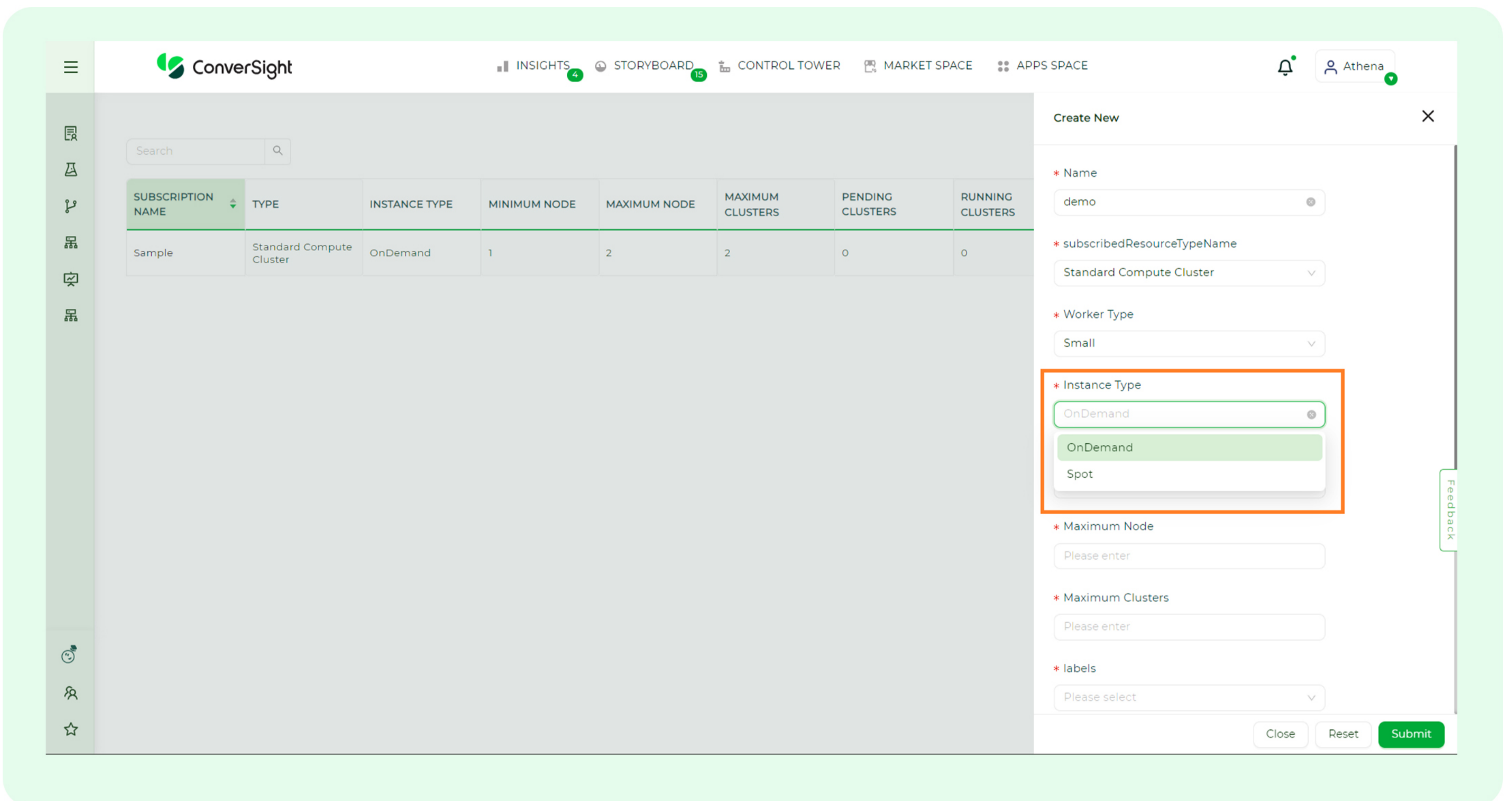
A worker is a component of an **n-node cluster**, where n denotes a specific number of nodes collaborating within the cluster. ConverSight provides users with the flexibility to allocate the capacity of the node according to their specific needs. Within ConverSight, three types of capacity options are available: Large, Medium and Small. This enables businesses to efficiently utilize their infrastructure, ensuring it is appropriately sized to handle their workloads and deliver optimal performance.

