Google Cloud

Getting started with FinOps on GCP



Authors: Sam Moss, Kinjal Tanna, Tan-Minh Truong

Contents

Scope	4
FinOps overview	5
What is FinOps?	5
Why does FinOps matter?	6
When should FinOps be implemented?	7
Who is responsible for FinOps?	8
How does FinOps work?	9
FinOps cultural principles	10
The variable nature of cloud	11
Business-value based decision making	12
Accountability	13
Collaboration	14
Blamelessness	15
Democratization of data	16
Agility	17
The FinOps Central Team	18
FinOps Lifecycle	20
Inform	21
Allocate	22
Shared resources	22
Unlabeled projects/anomalies	22
Report KPI & cost	24
Defining key reporting requirements	24
Sourcing appropriate data	24
Tracking through reports	25
Maintaining and improving reporting	25
Forecast	26
Creating forecasts in GCP	26

Reviewing forecasts	27
Maintaining forecasts	27
Optimize	28
Prioritizing and implementing optimizations	29
Product efficiency	32
Refactoring workloads	32
Rightsizing	33
Pricing efficiency	34
Guardrails	35
Operate	36
Onboard workloads	38
Educate and enable	39
Learning services	39
Enablement services	40
Incentivize	41
Measure value	42
Funding	43
Closing the loop	44
Next steps	45
Appendix:	46
Appendix A: organizing Epics according to capabilities	46
References and resources	47

Scope

This technical paper outlines a FinOps framework for Google Cloud Platform (GCP), covering the teams required, the processes and behaviors to embrace, the outcomes expected, and an approach to implement it.

To baseline a shared understanding across the industry, this technical paper will align with the framework set out by the <u>FinOps Foundation</u> and build on top of the established industry phases of Inform, Optimize, and Operate with GCP-specific Epics.

The aim of this technical paper is to provide an actionable set of steps to help your organization implement FinOps and maximize your cloud investment¹ on GCP.

This technical paper will not cover:

Security

While security plays an important role in FinOps (e.g. billing IAM roles limit the possibility of malicious spend), security policies do not contribute to the aim of maximizing your cloud investment.

Basic cost management setup on GCP

While fundamental to all aspects of FinOps, there is extensive <u>existing documentation</u> already available. This technical paper will move your organization beyond the basic technical setup of cost management in GCP. It is assumed that <u>a billing export and a central billing store has already been set up</u>.

Product-specific optimization techniques

For example, "How can I optimize my BigQuery costs?". Product-specific techniques can be found in public <u>documentation</u> and the <u>cost optimization technical paper</u>. This technical paper focuses on the overall strategy of cost management and FinOps in GCP.

Accounting practices for cloud spend

While accounting practices can significantly impact the way cloud spend is managed, this technical paper will not cover how finance should treat and optimize cloud expenditure from this perspective.

¹ All references to GCP products are correct as of May 2021.

FinOps overview

What is FinOps?

Cloud FinOps is an operational framework and cultural shift that brings technology, finance, and business together to drive financial accountability and accelerate business value realization.

FinOps, though different in its aims, is similar in concept to DevOps.

Both aim to bring disparate functions of your organization together to create a better way of working that puts collaboration, agility, ownership, and value at its core. The development of Cloud FinOps is a recognition that in order to effectively manage the variable spend model of cloud, new practices need to be adopted. FinOps ensures you get the most out of every dollar (pound/euro) spent in the cloud.

Cloud FinOps encompasses a wide range of existing and net-new processes. It permeates almost every aspect of the chain of value illustrated below, which links spend in an organization to a business outcome. By considering the end-to-end process, FinOps practices put the focus on maximizing the returns on your cloud investment.



Why does FinOps matter?

Achieving the full benefits of cloud requires fundamental changes to both mindset and behavior around existing financial management practices.

It is a balancing act between allowing the organization to take advantage of the flexibility that cloud brings and ensuring that spend results in business value. Allowing too much flexibility with no framework will almost certainly result in wastage, but too much financial control can limit some of the benefits that the cloud is meant to provide.

This often plays out as differing motivations between your workload and finance teams. Your engineers are eager to take advantage of the flexibility and licence to experiment, but your finance teams may, in turn, fear the shock of cloud bills when those engineers aren't accountable for their costs.

FinOps practices help align teams under a single aim – maximizing cloud investment.

When should FinOps be implemented?

Whatever stage of the cloud journey you're at, FinOps practices will help you get the most value out of GCP. However, a tried-and-tested point of implementation is during the first stages of a cloud transformation program, alongside other foundational activities, such as security assessments or identity and access management (IAM) configuration.

A common mistake is to start considering FinOps only when the bill becomes significant. This usually results in some unexpected cost spikes at some point, leading to a "spend panic" moment (e.g. hurting value realisation by stopping all work). Moreover, the longer inadequate practices have been in place, the harder it will be to change them to adopt FinOps best practices.

No matter where you are in your cloud adoption, the first step to implementing FinOps is the same: start small, iterate through the lifecycle, and continuously improve.

Who is responsible for FinOps?

A key aspect of FinOps is to drive everyone to take accountability for their costs. To that end, every user of cloud should feel responsible for their spend and empowered to take action to optimize it.

Collective responsibility is neither apparent nor intuitive for teams who have not had to consider their infrastructure running costs before. Therefore, it needs to be fostered and driven by a centralized FinOps team (hereafter known as the FinOps Central Team), which is accountable for implementing and embedding FinOps practices and behaviors.

The relationship between the FinOps Central Team and the wider organization is therefore one of enablement. Keep in mind that the existence of a FinOps Central Team does not mean they are responsible for all aspects of cloud FinOps. It's important to establish clear expectations from the outset to avoid bad practices developing; the Central team should focus on establishing practices and frameworks to guide the organization.

How should FinOps be implemented?

Google's FinOps framework has been developed using the key dimensions of **People**, **Process**, and **Technology**.

The **People** dimension consists of those who are responsible for FinOps practices: the **FinOps Central Team** and **Cross-functional Stakeholders** who use cloud. It also encompasses the FinOps Cultural Principles, which underpin success.

The **Process** dimension is based on the **Google FinOps Epics** – the most important activities when implementing FinOps on GCP. In this paper, an Epic is a set of activities which you can apply to an individual workload, a group of workloads, or your entire cloud estate.

This paper organizes these **Epics** (seen in the diagram below) according to the FinOps Foundation lifecycle of Inform, Optimize, and Operate. This helps to show the order in which they should be performed to successfully implement FinOps on GCP.

