

WHITEPAPER

Cloud Economics: Realize cloud cost savings of up to 30%



Today, for most enterprises who have made progress in their Cloud journey, cost optimization is the biggest challenge that they are facing. In fact, as per 2019 State of the Cloud Report from Rightscale, it is the number one priority for CIOs.

Top Cloud Initiatives in 2019

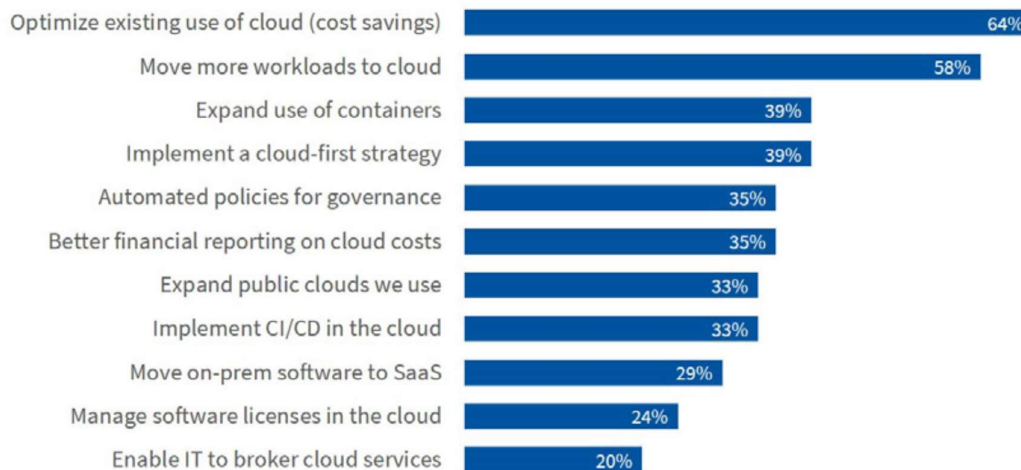


Fig 1: Rightscale 2019 State of the Cloud Report from Flexera

Given the problem statement, how should one approach optimization? Is this a one-time project? Should it be? Or, should it be an integral part of the ongoing cloud operations? Would tools solve the issue? In many cases CIOs say, "I have tools, but the issues remain". Why? These are some of the obvious questions that need answers.

Introducing Microland's Cloud Cost Management Framework

At Microland, we don't believe that cost optimization is a one-time exercise. It needs to be an integral part of ongoing cloud operations. Cost and governance should be the two pillars of today's public cloud operations. We follow our proprietary Cloud Cost Management Framework (CCMF) to help our customers tackle this issue on an ongoing basis.

CCMF is a baseline framework and is by no means intended to be prescriptive in nature. Rather, it provides a high-level direction on how this could be achieved. As cloud platforms are changing at a mind-boggling pace, this framework will need to be revisited frequently.