Worker Type	CPU	Capacity
Small	2	4GB
Medium	4	8GB
Large	8	16GB

The screenshot displays the 'Create New' form in the ConverSight interface. The 'Worker Type' dropdown menu is highlighted with an orange box, showing the following options: Small, Medium, Large, and Extra Large. The background shows a table with the following columns: SUBSCRIPTION NAME, TYPE, INSTANCE TYPE, MINIMUM NODE, MAXIMUM NODE, MAXIMUM CLUSTERS, PENDING CLUSTERS, and RUNNING CLUSTERS. The table contains one row with the following data: Sample, Standard Compute Cluster, OnDemand, 1, 2, 2, 0, 0.

2.3 Instance Type

Users have the flexibility to choose between On-Demand instances and Spot instances according to the specific needs of their applications. On-Demand instances are particularly beneficial for applications that require continuous availability without any interruptions. This makes them ideal for workloads where predictability in resource allocation is crucial. Conversely, Spot instances offer a different advantage, catering to workloads that are more adaptable and can tolerate interruptions. Spot instances allow users to bid on unused computing capacity, often at significantly lower prices compared to On-Demand instances. This makes them a cost-effective option for workloads that can handle occasional interruptions or fluctuations in resource availability.



The screenshot displays the ConverSight interface. On the left, a sidebar contains navigation icons. The main area features a table with the following data:

SUBSCRIPTION NAME	TYPE	INSTANCE TYPE	MINIMUM NODE	MAXIMUM NODE	MAXIMUM CLUSTERS	PENDING CLUSTERS	RUNNING CLUSTERS
Sample	Standard Compute Cluster	OnDemand	1	2	2	0	0

On the right, a 'Create New' modal is open. The 'Instance Type' dropdown menu is highlighted with an orange border, showing the following options:

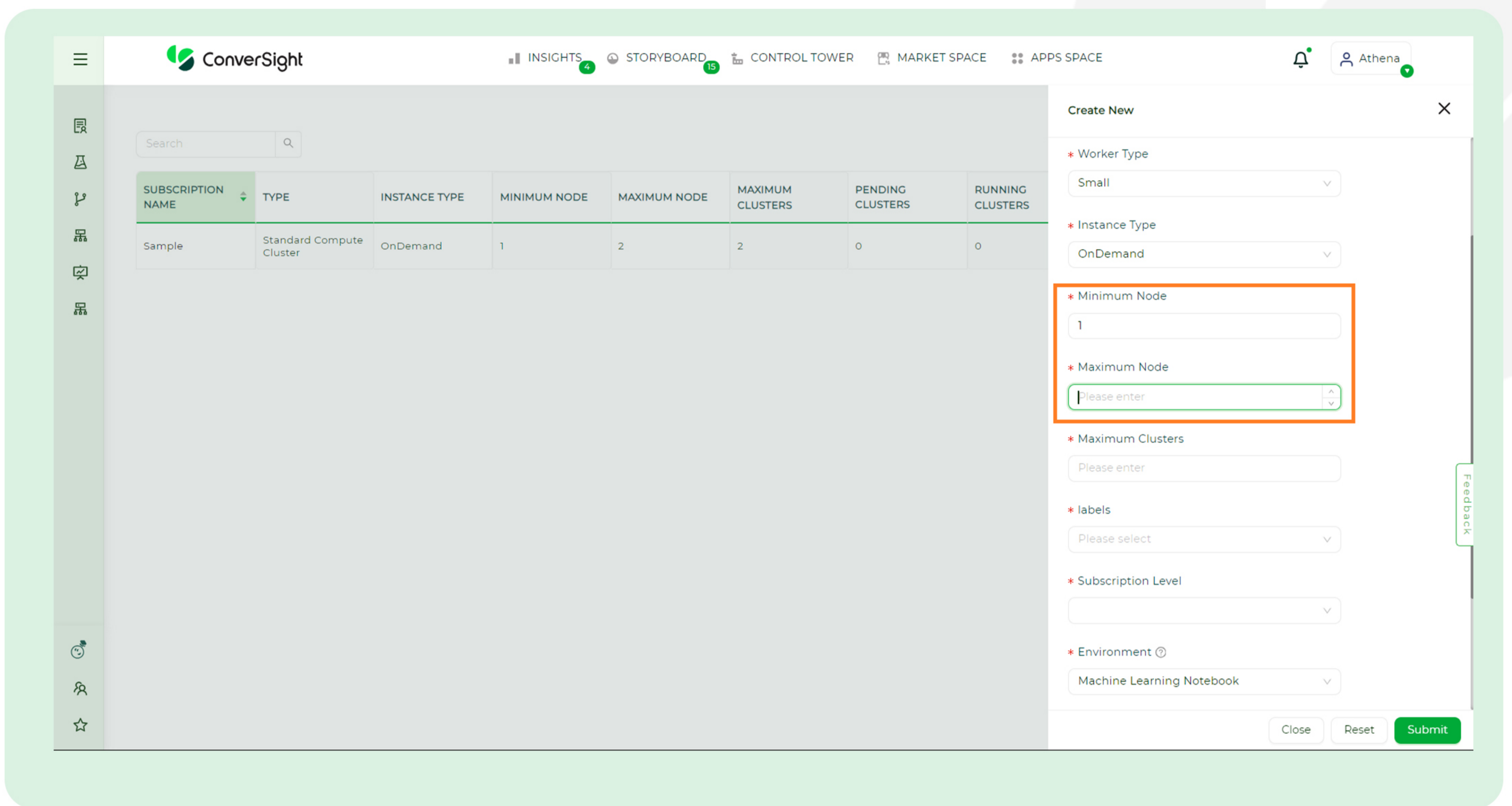
- OnDemand (selected)
- OnDemand
- Spot

Other fields in the modal include: Name (demo), subscribedResourceTypeName (Standard Compute Cluster), Worker Type (Small), Maximum Node (Please enter), Maximum Clusters (Please enter), and Labels (Please select). Buttons for Close, Reset, and Submit are at the bottom right.