The **Epics** can also be organized in other ways, for example by the capability they enable (see appendix 1). If you have an existing framework, you can map them to that instead.

This approach is iterative and continuous. Once the cycle ends, the process starts again with FinOps practices improving gradually each time. Lack of full-fledged maturity for each process need not be a blocker to adoption, as the organizational knowledge and confidence will grow through repeated iterations of the FinOps lifecycle.

Finally, the **Technology** aspect consists of GCP tooling which can be leveraged to support the **Epics**².



2 GCP Partners who leverage raw Google Billing Data also provide technology and tooling to support FinOps practices.

FinOps cultural principles

At Google, culture and mindset is considered a critical enabler to everything we do. FinOps is no exception to that. The cultural principles defined here are based on Google's culture of collaboration, openness, and blamelessness. They are not specific to any part in the life cycle, but rather should permeate every Epic in the Inform, Optimize, and Operate stages.

The variable nature of cloud

FinOps is successful when all employees of an organization embrace the variable cost nature of cloud.

Cloud brings a new cost structure, and costs are variable due to the pay as you go nature of consumption. There are new risks (e.g. runaway spend), new benefits (e.g. more data to leverage), and emerging thought patterns (e.g. fixed-cost mindset).

For example, a team who previously ran a server on-prem, when only needed for a few days a week, will now have to learn to spin that server up and down on GCP only for the hours needed instead, to reduce waste. In addition, the Finance team will have to learn to see fluctuation of costs as a positive i.e. the server being up and incurring charges is not an uncontrolled spike of cost, the server being down when not needed is a good cost control measure.

Bandwidth may now also be charged on a cloud model (price per GB used) versus purchasing a one-off amount. Embracing this change and, furthermore, optimizing for it is key to successful FinOps practices.

2 Business-value based decision making

FinOps is successful when financial decisions are made from a business value perspective.

When reviewing cloud spend, it can be tempting to make financial decisions from a purely cost reduction or cost optimization perspective. However, when moving to cloud, new benefits such as innovation and access to new capabilities (e.g. AI, ML) emerge.

All financial decisions should take these new considerations and opportunities into account. Cost is an important factor, but not the only one to consider, as new opportunities, speed to market, and velocity of change all affect the overall return on investment. Don't forget the increased staff productivity, reduction in downtime and performance improvements that can be additional benefits from moving to GCP.

Simply reviewing costs without taking into account the additional cloud context and increase in business value, or making like-for-like comparisons between on-prem and cloud workloads will likely not serve an organization well when addressing business value. When including all the variables above, these can quickly outweigh managed services, which are more expensive than building block laaS services.

3 Accountability

FinOps is successful when employees understand and are accountable for their spending in the cloud.

If an employee has access to cloud resources, they impact the usage (and hence the cost) of these resources. For an organization to be successful at implementing FinOps, the FinOps Central Team must help cloud users shift from seeing cost as someone else's responsibility, to seeing cost as their own responsibility. Employees across the business are now empowered to make cost decisions that are right for their teams, workloads, and the organization as a whole. Accountability also extends to the implementation of any recommended cost-effective actions that arise from relevant data or reports.

For example, the FinOps Central Team may surface a more cost-effective architecture pattern to a community. A project team, recognizing the relevance to their workload, could then proactively leverage this to improve it. Project teams must be empowered to act on recommendations and implement relevant changes.

4 Collaboration

FinOps is successful when teams work together.

The FinOps Central Team brings many disparate functions together to create a unified understanding of how to maximize cloud value. However, simply bringing the teams together is not enough – teams must join with an openness to collaborate, to break down silos, to discuss best practices, and to help each other learn and avoid the same cost mistakes again and again.

Teams may contribute to this aim in different ways. For example, the FinOps Central Team will have access to organization-wide cost data, but the Finance team may be missing the context of why Team A has chosen to take a different approach (e.g. using more expensive managed services) to Team B (e.g. running containers on VMs) when running similar workloads.

Fostering and engaging in a culture of collaboration can help all teams bridge the gap between decisions that are often made in silos. It also helps form strong repeatable patterns and learnings that bring the most business value to the organization.

Blamelessness

5

FinOps is successful when mistakes are treated as blameless opportunities to learn from.

"<u>Make it safe to fail fast</u>" – everyone is aware of this DevOps culture slogan, and for good reason. Encouraging an atmosphere of growth in your company also means acknowledging that mistakes, sometimes costly, can happen. This is as true for over-spending in the cloud as it is for any other part of your business.

Rather than finger-pointing and shaming individuals who overspend or waste resources, a blameless culture helps teams identify the cause to understand how spikes or issues occur in the first place. In a psychologically safe environment, individuals who spot issues are more likely to speak up, and better still, mistakes can be anonymized and shared across the business to ensure they don't happen again.

Don't mistake blameless culture for a loss of accountability. Employees are still accountable for the decisions they make and the money they spend – but when things go wrong, the emphasis is placed on the learning opportunity to prevent recurrence. <u>Blameless postmortems</u> (used by Google SRE teams) can be a useful tool to identify the root cause of an issue in the context of systems, rather than people.

Democratization of data

FinOps is successful when data is democratized and available to all.

For employees to make cost-aware decisions, they need a complete picture of relevant data as close to real time as possible. Data in reports and dashboards must show the cost and value impact of team members' decisions to the business as they occur.

Providing access to other teams' usage, cost, financial, and business data serves as a good baseline to spot more efficient patterns, promoting a shared understanding of how best to employ cloud services.

There may be some initial reluctance to share spend data – particularly if psychological safety in the organization is low. Additionally, in some cases, raw spend data may include information your organization is not comfortable sharing widely for legitimate business purposes. However, you should avoid a default 'no access to view' policy towards what teams are spending on cloud. The target and ideal state is to ensure there is a single source of truth and employees can rely on the data to inform their decision making. Agreeing on the form and the content of the source data can save time and effort. As a first step, consider making the raw billing data in the <u>billing export</u> available to all.

Agility

FinOps is successful when carried out in an agile manner. It requires a combination of an iterative approach, a commitment to continuous improvement, and automation.

For example, you should:

- Drive agile processes for budgeting and planning within IT by promoting iterative budget allocation, IT spending, and forecasting.
- Drive continuous improvement by sharing cost-efficient, successful architectural patterns widely among your employees.
- Improve the information you have at every available opportunity.
- Define rules on how a FinOps process should be automated and automate configuration through policy as code.