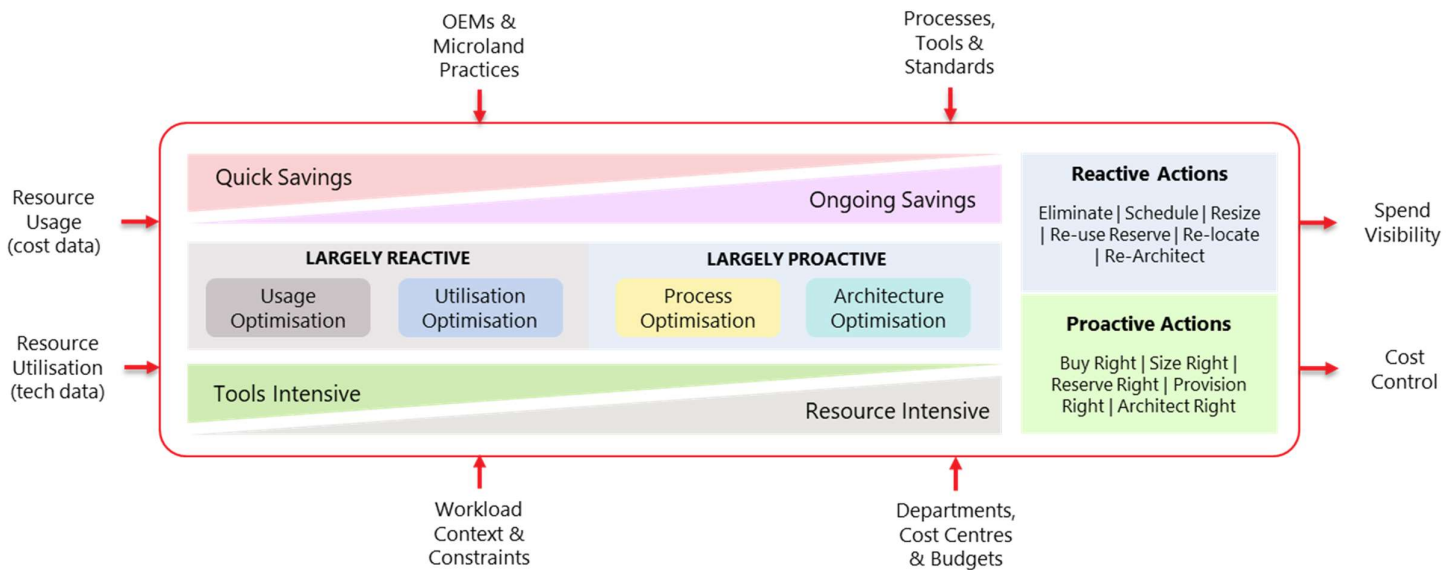


Fig 2: Microland's Cloud Cost Management Framework

Areas of Cost Optimization/Management

There are four broad areas for ongoing cost optimization and management listed below. Each one is progressively more intrusive, human resource intensive, time consuming, and expensive. In a brownfield environment, it is recommended to apply the model progressively from the first area to the fourth.

1. Usage Optimization

The objective is to identify resource usage patterns, find out their best run schedule and recognize resources that could be reserved by answering questions such as:

- Are there resources that are unused or orphaned which can be deprovisioned?
- Are there lean periods during which resources can be stopped and brought back up?
- Who (human/system) is using what and how much? Can we influence the demand?
- Can we reserve mostly used resources? Are there any unused reservations?

2. Utilisation Optimization

The objective is to find the best way to size, place or scale resources. While some recommendations could be made based on infrastructure utilization parameters such as CPU %, Memory %, Disk I/O rate and frequency, the potential pitfall is that none of these metrics provide us with the application or workload architecture context and therefore need to be validated with the business / app owners.

- Can we right size?
- Can we right place?
- Can we auto scale?
- How much is all the above is linked to workload characteristics and architecture?

3. Process Optimization

Optimizing processes is important to ensure that no unwarranted changes are made to the cloud environment that could negatively impact costs and cost management efforts. Answering questions such as the following would help enterprises in achieving this.

- Do you have a provisioning process for individual resources / reusable stacks that ensures the following is defined before / immediately after provisioning?
 - Expected usage, utilisation, expiry, scheduling, autoscaling, tagging, grouping
- Do you have a purchase process that ensures the purchase of cloud services is optimal? (subscriptions, licenses, marketplace items, reservations, discounts, etc)
 - Centralised vs. Distributed purchases, EA / CSP / PAYG rationalisation (in case of Microsoft Azure), centralised reservation purchases and discount negotiations
 - Usage data on reusable licenses, unused reservations, predicted new spends (new workloads)
- Do you have a defined process for multilevel (BU, Department, Project, etc.) budgeting, usage allocation, downstream billing/show-back and reconciliation to ensure your influence usage?
- Do you have an escalation mechanism to ensure recommended optimization actions on resources/workloads are implemented within time?

4. Architecture Optimization

Architecture level assessment is the most critical and the most difficult one to carry out from an optimization perspective. The reasons range from availability of resources (knowledge, skills & bandwidth), cost & time of the exercise and risk appetite.

However, in our experience, this is where the maximum cost savings happen in the long run.

- Do you have a team or person whose primary focus & responsibility is to analyse, rearchitect (simple/complex) and implement the changed / optimised architecture?
- Have you defined 'deployment architectural patterns' that could be used for various workloads and tagged them accordingly? (e.g. Traditional IaaS, Cloud native IaaS, HC PaaS, LC PaaS, N-tier on containers, Microservices on containers or combinations)
- Do you conduct initial and periodic workload architecture review and approvals from a cost angle before release to environment (Migration or Green field deployments)?
 - Cost effective alternates, reservations, licenses/AHB, scheduling times, normal /peak / seasonal utilization estimations, autoscaling, tagging, grouping
- Do you have a mechanism (not just tools) to monitor and compare actual usage and utilisation against estimations and fine-tune the architecture or revise estimates?
- Do you record the reason for selecting a deployment architecture during design and periodically review incorporating new technology introduction by the cloud platform that could optimise costs marginally or significantly?