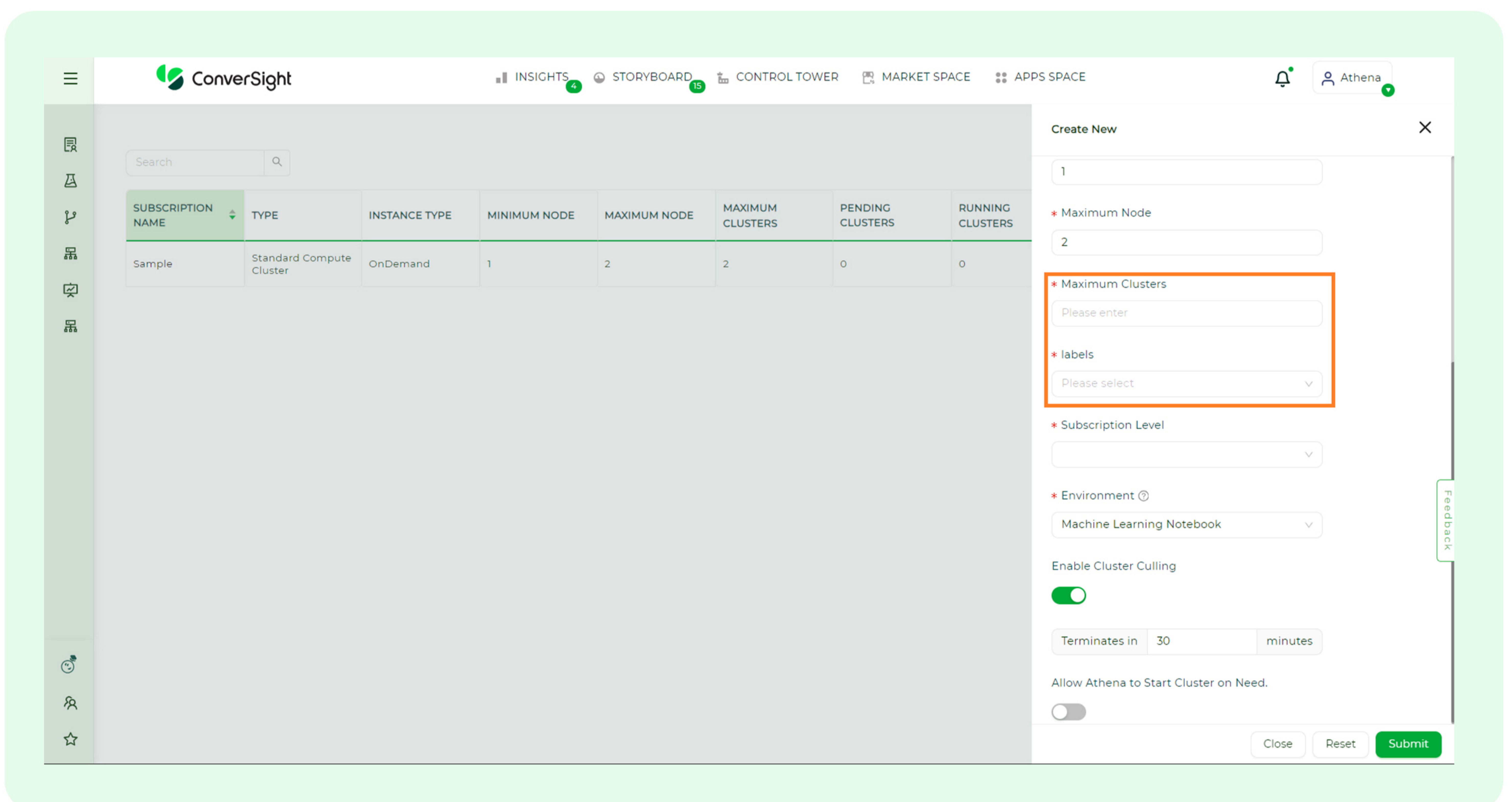
2.4 Nodes

Nodes in a cluster are the fundamental components that contribute to distributed computing and infrastructure management. They are virtual instances within the cluster. Nodes work together to handle tasks and distribute data, enabling efficient parallel processing. ConverSight enables users to customize the deployment of nodes in a cluster by allowing them to select the minimum and maximum number of nodes according to their specific requirements.

The execution of tasks starts with a minimum number of allocated nodes and additional nodes are automatically spawned as needed until the maximum number of allocated nodes is reached.



Users have the freedom to allocate as many clusters as they need, tailored to their specific requirements. Additionally, users can assign multiple labels to their clusters, allowing for enhanced categorization. However, it is important to note that at least one label is mandatory to ensure proper identification and management of the clusters within the platform.

