Eckerson Group | White Paper

Governing Cost with FinOps for Cloud Analytics

Program Elements, Use Cases, and Principles

KEVIN PETRIE

MAY 2023

RESEARCH SPONSORED BY

unravel™

About the Author

Kevin Petrie is VP of Research at Eckerson Group, where he manages the research agenda and writes independent reports about topics such as data integration, data observability, machine learning, and cloud data platforms. For 25 years Kevin has deciphered what technology means to practitioners, as an industry analyst, instructor, marketer, services leader, and journalist. He launched a global data analytics services team for EMC Pivotal in and ran field training at the data integration provider Attunity (now part of Qlik). A frequent public speaker and co-author of two books about data management, Kevin most loves helping startups educate their communities about emerging technologies.

About Eckerson Group

Eckerson Group is a global research, consulting, and advisory firm that helps organizations get more value from data. Our experts think critically, write clearly, and present persuasively about data analytics. They specialize in data strategy, data architecture, self-service analytics, master data management, data governance, and data science. Organizations rely on us to demystify data and analytics and develop business-driven strategies that harness the power of data. **Learn what Eckerson Group can do for you!**

About This Report

This white paper is sponsored by Unravel Data, who has exclusive permission to syndicate its content.

Table of Contents

Executive Summary

The emerging discipline of FinOps enables enterprises of all sizes to turn cloud usage-based pricing models to their advantage. They only pay for what they use, and only use what they need. Cloud-based analytics projects sorely need FinOps because their workloads can fluctuate widely, which makes it hard to budget and control costs. A FinOps program helps cross-functional teams govern cloud costs as they forecast, monitor, and account for the resources that go into cloud-based analytics projects. It helps them oversee the project lifecycle, including design, operation, and optimization. It helps control the cost of cloud-based applications as well as their underlying storage, compute, and network resources, reducing the risk of painful monthly bills. Data observability tools can assist FinOps while improving the health of data environments. Data management is the largest and fastest-growing cloud spending category, leading to increased focus on FinOps for data, also known as DataFinOps.

Implemented well, a FinOps program can drive measurable, achievable returns on investment (ROI) for cloud-based analytics projects. Cross-functional teams should adopt the following guiding principles to succeed with their FinOps program.

Key Takeaways:

- **Start with business requirements, then rinse and repeat.** Analytics leaders should design their analytics projects to stay flexible and meet business requirements that might change, even before project completion. Strategic initiatives, business priorities, and budgets might evolve slowly or change on a dime. Analytics and FinOps teams must stand ready to adjust their tools and processes when that happens.
- **Take small bites.** Like all technology programs, a FinOps program works best when implemented in phases. A FinOps team should start with one business unit or set of analytics projects and set modest, achievable objectives. If they demonstrate success with the first phase, they can gain the executive support needed to tackle bigger challenges and set more ambitious goals in subsequent phases.
- **Continuously improve.** FinOps programs depend on continuous improvement. Business leaders, analysts, and engineers all must find ways to sharpen best practices and reinforce a culture of accountability and vigilance. They must use data observability tools to drive this process with insights and automation.
- **Consider adopting a maturity model.** The FinOps Foundation offers a FinOps Maturity Model that can help cross-functional teams track program performance as they make adjustments to boost results. A model like this helps FinOps teams accelerate their journey to cloud cost governance.

The Cloud's Silver Lining: FinOps

The cloud provides enterprises with a flexible infrastructure and a rich ecosystem that foster innovative analytics. But the dynamic nature of cloud platforms also makes it hard to forecast, comprehend, and control costs. Data workloads can fluctuate widely, automatically consuming more cloud resources than expected. Traditional tools and methodologies do not provide the granular visibility that stakeholders need to predict, monitor, and control spending. Even when engineers can take useful measurements, organizational silos prevent them from collaborating and acting in a timely manner.