And while the FinOps lifecycle is iterative in nature, do not wait for the end of a cycle to improve – embrace improvement at every step of every cycle.



The FinOps Central Team

With the job of implementing and embedding FinOps practices and behaviors, the FinOps Central Team brings together and aligns different contributing teams. Each team will be responsible for many aspects of the framework we illustrate here. However, each team will also naturally focus on certain Epics.



While your organization may call this team a different name, like the Cloud Business Office or Cost Management Office, the role of this centralized team remains the same.

Specifically, the FinOps Central Team will be expected to:

- **Demonstrate GCP knowledge** to interact with engineers to identify wastage and discuss cost optimization activities with them.
- **Demonstrate financial knowledge** to interact with the organizational finance and business teams, and help clarify actual spend, forecasted spend, and budgeting.
- Establish and run common activities that are more efficient when centralized, for example purchasing Committed Use Discounts.
- Make reliable cost data available, enabling workload owners to take ownership over their share.
- Align practices among finance, workload owners, and business teams.
- Maintain neutrality between parties, taking into account all viewpoints.

- Speak the language/vocabulary of all the other cloud stakeholders, including finance and engineering.³ The team also helps establish and define what a 'dollar' of Cloud spend specifically means to help build trust and confidence in the data.
- Act as an ongoing catalyst for change, driving FinOps adoption and promoting FinOps principles in the organization.

Anti-pattern: the FinOps Central Team becomes the Cost Optimization Team.

A common anti-pattern is to have the Central Team take full responsibility for identifying and implementing all **cost optimizations** to workloads in the cloud, based on the assumption that they are best placed to spot high value opportunities that can improve efficiency. This can work well in the early stages when usage is relatively low, but falls apart as usage scales. The Central Team will struggle with scaling, and project teams are unlikely to agree to others changing their solutions.

What the FinOps Central Team doesn't do may be as important as what they do. As stated earlier, the aim for this team is to *enable* good practices. There are some activities which are important to FinOps but are best performed by other teams, for example enacting cost optimizations or refactoring workloads. Those activities are always best performed by the workload owners – they know most about it and can easily make the change.

Keep in mind that the FinOps Central Team doesn't need to consist of full time roles, at least to begin with. In the early stages of an organization's cloud journey – when there are only a small number of workloads – it could even be wasteful. The key point is that there is an individual or team responsible for the above, even if this represents only a small time commitment. As the organization's usage grows, the team can grow as needed.

In addition, large organizations may wish to balance both a centralized and federated approach.⁴ It is not wrong to have local or regional FinOps function to be closer to workload teams and a central team to tackle company-wide topics. The right balance will be found thanks to the iterative self-balancing nature of the FinOps approach described below.

A typical organizational construct for any cloud adoption program is a Cloud Adoption team (a.k.a. Cloud Adoption Office or the Cloud Centre of Excellence). This team provides common services which are necessary for success (e.g. learning and enablement) and benefit the whole organization. The FinOps Central Team is best established there in order to leverage other central services.

³ The FinOps Central Team may want to achieve this aim by creating a working vocabulary list that defines key terms used in the organization. They should drive to a point where other teams can have meaningful conversations without their necessary presence to translate terms.

⁴ Federation here means the distribution of accountability across the organization versus. centralized services

FinOps lifecycle

The FinOps Foundation lifecycle consists of three stages: Inform, Optimize, and Operate.⁵ The Google FinOps Epics are organized according to these stages.



Together, the stages and Epics form a set of steps you can take to implement FinOps practices on GCP. Each of the stages has a different aim. The **Inform** stage aims to make information more visible, such that **Optimizations** can be made to your cloud usage. These optimizations are then embedded into the organization in the **Operate** stage, at which point the cycle begins again.

While your organization can start at any point in this lifecycle, Inform is a natural place to start in absence of any specific requirements. Although starting with Inform gives you the advantage that all future phases will be data led, each stage of the lifecycle is a necessary part of practicing FinOps successfully.

It is recommended to start small. If some Epics do not fit your business right now, feel free to skip them and integrate them on future iterations – **embrace progress, not perfection**. Once you've completed the lifecycle for the first time, use your data-led learnings to go around again, with specific measurable goals in each epic you choose.

5 These phases were originally outlined in Cloud FinOps: Collaborative Real-time Cloud Financial Management (Storment and Fuller).

Inform

Inform aims to provide transparent, accurate, and actionable data to all.

As a bare minimum, a mechanism must be implemented that allows cost to be allocated to the right teams, so they can easily see what they are using/responsible for. This information needs to be made visible through dashboards and reports, so teams can drill into and slice the data in meaningful ways, perhaps tracking KPIs as well as costs. Teams also must be able to see their forecasted spend to be able to make decisions on any necessary changes or alert the right people of the expected costs.

Inform consists of three Epics:

Allocate

The transparent mechanism to ensure 100% of costs have owners

The ability to make more cost data available, consumable, and

Report KPI & cost

actionable

Forecast

The ability to project and track future spend

The FinOps Foundation <u>State of FinOps Report 2021</u> found that when members were asked what their largest challenges on cloud were, 3 out of the top 4 challenges included full allocation of costs (26%), dealing with shared costs (33%), and accurate forecasting (26%). This demonstrates how getting the Inform phase right is key to driving meaningful and long-lasting business change.



Inform: Allocate

Allocation is the practice of assigning all the costs between your teams to ensure each team is paying for what they agreed on in a fair and transparent way.⁶

Allocation is affected by:

Billing accounts

The relationship between a GCP project and a billing account is many-to-one, meaning that multiple GCP projects are attached to one billing account for the entire organization.⁷

Best practice

Use a single billing account for simplicity. Devise internal chargeback mechanisms to allocate costs amongst departments.

Labeling strategy

Infrastructure on-premise may have previously had a clear chargeback mechanism. <u>Labels</u> in GCP ensure that the right teams are paying for the cloud services they use. Labels provide a way to tag projects (and specific resources) with a key-value pair.⁸ By introducing a consistent labeling structure, detailed granular costs are categorized and the results can then be made transparent in Reporting.

A labeling strategy should follow your organization's chargeback mechanism, which will help identify the kinds of labels necessary so it is easy to see who to charge during invoicing.

Best practice

Cost allocation labeling should be done at the project level. A set of globally-applicable labels that are applied to all projects at creation should be agreed upon. Avoid multi-tenant GCP projects in which you need to label costs at the resource level as this adds much complexity.