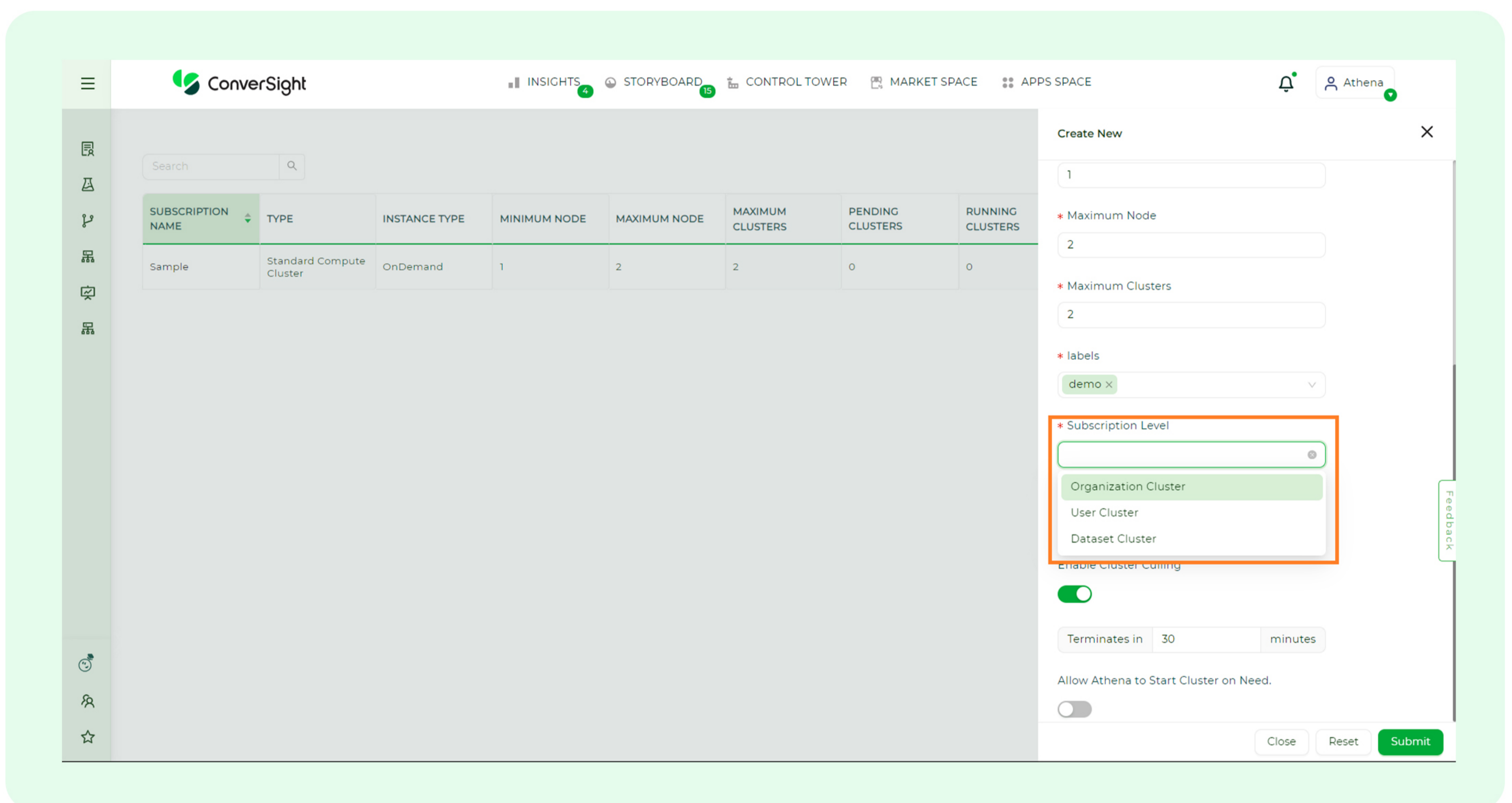




2.5 Subscription Level of Cluster

In the ConverSight platform, clusters are available under two distinct subscription levels: Organization Level and User Level. The **Organization Level subscription** is designed to cater to the needs of each organization. It provides a single computer that can be assigned to all users within that specific organization. This subscription level ensures efficient resource allocation and collaboration among users within the same organizational context.

On the other hand, the **User Level subscription** offers a more personalized approach. Each user is provided with their own dedicated cluster. This user-specific allocation enables individuals to have exclusive access to computing resources tailored to their specific requirements. It allows for customization and fine-grained control over computational tasks, providing users with a seamless and optimized computing experience.



2.6 Environment

Within the ConverSight platform, users possess the capability to tailor the Notebook environment of their clusters to suit their specific requirements. ConverSight offers two default options:

- Standard ConverSight Notebook
- Machine Learning Notebook

The Standard ConverSight Notebook serves as a versatile foundation, incorporating essential packages necessary for executing various tasks, workflows, CS Apps and performing fundamental analytical operations. Conversely, the Machine Learning Notebook is meticulously crafted to cater to the sophisticated needs of advanced analytics and machine learning endeavors. It provides users with an extensive suite of packages tailored specifically to these demanding tasks. Moreover, ConverSight empowers users further by allowing them to create Custom Notebook environments, offering flexibility to configure the environment according to their unique preferences and requirements. This capability enables users to optimize their workflow and maximize productivity within the ConverSight ecosystem.

