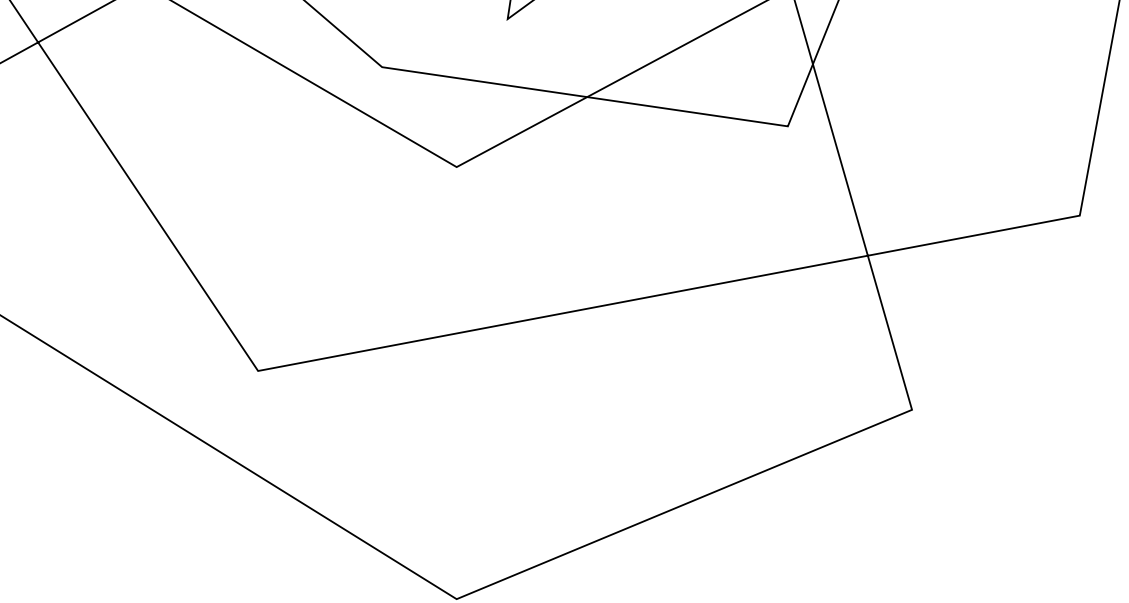


CODECO

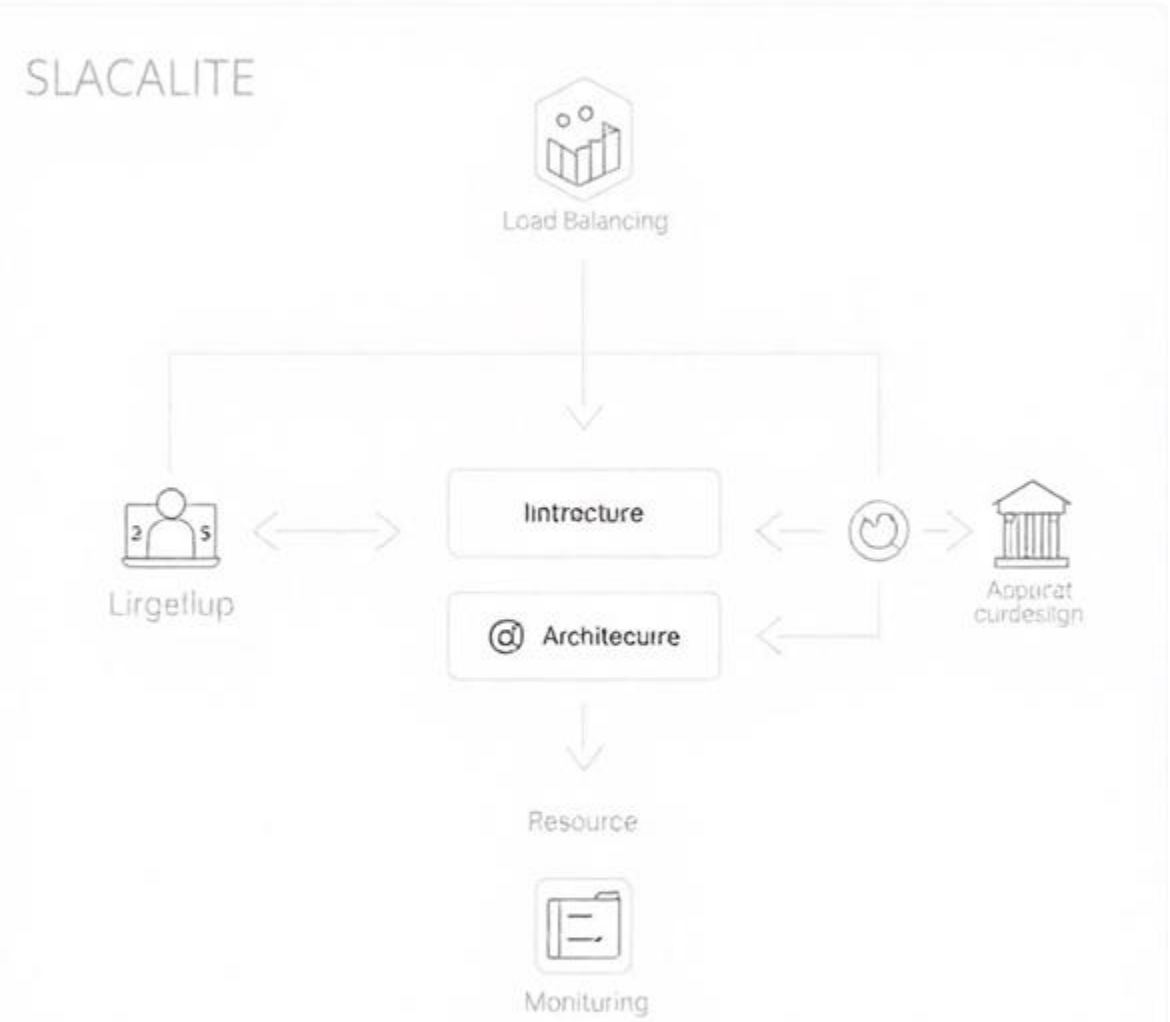
SCALABILITY AND ELASTICITY IN EDGE-CLOUD RESOURCE ORCHESTRATION

Funded by
the European Union

SCALABILITY IN EDGE-CLOUD RESOURCE ORCHESTRATION

Scalability is a fundamental aspect of edge-cloud resource orchestration, enabling systems to efficiently manage varying workload demands. It encompasses the ability to increase resources to accommodate growth without sacrificing performance. There are two primary types of scalability:

- **Horizontal scalability**: Adding more resources, such as servers, to distribute the workload across multiple nodes.
- **Vertical scalability**: Enhancing resources on a single node, like increasing CPU or memory capacity.



KEY CONSIDERATIONS FOR SCALABILITY

- **Designing a scalable architecture**: Create an infrastructure that can grow based on workload demands.
- **Load balancing**: Distribute workloads evenly across resources to maximise scalability.
- **Efficient resource utilisation:** Allocate resources wisely to prevent overutilisation or underutilisation.
- **Monitoring and performance analysis**: Continuously assess the system to identify bottlenecks and improve scalability.