You will need to understand cost center reporting requirements before defining this labeling strategy. For example, do costs need to be aggregated and split by department? Business unit? Environment?

⁶ It is likely that this would include many factors beyond cloud usage costs (such as FTE operational costs) to ensure that a complete and accurate understanding of the cost to run the workload is acknowledged. This whitepaper however, will cover allocation from a pure GCP viewpoint to explain how these specific costs (usage & shared usage) can be allocated to teams.

⁷ However, there are some circumstances in which multiple billing accounts may make sense for the organization (loosely structured conglomerate, tax implications based on the country of the billing account, an educational organization with various entities that should not affect each other, or reseller accounts).

⁸ Tags in GCP look to evolve GCP's labels offering and are currently in development as of June 2021.

Applying and validating labels

Applying labels manually leaves room for user error. Automating this process⁹ will lead to higher compliance. If labels are not applied on projects, costs of that project will not be visible in reports, which will require intervention to fix.

Best practice

Apply labels according to the labeling strategy at project creation via an automation process.

Shared resources

When using GCP, there are likely to be cloud resources that are shared across your teams (for example, common data stores, multi-tenant clusters, shared services, support subscriptions, etc). Consider:

- Division of shared services spent amongst departments/teams. Labeling shared projects as "shared" allows you to see which projects are in use by all. The relative cost can then be split to different departments.
- For a subset of services, the <u>Requester Pays</u> configuration allows you to charge operational costs to the user, rather than the owner.

Best practice

Ensure that these shared services are also allocated relative to the use of GCP.¹⁰ As your organization matures, it should review the allocation of both costs and savings made through each iteration to teams.

Unlabeled projects/anomalies

To guard against configuration drift or noncompliance, periodically scan your environment for labeling violations. In GCP, various policy-as-code or CSPM tools exist that can monitor for misconfigured resources, including labeling violations. A commonly used tool in the GCP ecosystem is <u>Forseti Config Validator</u> (other third-party/multi-cloud tools are also available).

Best practice

Ensure tooling is in place for labeling violations and create relevant business processes to follow in these scenarios.

⁹ Consider the use of <u>Google's Cloud Foundation Toolkit</u> if you're just setting out – there are options for both Terraform and Deployment Manager, and the <u>Project Factory module</u> ensures labels are applied automatically during provisioning.

¹⁰ There may be resources that cannot be shared based on a simple percentage usage model, the FinOps Central Team may need to work with application owners to come to a solution of how best these resources can be divided up (for example, networking, monitoring, taxes, upfront reservation charges, etc.)

Inform: Report KPI & cost

1

Making data available and usable by your stakeholders is paramount. The report KPI & cost Epic centers around taking all the relevant data sources (financial, business, usage) to visualize your data in easy, clear, and actionable dashboards or reports.

Reports allow individuals to see the effect of their changes on the spend of the business. Clear reports surfacing relevant data will drive both high and low-level habit changes, resulting in a closer alignment between spend and business value. <u>The GCP Console Reports</u> provide a ready-made, comprehensive, and centralized view of spending, but you may need to build additional dashboards and add in your own custom metrics or KPIs to put spend into context. Workload teams may also wish to build their own dashboards from the raw billing data to cater for their own requirements.

You can also complement reports with the templatised <u>Cost Insights Dashboard</u> that provides organizational usage and revenue data alongside surfacing insights to optimize your spend in GCP. This includes data from the <u>Recommendations Hub</u>, <u>Cloud Billing Usage</u> and <u>Custom Pricing Data Export</u>.

Defining key reporting requirements

This should combine business and user requirements:

- Business requirements (cross-business units, multi cloud)
- User requirements (personas, latency, method of access)

Identification of your user groups is crucial to developing targeted reporting and dashboards that equip individuals to make cost optimization decisions. A subset of these reports should also surface key performance indicators (KPIs) to put spend into context.

Sourcing appropriate data

There are a number of different data sources needed to make reports cohesive and complete. Initially, raw usage data drawn from the <u>billing export</u> may be the only data source available to the FinOps Central Team. As your organization matures, consider integrating additional sources such as business metrics (see Measuring Value). This may result in a need to upgrade tooling to leverage these data sources/inputs in the most efficient way.

Building reports

Reports should be focused, easy to understand, and should identify new spending trends. By connecting them to automatic alerting,¹¹ stakeholders are able to react quickly when necessary.

Reports should show:

- New spending trends
- Largest spenders:
 - Identify highest areas of spend by project
 - Identify highest areas of spend by product
 - Identify areas of irregular spend
- Important org-wide insights such as:
 - Anomaly detection
 - Trends over time
 - Seeing if any trends occur in a set pattern (i.e. month on month)
 - Helping compare costs between internal and external workloads (i.e. perform benchmark analysis)
 - Providing business case tracking and value realization (against on-prem)
 - Validation that GCP bills are as expected and accurate

Once the FinOps Central Team has created standardized templates, project teams can copy and then customize as per their use case.

Best practice

- Ensure the reports are able to provide data on the differing dimensions.
- The FinOps Central Team must be wary of perfection reporting relies on accurate allocation. It is better to prioritize correct but incomplete data, rather than a large volume of unreliable data.
- Consider using the customizable <u>GCP Billing Visualization Dashboard</u> that is pre-built in <u>Google Data Studio</u>.¹²
- Consider utilizing the templated GCP Cost Optimization Dashboard.
- Reporting should have as low a latency as possible.

¹¹ Alerting should not be confused as the Alerts feature in Guardrails which is a specific set of programmatic tools connected to Budgets but rather as any mechanism that triggers a notification to a stakeholder (ticket, email, etc.).

¹² The Terms Of Services for Data Studio is different from Google Cloud Platform.

Maintaining and improving reporting

Initially, dashboards will need to be maintained by the FinOps Central Team. Feedback should be encouraged to help them ensure the reports and dashboards are relevant, creating a feedback loop of trust and enthusiasm.

As maturity grows, the FinOps Central Team can work with individual business units to further add data sources, calculations, and graphs, allowing users to create and customize their own dashboard with less oversight specific to their audiences (engineer versus executive level).

Best practice

- Ensure there is a lightweight business process for surfacing feedback to the FinOps Central Team.
- Ensure there is clarity on which dashboards are centrally maintained versus what support is available for custom-build dashboards.