SUBSCRIPTION NAME	TYPE	INSTANCE TYPE	MINIMUM NODE	MAXIMUM NODE	MAXIMUM CLUSTERS	PENDING CLUSTERS	RUNNING CLUSTERS
Sample	Standard Compute Cluster	OnDemand	1	2	2	0	0

Create New

1

* Maximum Node
2

* Maximum Clusters
2

* Labels
demo x

* Subscription Level
User Cluster

* Environment
Machine Learning Notebook
Machine Learning Notebook
Standard ConverSight Notebook

Terminates in 30 minutes

Allow Athena to Start Cluster on Need.

Close Reset Submit



2.7 Cluster Culling

Cluster Culling empowers users to define a specific time at which a cluster will be automatically terminated. This feature streamlines cluster management by eliminating the need for manual intervention. By specifying the desired termination time, users can automate the process of shutting down idle clusters within a predetermined timeframe. This automated approach significantly reduces costs associated with running underutilized clusters, enhancing overall efficiency and resource optimization within the system.

The screenshot shows the ConverSight dashboard with a 'Create New' modal open. The modal contains the following fields:

- Subscription Name: 1
- Maximum Node: 2
- Maximum Clusters: 2
- Labels: demo
- Subscription Level: User Cluster
- Environment: Standard Conversight Notebook
- Enable Cluster Culling:** (highlighted with an orange box)
- Terminates in: 30 minutes
- Allow Athena to Start Cluster on Need:

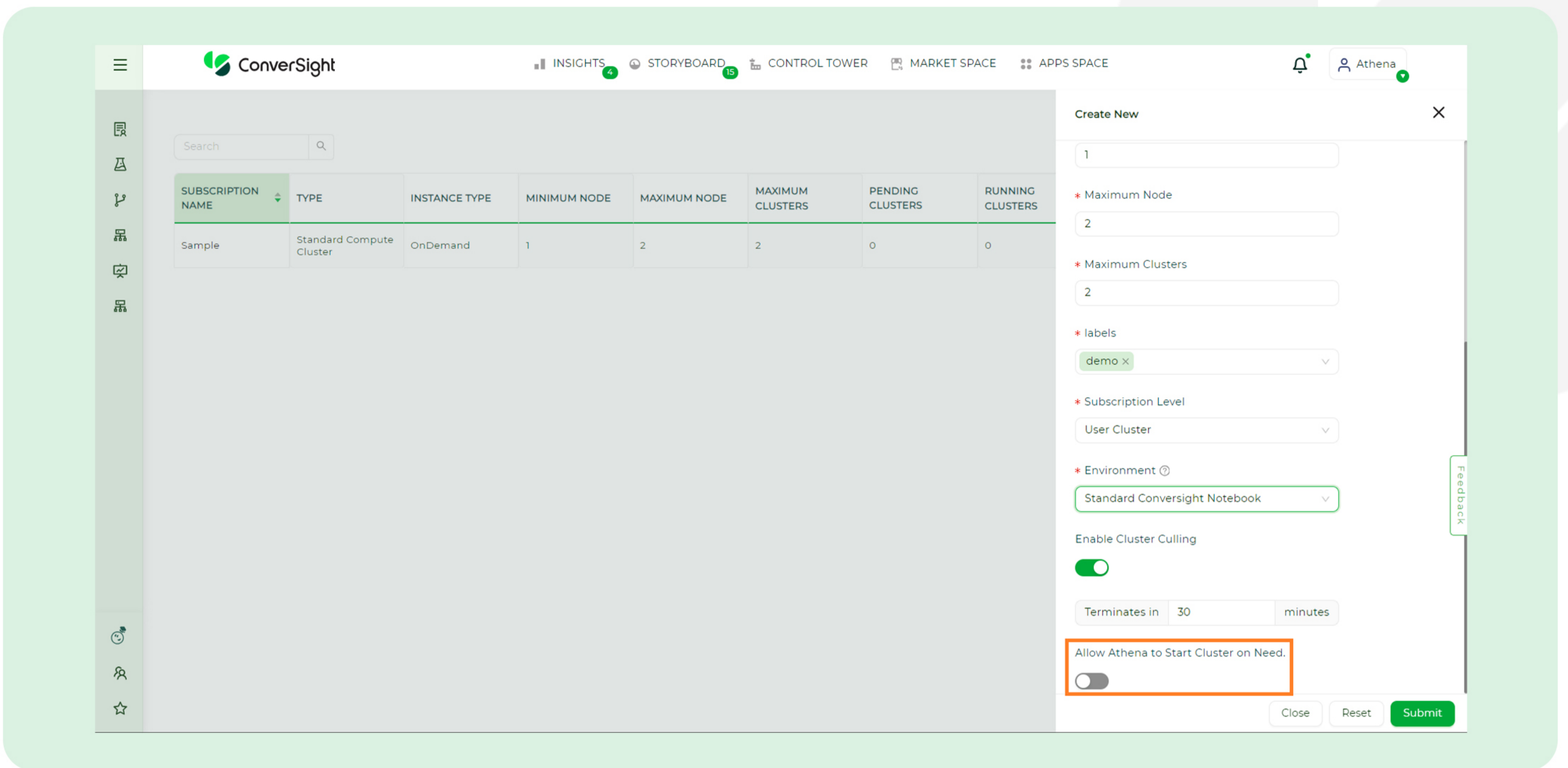
The background shows a table with the following data:

SUBSCRIPTION NAME	TYPE	INSTANCE TYPE	MINIMUM NODE	MAXIMUM NODE	MAXIMUM CLUSTERS	PENDING CLUSTERS	RUNNING CLUSTERS
Sample	Standard Compute Cluster	OnDemand	1	2	2	0	0

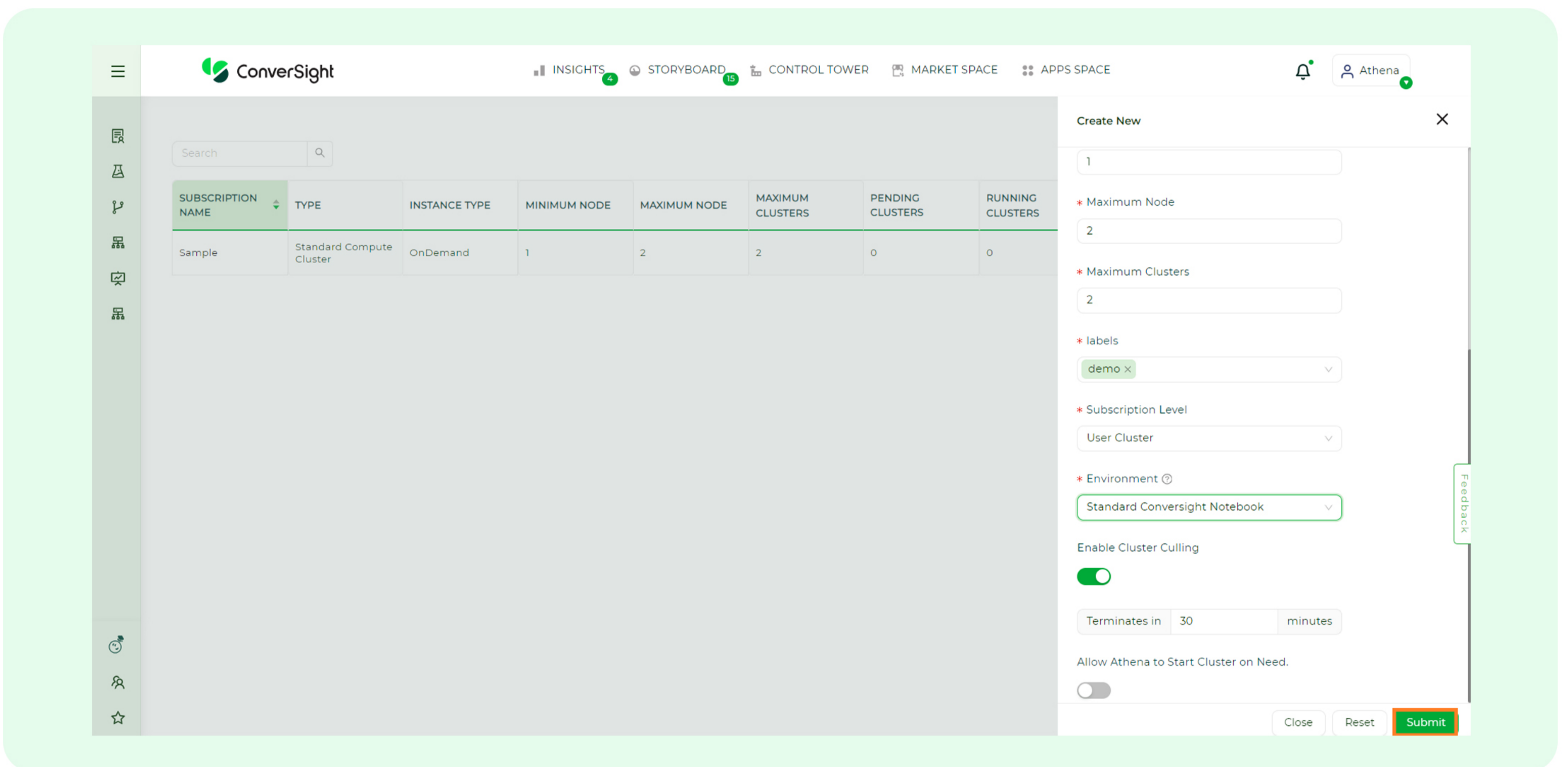


2.8 Cluster Initiation with Athena

When the toggle **'Allow Athena to Start Cluster on need'** is enabled, it signals the system to initiate a cluster according to the specified subscription level. However, it's important to note that the cluster will only commence if there are no active clusters already running on the chosen subscription level. This ensures that resources are utilized efficiently and avoids unnecessary duplication of clusters. By allowing only one cluster per subscription level, the system can effectively manage and allocate resources based on demand, ensuring optimal performance and avoiding unnecessary costs.



Once completed click on the **'Submit'** button.



The cluster will be generated according to the provided specifications and can be effortlessly modified to meet future requirements.

3. Key features of Compute Cluster

Compute Cluster empowers users with the tools needed to handle complex computational tasks seamlessly. The Compute Cluster provides a range of key features for efficient data processing and analysis :

Enhanced Performance: Clusters leverage multiple interconnected nodes to distribute computing tasks, resulting in improved performance and faster processing of complex data.

Scalability: Clusters can be easily scaled up or down to meet changing workloads and resource demands, ensuring efficient utilization of computational resources.

Cost Efficiency: By optimizing resource allocation, clusters help minimize unnecessary costs, making computational power more cost-effective.