This leads to cost overruns that break budgets, derail analytics projects, and slow innovation. Chief data officers, heads of analytics, and business leaders need to understand this risk and to appreciate that its root cause lies in how they process and consume data. They must address that root cause as they plan and execute projects for business intelligence, data science, and data-driven applications. They must maintain this focus as part of an iterative cycle that designs, operates, and adjusts projects as they continue to migrate to the cloud.

Variable cloud workloads lead to cost overruns that break budgets, derail analytics projects, and slow innovation.

Enter FinOps

The emerging discipline of FinOps can help enterprise teams start this journey. FinOps enables IT and business stakeholders to collaborate on data-driven spending decisions that control the cost and increase the value of cloud-based projects. It instills best practices, encourages automation, and defines accountability for stakeholders that range from business and finance managers to IT and data engineers. The **FinOps Foundation** describes the ideal future state: "…everyone takes ownership of their cloud usage… [they] enable faster product delivery, while at the same time gaining more financial control and predictability."

Cloud-based analytics projects need FinOps because their workloads can fluctuate widely. A FinOps program guides cross-functional teams as they forecast, monitor, and account for the cloud resources that go into analytics projects. It helps them oversee the project lifecycle, including design, operation, and optimization. It helps control the cost of cloud-based applications as well as their underlying storage, compute, and network resources, reducing the risk of painful monthly bills.

Implemented well, a FinOps program drives measurable, achievable returns on investment (ROI) for cloud-based analytics projects. Enterprises can streamline their cloud infrastructure, increasing efficiency. They can make their business and IT teams more productive by reducing the need to fight fires, which frees up time to tackle new projects.

Implemented well, a FinOps program drives measurable, achievable ROI for cloud-based analytics projects.

Data observability. All this depends on having the right team members make the right calculations at the right points in time. This is where data observability plays a critical role. Data observability tools help monitor and optimize both data quality and data pipelines, including the infrastructure that supports them. They gather granular metadata from many elements: technologies (Apache Spark, Kafka, Airflow, etc.), services (Amazon S3, EC2, IAM; Azure ADLS, VMs, Active Directory; Google Cloud Storage, GKE, Cloud Identity), and platforms (Amazon EMR, Azure Databricks, Databricks on AWS, Google Cloud BigQuery, Dataproc, Snowflake, etc.).

Data observability tools correlate metadata such as logs, metrics, and traces from all these elements so they can understand how data pipelines perform across data sources, compute clusters, partitions, jobs, destination tables, infrastructure and applications. Based on this analysis, these tools help automatically predict, monitor, and optimize the data workloads that support analytics projects. In this way, data observability helps enterprises govern cost while improving the health of data environments.

What are Data Observability Tools?

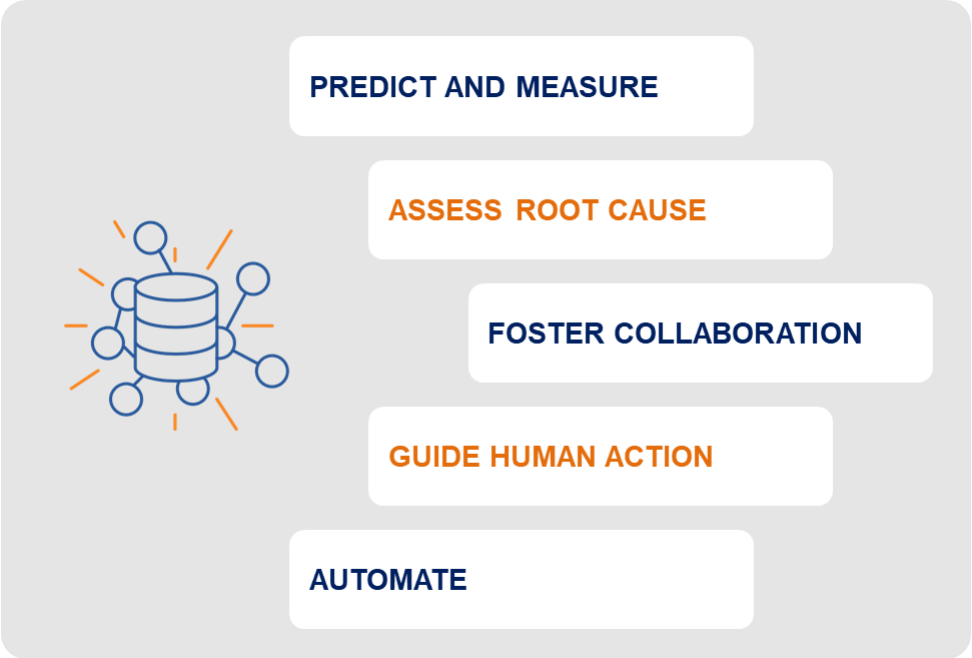
Businesses are increasingly dependent on data to provide insights, make decisions, spot trends, and predict changes. As data pipelines have grown more sophisticated, data observability tools have emerged to help data teams to visualize, comprehend, and optimize aspects of modern data sacks, including data performance, cost, and quality. Data observability tools enable data teams to not only sift through, correlate, and leverage AI and machine learning on massive amounts of metadata, logs, metrics, and billing data to provide actionable insights and automated governance needed to ensure the business reliably delivers value from its data assets.