Inform: Forecast

Forecasting allows a customer to track their planned spend of cloud resources against their actual spend. Forecasting should encompass two broad types of data:

- Future spend based on extrapolating past spend¹³
- Future spend based on known changes¹⁴

Creating forecasts in GCP

Forecasting can be done on any time scale relevant to your business (i.e. monthly in line with sales targets, versus annualized budgets).

Google recommends that the FinOps Central Team gain familiarity with the Cloud Console <u>forecasting tooling</u>. This is a smart set of reports based on the previous cost history, which allows FinOps to predict the rate of spend in the short term.

As your organization matures, additional tools to incorporate additional types of data (e.g. data surfaced from specific change management events) may be adopted as well. <u>BigQuery ML</u> and <u>Data Studio</u> are two Google-native tools that are heavily customizable for forecasting needs and provide more advanced features (such as linear regression models). Alternatively, many third-party tools exist in the market to provide a multi-cloud view.

¹³ An example of this would be a workload with steady increase of consumption (1% MoM for last 12 months) and shows no evidence of change.

¹⁴ An example of this would be a workload going into production in 3 months and an expected increase in spend as it goes live.

Best practice

Review the tooling choices that provide forecasting abilities and decide which will be used by the business. If necessary, ensure a customized forecast report is also available when setting up reports.

Reviewing forecasts

It is important to define clear processes for cloud spend forecasts that look to be under or over their allocated budgets.

While any month that comes in under budget might initially be seen as a successful move to GCP, review whether this is due to any specific action taken by the team, or if forecasting done by the Finance Team needs tweaking. Any extra, leftover budget might then be used as incentive (see Incentivization) to encourage the right behavior.

It is also critical to define what happens in a scenario where cloud spend in a forecast is anticipated to be over budget. While the reasons behind this may be positive (more visitors and usage of the platform) or negative (an expensive service has been left on), this question should be asked in the Optimize phase when analyzing spend.¹⁵

Don't forget to define who will be monitoring the spend forecast, and what steps they should follow at given intervals – this will ensure data is being surfaced at the **right time**, to the right people.

Maintaining forecasts

It is also important to re-forecast on a regular basis. The ephemeral nature of many cloud services means that forecasts can quickly become out of date. This will depend on the organization, but monthly is a good starting point.

Best practice

- Ensure business processes are in place for projects that are forecasting ahead of budget.
- Reforecast frequently.



15 Forecasts can also be monitored using Automation (see Guardrails in Optimize for more information)

2 Optimize

The Optimize phase is concerned with taking the information that was surfaced in the Inform phase, and translating that into a set of actions to optimize your cloud spend.

There are three broad methods of making workloads in GCP more efficient,16 and these correspond to the Epics in the framework:

Product efficiency

Pricing efficiency

Guardrails

Optimize your resource usage.

Reduce the rate paid for GCP resources.

Controlling or actively limiting your spend.

These Epics form the levers you can pull to optimize cloud spend – for example, you could develop organizational best practices for using specific products or refactor a workload to take advantage of on-demand auto-scaling.

However, pulling any of these levers necessarily comes with a price-tag attached. The business has to invest effort to put in place the identified optimization, and so it is important to ensure that this effort is <u>proportional to the</u> <u>benefit generated</u>.

Any enterprise with a significant cloud presence is likely to have many possible actions that could be taken to optimize spend, so prioritizing which action to take is critical. That prioritization should be based on:

- An analysis of your existing spend
- Organizational and business objectives
- Effort versus benefit ratio

A fair analysis of those dimensions requires a cross-functional view. Business stakeholders provide the best view on business objectives and benefit, finance and FinOps can provide the best view on overall spend, while workload owners and platform teams can provide the best view on implementation effort.

16 FinOps practices advocate that usage, cost, financial, and business data be considered as a whole to allow your organization to focus on all types of optimizations – not just on cloud usage cost reduction. For the sake of simplicity, this paper will focus on optimizations you can make based on your actual cloud usage.

Prioritizing and implementing optimizations

Exactly how each organization arrives at a prioritized list can vary and there is no right answer. However, one method which may help in the initial stages is to hold a dedicated cross-functional session to look at spend and prioritize action as below:¹⁷





In the **Analyze** step illustrated in the diagram above, the FinOps Central Team analyzes cloud spend to identify relevant information prior to the cross-functional working session. As the organization's FinOps practices mature and ownership of costs is instilled, cross-functional teams will input as well.

They gather information on:

- · Progress against previous cost optimization targets (if any have been set)
- New spending trends
- Changes or anomalies in high spend projects

In the **Prioritize and Target** step, the cross-functional team uses this information, coupled with the overall business objectives, to prioritize focus areas and agree on a corresponding set of cost optimization activities. Deciding which activities to undertake should be determined based on an effort versus saving calculation. Some examples of balancing effort with spend is shown in the following diagram, drawn from the <u>GCP cost</u> optimization technical paper.



17 As the organization matures, this process will become increasingly inefficient. Optimizations are ideally identified and implemented on a frequent, continuous basis without the FinOps Central Team being involved. As a result, this session may become obsolete, less frequent or only required for very large scale, cross functional changes. It is, however, a helpful starting point to instill the right behaviors while the organization matures. Once you have prioritized a focus area and one or more activities, **targets** should be created to track your progress. They should be time-bound and metric-driven.

Tip: Don't focus targets on usage reduction

If you had the following activity identified:

• Rightsize GCP resources for over provisioned instances due to identified wastage.

An effective target might be:

Ensure that total CPU utilization averages XX% by Q3 202X. This target allows for expected growth and focuses on the average utilization as a measure of efficiency. If Compute Engine usage has gone up but the average utilization is now high, it indicates a lower level of wastage.

An ineffective target might be:

Reduce overall Compute Engine instance use. The business wants to migrate workloads, so an increase is to be expected. Seeing a trend of reducing compute usage could be the result of teams being more efficient but, equally, it could also be a result of migrations not going as planned (unrelated to cost optimization efforts). Targets which focus on reducing usage will not enable you to understand whether cost optimization activities are effective.

Finally, in the **Implementation** step, the identified optimization is actioned. The overall cadence of how often you perform these three steps depends largely on the speed at which optimizations can be made and the urgency with which they must happen. A good starting point is to hold the prioritization and target session once per month and then increase/reduce the frequency as needed.

2 Optimize: Product efficiency

The product efficiency Epic refers to all aspects of how efficiently you are employing GCP products in terms of their cost-to-benefit ratio. Whether your product usage is optimized or not boils down to whether spend can be reduced while maintaining or increasing the value delivered.