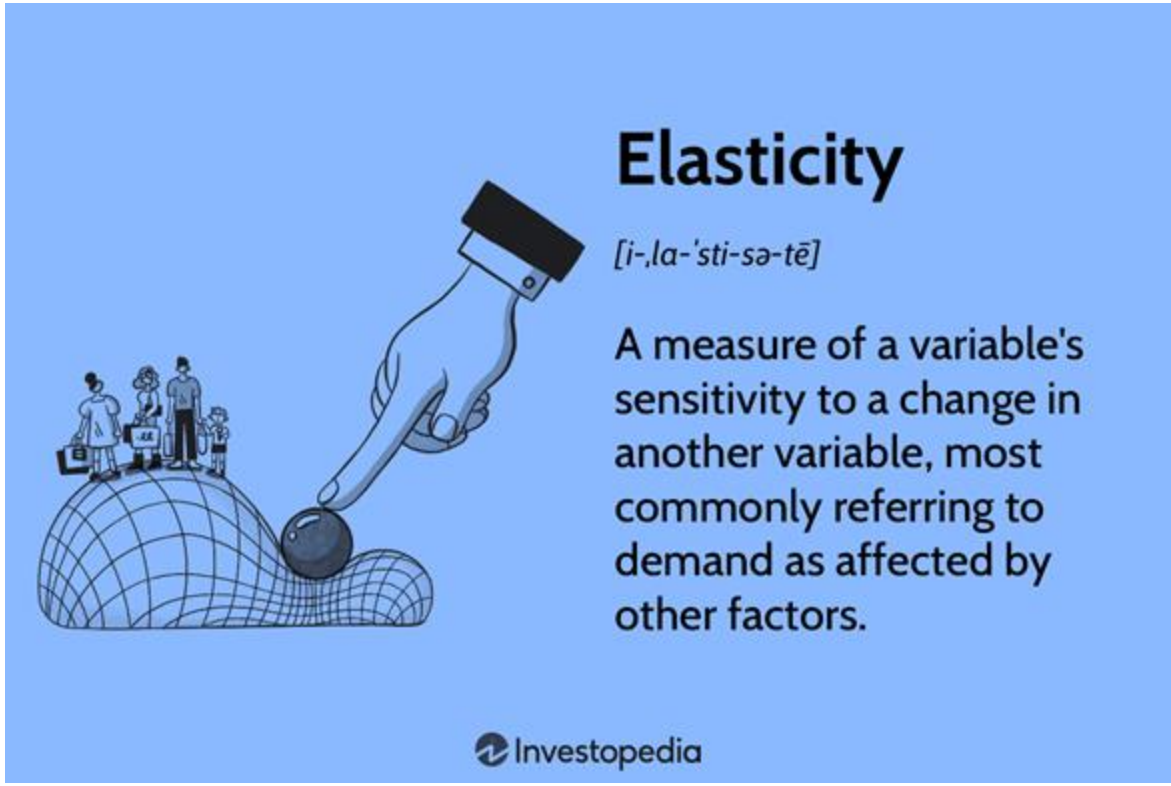
Funded by the European Union

ELASTICITY IN EDGE-CLOUD RESOURCE ORCHESTRATION

CODECO

Elasticity is crucial for dynamically adjusting resources in response to workload demands. This ensures efficient resource allocation and operational flexibility. Key elements include:

- **Automatic scaling**: Implement auto-scaling mechanisms to add or remove resources based on predefined rules.
- **Dynamic resource allocation**: Increase resource allocation during peak times and reduce it during low-demand periods.
- **Pay-as-you-go model**: Optimise costs by scaling resources as required, avoiding unnecessary expenses.



Funded by the European Union

KEY CONSIDERATIONS FOR ELASTICITY

- **Defining scaling policies**: Set clear policies and rules for scaling based on workload patterns and performance metrics.
- **Monitoring and prediction**: Analyse workload trends continuously to accurately predict resource needs.
- **Real-time adjustments**: Ensure prompt changes in resource allocation to adapt effectively to workload fluctuations.

SUMMARY:
OPTIMISING EDGE-CLOUD RESOURCE ORCHESTRATION WITH SCALABILITY AND ELASTICITY

CODECO

Scalability and elasticity are essential for optimising edge-cloud resource orchestration.

Scalability focuses on accommodating workload growth.

Elasticity emphasises the dynamic adjustment of resources based on demand.

Funded by the European Union