This report seeks to educate data and business leaders about the opportunity to govern cost and improve the ROI of cloud-based analytics with a FinOps program, supported by data observability. In today's competitive environment, the opportunity is fast becoming an imperative. If you don't govern your cloud costs, your competitor will—and boost profits at your expense.

Requirements

So, what makes FinOps effective? A FinOps program must use tools and processes to meet the following requirements. (See figure 1.)

Figure 1. FinOps Program Requirements



- **Predict and measure.** Cross-functional FinOps teams need to predict what resources a cloud analytics project will consume, then measure and compare actual consumption as well as their associated costs.
- **Assess root cause.** When costs exceed a forecast—or seem likely to do so—the team needs to assess the root cause of the overrun. Perhaps a developer missed an error in her application code, a CloudOps engineer misconfigured a compute cluster, or a data science team created a sandbox without budget approval.
- **Foster collaboration.** FinOps teams depend on transparent communication and efficient handoffs between multiple roles. A typical team includes a finance manager, FinOps manager, project manager, data engineer, CloudOps engineer, data analyst, and/or data scientist.
- **Guide human action.** These stakeholders need guidance from their data observability tools as they make decisions and take actions. For example, the CloudOps engineer might need to understand the cost implications of a compute cluster configuration before they put it into production. And the data engineer might want advice about how to fix a misbehaving data pipeline.
- **Automate.** Sometimes the engineer needs her data observability tool to take automated action. She might need to fix a known issue before it impairs production workloads or have a rogue data application shut down before it breaks the monthly budget. Automated alerts, meanwhile, can help FinOps and finance managers intervene when costs exceed budget thresholds.

A FinOps program must predict and measure, assess root cause; foster collaboration; guide human action, and take automated action.