High Availability: Clusters offer high availability and fault tolerance, ensuring continuous operation even if individual nodes experience failures.

Parallel Processing: With the ability to perform parallel processing, clusters can handle multiple tasks simultaneously, accelerating data analysis and computation.

4. Conclusion

In conclusion, Compute Clusters are a pivotal component in ConverSight, offering substantial benefits for enhancing computing capabilities. By leveraging the power of interconnected nodes, Compute Clusters enable efficient parallel processing, improved performance, and scalability. The utilization of Compute Clusters in ConverSight facilitates the seamless handling of complex calculations, data-intensive tasks and resource-intensive applications. The high availability, fail tolerance and resource sharing capabilities of Compute Clusters contribute to uninterrupted operations and optimized resource utilization. With the ability to scale horizontally and adapt to evolving computing demands, Compute Clusters in ConverSight provide a future-proof infrastructure for meeting the increasing computational requirements of modern organizations.

Join our customers who have accelerated growth with ConverSight



About ConverSight

ConverSight's Adaptive Analytics platform uses conversational AI, Natural Language Processing and machine learning to converge the distance between humans and data through data stories, presenting the meaning of data in the most effective, personalized and efficient form possible. ConverSight's patented AI business assistant, Athena, connects distributed databases to answer questions and Augment the consumers through 4 key functions: Information on demand, Automated Story Telling, Proactive Insights, and Recommended Actions.

For more information, visit www.conversight.ai