You can achieve this through either refactoring or rightsizing the workload.

Refactoring workloads

When looking to optimize usage, it's important to first consider whether you're using the right product for the workload. Setting guardrails or more efficiently deploying infrastructure is of limited use if a workload's architecture or operational model is fundamentally inefficient.

In these scenarios, <u>re-architecting</u> the workload to make use of different GCP products can be of significant value. If possible, moving to <u>serverless products on GCP</u> (like Cloud Storage, App Engine, BigQuery, and Cloud Functions) that allow scaling to zero can also help to ensure efficient usage. You can use the <u>pricing calculator</u> to compare hypothetical architectures in order to get a rough idea of respective monthly running costs. You can read more about our solutions and explore which might be best on our dedicated <u>Solutions</u> page.

Assuming you have the most optimal product for your workload based on your requirements, the next step is to consider whether your usage of that product is optimal. For example, a set of analysts which use 'SELECT *' frequently when querying large BigQuery tables may not actually require all data to be returned. If that's the case, since the service is charged per data processed, this query is inherently wasteful.

The upshot of this is that for any product which you are heavily using, you should define standards of efficient use which can be used by your workload teams. Google provides guidance on efficient usage for some products already (e.g. <u>controlling costs in BigQuery</u>).

Where possible, review these best practices, customize them as context requires, then communicate them to cloud users. Don't forget, turning VMs off when not in use is also an efficient way to use the product, save money where possible, and better for the environment!

Rightsizing

Rightsizing is a method of ensuring your infrastructure is scaled appropriately for its intended usage. The simplest example here is when your estate is oversized – you have 50 VMs, but each individual VM is not fully utilized and the workloads involved could run as effectively on less. In that case, the answer here may be to simply delete or resize some of your instances. The <u>Rightsize Recommender</u> can support this analysis and provides insights as to whether you could save money without sacrificing performance by provisioning smaller VMs.

If the product you're using supports dynamic auto-scaling, you may wish to consider re-architecting your workloads to take advantage of that to achieve a better cost and performance benefit. There are many GCP products which enable scaling up and down of infrastructure. With compute-heavy workloads, for example, an organization can take advantage of managed instance groups on GCE or by <u>containerizing workloads in Google Kubernetes Engine</u>.

Best practice

- Efficient GCP product usage is reliant on using the right product in the right way.
- The templated <u>GCP Cost Insight Dashboard</u> provides insights into efficient product use for commonly used GCP Products.
- Reliability, security, and performance are all key drivers for refactoring workloads and the relative cost-to-benefit ratio must be established with all these factors in mind. The <u>Re-architecting To Cloud Native</u> technical paper provides further guidance on this topic.

Optimize: Pricing efficiency

Outside of usage reduction, organizations can also employ measures to **increase pricing efficiency** for GCP resources.¹⁸ This can be done through GCP discounts.¹⁹

Preemptible VMs

<u>Preemptible VMs</u> are VMs which do not have guaranteed availability. Good for workloads which are both fault-tolerant and not bound by a strict time schedule due to the significant discount available.

Custom VMs

<u>Custom VMs</u> are useful when the exact size of the workload is known. By tailoring a <u>custom machine type</u> to your specific needs, you can realize significant savings.

Committed use discounts (CUD)

<u>CUDs</u> allow you to purchase committed use contracts in return for deeply discounted prices for VM usage. You can analyze the effectiveness of your CUDs by following the steps outlined <u>here</u>. <u>Shared CUDs</u> allow multiple projects to consume the same contract.

Best practice

- Taking advantage of discounts is naturally less important when GCP usage is low.
- As an organization's resource usage grows, discounts represent an increasingly significant opportunity to save money and should be actively pursued.

18 As stated in the scope of this document, this paper does not intend to cover all the ways an organization can optimize costs for each product. Rather, it sets out a framework for applying broad controls across all GCP services. In addition to reviewing the broad controls an organization can apply across GCP, organizations should also familiarize themselves with the cost optimization techniques specific to that product. For a discourse on how to optimize some of GCP's largest products, refer to the <u>cost optimization whitepaper</u>.

19 Discounts available as of March 2021

2 Optimize: Guardrails

Guardrails are passive controls in GCP that allow you to control what your organization spends.

Budgets

Enables you to track actual Google Cloud spend against planned spend.

Capping

Set an absolute limit on the amount of resources that can be consumed – implementation of capping in GCP varies between products.

Alerts

Threshold rules that are used to trigger email notifications. Budget alert emails help you stay informed about how spend is tracking against budget.

Quotas

Hard limits on resources to help control spend in case of malicious intent or misuse. Applied on a project level, per resource type and location. Refer to the productspecific documentation for more information.

Budgets & alerts best practices

- Apply default budgets and alerts to projects which you wish to cap absolutely (for example, training or sandbox projects) or to all projects at a high-water mark level as a way to catch extreme overspend.
- <u>Define budgets</u> to correspond to a financial budget you wish to track. For example, a department may have an overall cloud budget. The <u>budget scope</u> could be modified to include all projects in that particular department.
- Devolve budget and alert maintenance to workload teams to ensure that budgets can be effectively maintained.

Capping & quota best practices

- Apply <u>capping</u> in situations where there is confidence that blocking resource usage will have a negligible impact on the business. **Important note:** Capping will remove all resources and data from the associated projects. Care should be taken in how this pattern is applied.
- Commonly, it is sandbox/training environments in which this pattern is applied.

Operate

The Operate phase aims to embed FinOps practices so that they become business as usual. One-off optimization exercises may be of some value to the business, but the end goal is for FinOps to be the new normal.

To ensure FinOps practices become embedded, existing processes need to be adapted. The FinOps Central Team initially (and all cloud users eventually) need to ensure that workloads are reviewed for cost efficiency and that cost management expectations have been set during the onboarding process. Those responsible for the workload need to be educated on FinOps practices and enabled to make cost conscious decisions. The FinOps Central Team will also need to measure whether that workload is truly generating the intended value outlined in the initial business case, and use that and other centrally-defined value metrics to influence which workloads are given funding in the future.

Operate includes 5 Epics:

Onboard workloads

workloads in a cost effective manner.

Educate & enable

The ability to effectively onboard new The ability to effectively educate and enable the business in FinOps practices.

Incentivize

The mechanism by which your employees are encouraged to make cost-conscious decisions.

Measuring value

The ability to determine whether a workload is providing business value.

Funding

The mechanism by which the FinOps team supports the business in making informed decisions about which new workloads should be migrated or developed in the cloud.