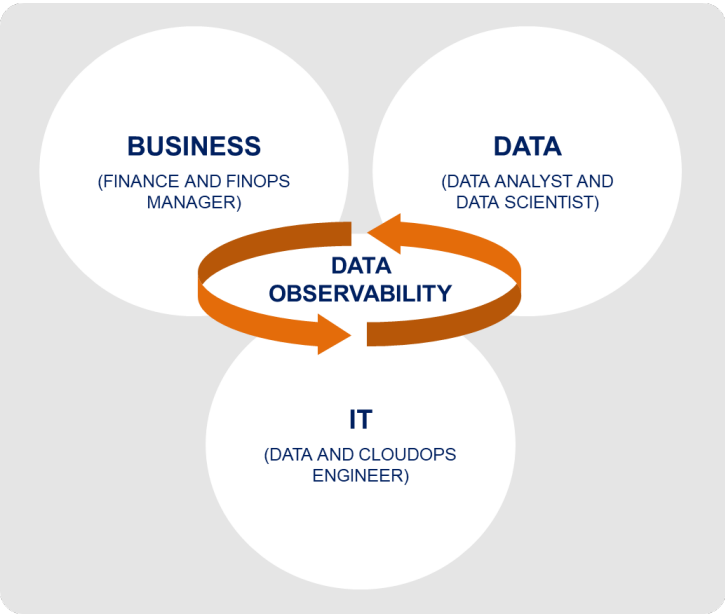
Challenges

To meet these requirements, FinOps teams must overcome a variety of challenges. They can struggle to make sense of complex data environments with interdependent elements that often change. They might subscribe to an opaque service from their cloud provider whose default auto-scaling features break through budgets. Siloed teams, processes, and tools can prevent engineers from communicating issues back to data consumers. Silos also can keep those consumers in the dark about the cost implications of their actions. Traditional observability tools might monitor operational applications but fail to shed light on the root cause of misbehaving analytics workloads.

Use Cases

The following three use cases illustrate how stakeholders use data observability tools to govern cloud costs as part of a FinOps program. We categorize these use cases as "business" (led by the finance and FinOps manager); "data" (led by the data scientist and data analyst); and "IT" (led by the data and CloudOps engineers). (See figure 2.)

Figure 2. FinOps Use Cases



Business (Finance and FinOps Manager)

Suppose the finance manager with a national chain of grocery stores oversees the budget for both the business intelligence and data science teams. The finance and FinOps managers measure actual

cloud spending vs. their budgets, then collaborate with stakeholders to fix problems. To start, a central dashboard helps them spot a project with a cost overrun. They pull up a real-time list of the most inefficient jobs for that project, including their costs and recommended actions to streamline. The FinOps managers share these recommendations with the data engineer and CloudOps engineer, who select and accept the recommendation that best fixes the cost issue while still meeting service level agreements (SLAs) for performance. The project gets back on track.

Data (Data Scientist and Data Analyst)

Now suppose a logistics company has a combined business intelligence and data science project that optimizes supply chains to handle disruptions such as new trade restrictions, shipping price increases, or factory shutdowns. To stay current with industry developments, the data scientist trains ML models on the latest industry data. The data scientist uses her data observability tool to configure an automated alert that fires when the monthly costs for a given project, job, or cluster exceed specified thresholds, or threaten to do so. When she receives an alert, the data scientist discovers that a long-running SQL query in a data prep phase is due to an inadvertent cross-joining of tables. She updates the join condition to improve query performance and unblock ML training.

IT (Data and CloudOps Engineer)

For our final use case, consider Maersk, a Danish shipping company that manages ocean and inland freight routes across the globe. Maersk's data engineering team **uses Unravel's data observability tool** to optimize pipelines, applications, and queries before pushing them into production. For example, Unravel automatically inspects Spark code in development and coaches programmers on ways to streamline or improve it. Programmers and data engineers deploy code only after predicting and approving its likely consumption of cloud resources. These techniques help Maersk analyze its global supply chain–comprising about 15 million sensor-tracked containers and thousands of ports at a given point in time–in a more timely and cost-effective manner.