The first two optimization areas, usage (1) and utilization (2), will typically yield more benefits to start with as they are low hanging fruit. Tools play a major role in this. Beyond that, we start seeing diminishing returns from focusing solely on usage and utilization optimization. Having said this, it is necessary to keep a constant eye on usage and utilization, because over the time there is a possibility of increased cost leaks due to changes introduced by the Cloud provider and within the enterprise's environment.

It important to understand that it is neither practical nor sustainable to carry out optimization exercises in the first two areas without the help of cloud cost optimization tools. To start with, we

will need tools that can discover all the cloud resources, their usage and costs, tagging & grouping them in various ways that enables you to apply rules to identify 'idle', 'reservable' or 'optimizable' resources.

The most difficult part in a cost optimization project / management program is to implement the recommendations to realize savings. Implementation may not be possible or delayed due to technical or nontechnical reasons which may be fully outside the control of the party who is responsible to identify leak points and provide pertinent and realistic recommendations. It is important to have a mechanism to influence the prioritisation and time taken to implement this.

Microland's Offerings

To help enterprises with cloud cost optimization, we have the following set of tiered offerings.

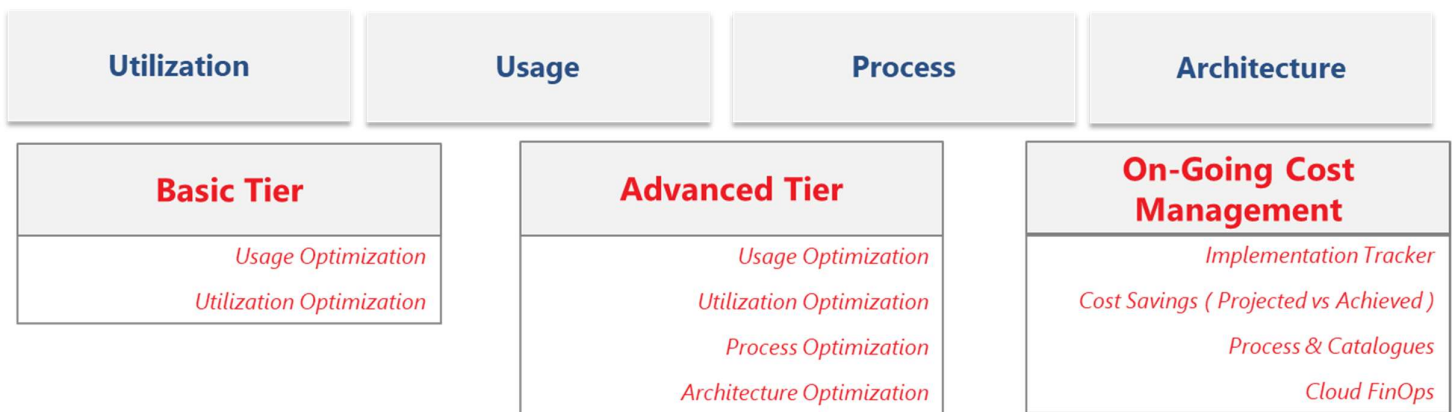


Fig 3: Microland's Cloud Cost Management Service Offerings

Why Microland?

Following are the reasons why we believe that we are uniquely positioned to offer these services:

- Experience in working with multiple cost optimization tools across multiple cloud providers
- Unique framework and methodology
- Proprietary and constantly evolving optimization library of rules across IaaS and PaaS services
- We have, on an average, saved 28% of monthly spends amounting to millions of dollars for various global enterprises.

About the Author:

For more information visit www.microland.com/cloud-economics, or contact Sunil Sarat, SVP, Enterprise Solutions & Head of Cloud & Data Center Practice directly at sunilps@microland.com. He is a technology leader, who has built offerings & practices in the Cloud & IT Infrastructure space. He is a thought leader as well, having authored books, blogs, and whitepapers. He has also spoken at various tech conferences worldwide.

Microland Contact Information:

Microland Limited | Birmingham

Suite 4, Two Devon Way, Longbridge Technology Park, Birmingham - B31 2TS, United Kingdom
+44 12 1222 5495 | www.microland.com | info@microland.com