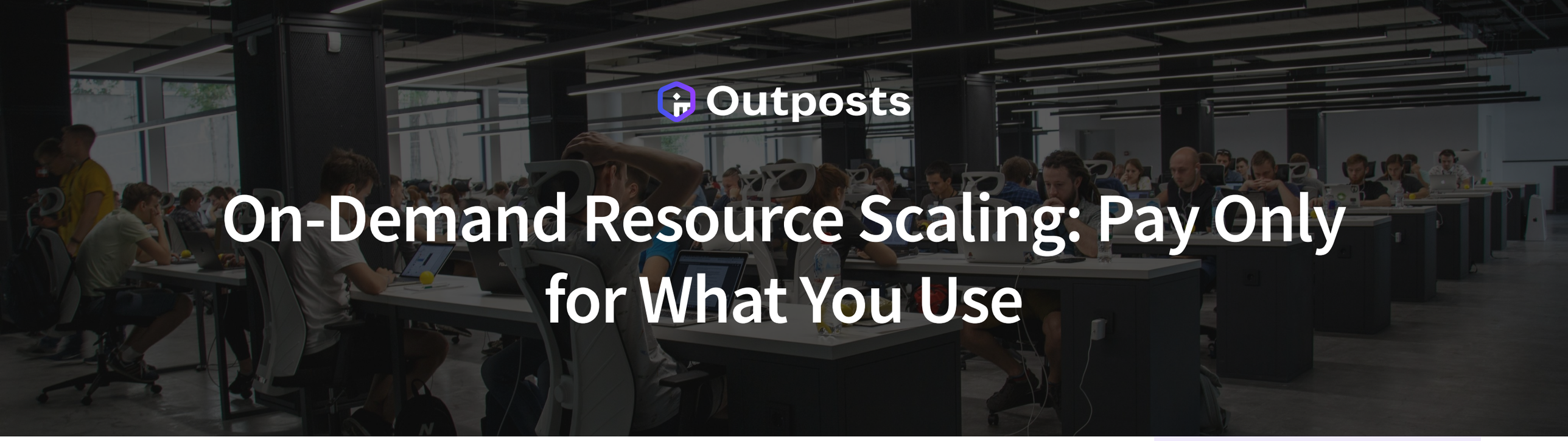

Outposts

On-Demand Resource Scaling: Pay Only for What You Use

Using cloud computing is popular, but it can be expensive if not managed properly. Imagine you have a website running on 10 small machines, but each one is only using 10% of its capacity. Instead, you could run that same website on just 5 medium-sized machines using 40% of their capacity each. This way, you get the same performance but pay less. This is what "sizing resources for cost efficiency" means.

In addition, you can automatically adjust the amount of cloud resources you use based on your actual needs. For this, we use horizontal pod autoscaling. With a load balancer distributing traffic evenly across pods, the system can detect when the load exceeds a certain threshold, like 70%, and automatically spin up a new pod to share the load. This process continues until the load stabilizes or the number of pods hits a predefined limit, preventing an infinite scale that could lead to excessive costs.

Cost efficiency is also achieved by turning off environments when they're not required, such as development or testing environments. We had a client who was initially on DigitalOcean but had to migrate to AWS, which is generally more expensive. Our team set up a Lambda function to scale their cluster up and down based on a schedule, taking advantage of AWS's per-hour billing. If the client only needed the environment during business hours, we effectively halved their costs.

In another case, a client was running analytics on their production database, which led to an unexpected spike in resource consumption. To prevent this from happening again, we set up a replica database that could be spun up on-demand through a Slack integration. The client could send a command to spin up the replica when they needed to run analytics. Once they finished, they could send another command to shut down and delete the replica database again. This way, they only paid for the replica resources when actively using them instead of having them running 24/7.

Sizing cloud resources for cost efficiency requires a customized approach tailored to each organization's specific use case. By leveraging tools like autoscaling, scheduling, and on-demand resource allocation, businesses can optimize their cloud spending while ensuring high performance and availability.

So, the main idea is customizing your cloud setup to automatically grow and shrink resources as requirements change. With some adjustments, the cloud can save you money by only using and paying only for what you use.

Nataliya Piskun

Chief Operating Officer

OUR CONTACT

+357 25 059376

hello@itoutposts.com

www.itoutposts.com

TOP RATED DEVOPS COMPANY

50+

projects delivered remotely

90%

of certified engineers in the company

2 years

average client engagement duration

4.9/5

customer satisfaction score

OUR AWARDS