As you mature, you will likely identify and require changes to processes related to teams outside the scope of this technical paper. The Finance team, in particular, is often required to embrace the Agility Culture Principle and change long-held planning, budgeting, and forecasting processes. For example:

- Review and create an adequate accounting treatment for cloud usage, then ensure it is transparent for all to understand the end-to-end financial landscape and avoid cloud cost optimizations that have little financial impact.
- Leverage the more frequent and granular cost data provided by cloud to drive insights, decisions, and actions more often.
- Document and share the new financial processes so all teams can embrace them.

Operate: Onboard workloads

Onboarding workloads includes all change management processes that help create/migrate new workloads to the cloud.

Workloads must now be cost efficient from an architectural perspective, and this should be given the same consideration as you would do towards data access or security. Allowing employees to review earlier in the cycle ensures designs are validated and best practices are adhered to before production.

To help with this aim, Google recommends that the FinOps Central Team helps to implement new acceptance criteria for workloads in the cloud, adding architectural reviews early on and throughout their development, as well as helping teams embrace the new ways of working.

Ultimately, this will drive your organization to:

- Review the workload's architecture in the early stages of the project to discuss best practices and proven patterns for optimal costs. Most delivery governance frameworks have architecture or security reviews throughout delivery. Ensuring cost reviews are also inserted into this process allows projects to demonstrate they took costs into consideration, have sound business rationale for their solution, and avoid cost shocks once in production.
- Estimate the future costs of the workload in order to enhance and update the business case. It will also feed into both the FinOps forecast (for example adding the dev/test/prod projected costs of the workload to the forecast, based on the project delivery plan) and the organization-wide Finance budgeting and forecasting processes.
- Validate any running cost projects as part of volume and performance testing that would be carried out.
- Update any existing Operational Acceptance Criteria the business may have to ensure the above points are incorporated into existing governance.

Onboarding is also a natural point where the FinOps Central Team can lend their expertise to workload owners by:

- **Recommending training** on both relevant technical topics and internal processes. For example, any product specific internal guidance which may be relevant or the cost management responsibilities of that team.
- Setting expectations as to future ongoing interactions with the FinOps Central Team and the associated community, for example in case of cost anomalies or sharing lessons learnt.
- Document effective patterns that have been used efficiently by other workload users/teams for specific use cases, and ensure these are fed into best practices and are communicated/reinforced as part of the iterative nature of FinOps (see <u>Cultural Principles</u>).

Operate: Educate & enable

Once a workload is onboarded, your teams must be continuously upskilled and aware of the latest developments and findings. This Epic ensures the organization has the right services and tools available for teams to implement and adhere to FinOps best practices.

Learning services

Learning services are the principle means by which an organization delivers formal upskilling. This service is likely in place in your organization, and will be providing on-demand, instructor-led, and hands-on training to upskill business and technical teams in cloud topics. It should also have an online presence (intranet or internal portal).

With FinOps, the aim here is to educate the organization on both the **Why** (Why does it matter? What is the impact of wastage or inefficient workloads?) and also the **How** (How can I optimize my resources? How can I see my cloud? What new technologies can make my application more efficient?).

FinOps will leverage learning services in several ways. For example, they may:

- Suggest on-demand training (from Google's catalog):
 - Understanding your Google Cloud Platform (GCP) costs (via Coursera)
 - Optimizing your Google Cloud Platform (GCP) costs
- Request or self-organize custom made training/tailored programs/events/upskilling based on observing the organization:
 - How to set up BigQuery for engineers in our organization
 - How to use BigQuery for business analysts

Training should not be limited to technical cost optimization but cover all aspects of FinOps (e.g. how to build new reports) in order to drive adoption of the practice.



Enablement services

Enablement services allow teams to engage and collaborate on specific topics. From a FinOps perspective, it includes all services and tools that cover reaching out across the company at scale in order to drive adoption of FinOps behaviors and processes. Tools and services for enablement should already be in place (an internal portal, a platform for communities, chat channels/official communication channels) and the FinOps Central Team will leverage these to achieve this goal.

Starting small and growing as the GCP footprint increases, will ensure that the effort invested is appropriate to the value generated.

Due to the heavily context-dependent nature of a successful enablement program, there are no single, one-size-fitsall patterns. Google recommends starting with:

- 1. A FinOps portal/knowledge base that includes best practices, introduction to the team, lessons learned, etc.
- 2. Access to successful architectural crowd-sourced patterns/how-to guides in a version controlled repository.
- 3. A FinOps community space that fosters collaboration and upskilling to a wider-cross functional group.
- 4. Set communication channels (newsletter, etc.) that show successes and changes made. These should be tailored to the audience to keep teams efficient and focused on value realization.
- 5. Access to GCP events (Google Cloud on-air/Google Next Conference, for example). There are many events, explanatory videos in the GCP youtube channels, webinars that announce new GCP features, and best practices to implement and utilize them.

Best practices

- Identify cost champions across all major parts of the company to spread knowledge, and ask these champions to attend internal events.
- Post on community channels and keep the enthusiasm flowing by recognizing contributions.
- Recognize teams that achieved stretched goals in established forums (like company-wide meetings).
- Ensure the portal exists; be up-to-date, searchable, and findable; and be the first place of call for GCP FinOpsrelated knowledge, widely shared across the organization.

Operate: Incentivize

The incentivize Epic aims to motivate teams to optimize their costs in relation to the value their workload provides by rewarding successful teams with tangible benefits. Your company is likely to already have some form of employee incentives in place, so focus on adding additional incentives based on how well a team achieves targets set by the FinOps team.

Work with Finance and Leadership teams to develop these incentives, using the following list as examples:

- Remind all teams and individuals about the business priorities and orient them towards maximizing these benefits, not just cutting costs.
- Agree to allow teams that save ongoing cost to reinject the savings into their change budget.
- Gamify any or all of the Epics, or compare teams KPI performance against each other to create friendly competition and an <u>open culture of trust and learning</u>.
- Adapt allocation of costs to allow for teams that created patterns to get recognition based on a percentage of cost savings achieved by other teams that used the pattern.
- · Feature teams and individuals in newsletters for visibility and recognition.
- Maintain a discreet central budget to support changes with broader impact than the individual project delivering it (e.g. funding a reusable pattern for all).
- Pay centrally for some costs for a given period, using Showback (i.e. paying for the cost, but making a point to show to all who should have paid for it to ensure all are aware of the cost they are generating for your organization).

As usual, rules should be clear, transparent, and made available to all in order to drive expected behaviors.



Operate: Measure value

The measuring value Epic is concerned with measuring the business benefits that a workload in the cloud provides against both costs and risks. Reporting that a workload costs hundreds of thousands or millions per year is useful but that figure says nothing about whether that workload is providing value.

One effective way to show the *value* is to use <u>Unit Economics</u>. This is the practice of describing cost in terms that reflect the motivations of a particular team in order to bring context to it.

For example, for an e-commerce website comparing the cost of a high-performing website (with under 5ms load time) versus unique site visits would allow a company to understand the cost to serve and make informed decisions on which infrastructure components to use. Instinctively, the choice may be to reduce the infrastructure cost, but by adding the context of the unique site visits, a company can see the value that high performance page load time delivers in terms of increased active users. This can be directly correlated and understood as conversion rate from users to buyers, and shows that high performance leads to a direct increase in dollars spent.

This metric needs to be given in the correct context so that each team understands how they contribute to the end goal and how their component parts benefit the value sought. For some teams, this may mean reducing their costs, as wastage present. For others, it may mean increasing their speed to market or reliability as the value generated is worth the cost.

Consider having a centralized value mechanism which aligns with the strategy of the business. On top of this, if needed, you could also have specific team-based value metrics which are still aligned with the overall strategy but more relevant to individual teams.

- **Centralized value metrics**: These are defined by the business and FinOps teams and based on the business strategy. These metrics should be the **single source of truth** for anyone in the business wishing to understand whether a workload is delivering business value. For example, we spent \$40,000 on this workload last month, which translates to \$4 per daily user. These metrics should be used to inform investment decisions which relate to cloud spend. (see Funding).
- **Team based value metrics:** These are defined by the FinOps team. These metrics are designed to motivate teams toward a particular goal. If you've identified that a department has high wastage, you might report wastage

in terms of the cost of a signature product your business sells. For example, say a pizza restaurant chain spent \$20,000 on idle VMs last month – or the same as an outlet selling 4,000 pizzas. Making this more relevant to the specific teams/users will provide context to why this behavior needs to change or be encouraged.

Be careful – While the latter set of metrics can be very useful in motivating a team toward a particular goal, be wary of providing metrics which encourage teams to simply reduce costs without considering the wider impact of doing so. Reducing spend in an area may result in costs shifting to *somewhere else in your business or even limiting growth*.

Operate: Funding

The last step in the cycle encourages you to reconsider the process by which investment decisions are made with respect to the cloud. At an early stage, companies typically have one business case per workload to be migrated to the cloud or to be created in the cloud. With a better understanding of the value cloud adds, the FinOps Central Team will be able to improve individual business cases, and also elevate the focus to take into consideration enterprise-wide factors and support enterprise-wide funding.

Develop more comprehensive business cases for cloud-based workloads

- Consider the benefits provided by cloud such as increased availability, increased agility, or reduced go to market time. Benefits should also be predicted against the centralized value metrics.
- Do not limit business cases to a comparison of hosting prices between on-prem and cloud.

Create an enterprise-wide cloud business case. Support enterprise-wide optimizations and large changes that are only justifiable through a broad lens. For example, consider:

- Migrating all the applications that underpin a particular business service.
- Mitigating large delivery risks through environment agility (e.g. spinning up a whole company in a test environment to test major changes).
- Exiting data centers.
- Reducing licencing tie in.
- Reducing vendor tie in.

With a strong definition of value, aligned business cases, and a strong mechanism to track value realization, you will be able to make more granular and frequent investment decisions. For example:

- As your data warehouse in the cloud proves to generate more value than the one on-prem, you may decide to accelerate its migration.
- As a new innovative service proves to generate more value than expected, you may decide to inject more funding.
- As a workload generates value faster than anticipated, you may decide to ask the owning team to switch its attention to another workload in need of optimization.

In a nutshell, leverage the frequency and granularity of value information to continuously maintain your portfolio of investments.

4 Closing the loop

Once a complete iteration has been made, the lifecycle begins again with further improvements and optimizations, gradually advancing your organization's FinOps maturity.

In each cycle, it is important to consider the progress that has been made and the overall value that FinOps has provided in order to promote successes to both leadership and the wider organization. Why is this important? Because both FinOps practices and the Central Team will need ongoing investment in order to continue generating value.

When communicating the generated value, two typical examples exist:

- Show the **similarity** of use cases across two similar business units/projects and the differences in value FinOps activities generated. This will demonstrate the potential for change and drive remediation action for the underoptimized party. Or, in the case of a comparison between on-premises and cloud workloads, justify a migration.
- Show the **risks** that are mitigated and at what cost. For example, increasing reliability by duplicating the solution in several regions while mitigating the risk of the service being unavailable for customers, further reduces the risk of the regulator fining the organization (the dollar amount could be documented), and the press highlighting the impact in the media, and your customers increasing their complaints.

Making clear what value has been provided by FinOps will allow you to appropriately scale investment and sponsorship with usage.



Next steps

When considering your next steps toward FinOps, it's important to remember that an iterative process is needed. No organization can go from zero to fully realized FinOps in a short period of time.

With that in mind, it's best to focus initially on establishing the fundamentals:

Inform

Build cost transparency and clarity through the Inform phase by implementing a consistent set of labels and some dashboards.

Operate

Educate and start to transform through the Operate phase by ensuring individuals know where to get the right training, and that the right teams are engaged.

Optimize

Implement some high-value/low-effort optimizations in the Optimize phase to build experience and engagement for future, more involved activities.

Principles

Approach each activity with the FinOps Cultural Principles in mind.



Appendix A: Organizing Epics according to capabilities

Google has also organized the Google FinOps Epics according to the capabilities they drive. See the following <u>technical paper</u>.

Building Blocks of FinOps



Cost Optimization

Provide financial visibility and recommendations of IT resources usage to optimize cloud spend

References and resources

FinOps Foundation Framework FinOps Foundation

Storment, J. R., and Mike Fuller. <u>Cloud FinOps: Collaborative, Realtime Cloud</u> <u>Financial Management</u>. 1st ed., Sebastopol, O'Reilly Media, 2019.

<u>Understanding the Principles of Cost Optimization</u> Google Cloud

Cost optimization on Google Cloud for developers and operators Google Cloud

Kubernetes: Your Hybrid Cloud Strategy Google Cloud

<u>Re-architecting To Cloud Native</u> Google Cloud