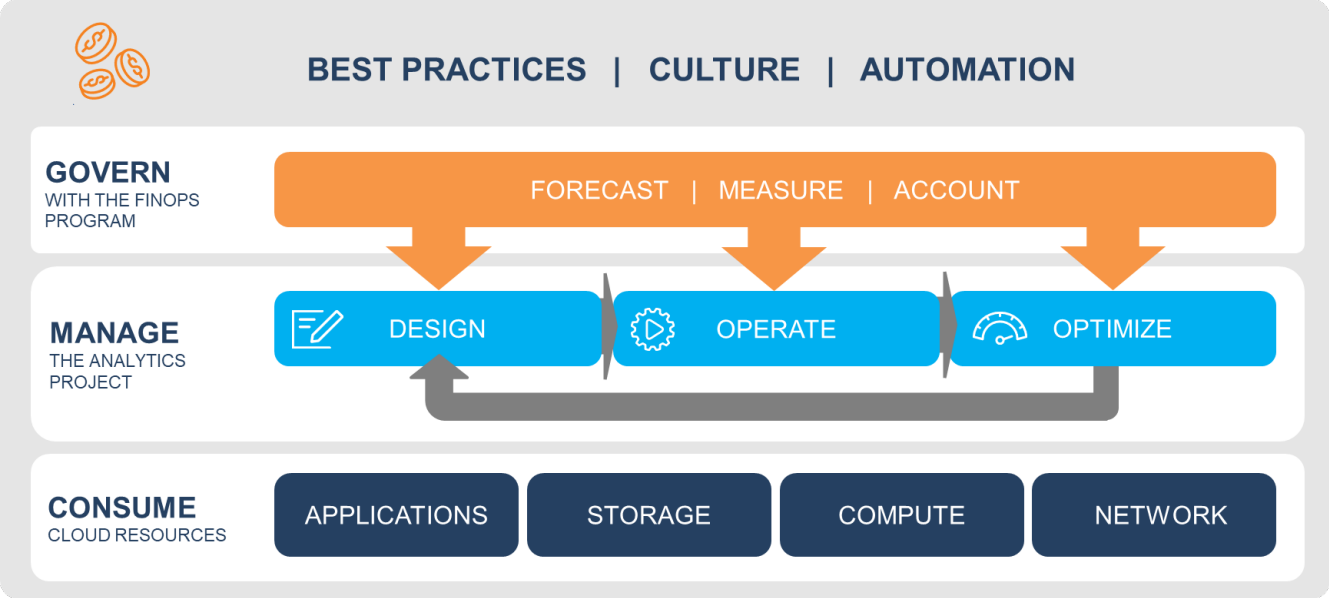
Business Impact

To understand the impact of FinOps, consider the outcome of projects that operate without it. Some do just fine and stay within budget. Other projects, however, fail with serious consequences. The Basecamp project management platform, for example, **moved back on premises** after its cloud bill broke the budget. On the positive side of the ledger, a rising number of enterprises improve the ROI of cloud-based projects by using FinOps. The retailing firm **Kroger** used Unravel's data observability product to control cloud costs, freeing up dollars to invest elsewhere, according to their VP of Data Solutions Jeff Lambert.

The Life Cycle of a FinOps Program

To put these use cases in context, let's explore the life cycle of a FinOps program and its touchpoints with cloud analytics projects. We define the primary cloud resources involved, then the key stages of an analytics project, including design, operation, and optimization. At each project stage, the FinOps program helps govern all these activities by enabling stakeholders to forecast, measure, and account for the consumption of cloud resources. We work from the bottom up in figure 3.

Figure 3. The Life Cycle of a FinOps Program



Consuming Cloud Resources

Enterprises adopted elastic cloud infrastructure to improve financial flexibility by converting capital expenditure (CAPEX) to operating expenditure (OPEX). Rather than purchasing hardware and software for their own data centers, they rent only what they need. Analytics projects consume four primary types of cloud resources: storage, network, compute, and applications.

- **Storage.** Elastic object stores such as Amazon S3, Azure Blob Storage, and Google Cloud Storage support all types of multi-structured data, such as tables, flat text files, and images, and classify them for querying with identifiers or other metadata.
- **Compute.** Cloud instances such as Amazon Elastic Compute Cloud (EC2) and Azure Virtual Machines provide virtual servers that increase utilization by having multiple users, applications, or enterprises share a given physical compute cluster. Compute charges cause a lot of the economic pain given the bursty nature of some workloads.

- **Network.** Cloud providers shuttle data across their platforms using hubs, routers, and switches. Services such as Amazon Virtual Private Cloud (VPC) and Google Virtual Private Cloud logically isolate these resources to help data teams improve security and ease of use. Cloud providers generally charge users by volume of data traffic.
- **Applications.** A range of applications, mostly software as a service (SaaS), also support analytics projects. These include business intelligence tools such as Power BI or ThoughtSpot, artificial intelligence (AI/ML) platforms such as DataRobot or Dataiku, and various operational applications that contain embedded ML models or other analytical functions.

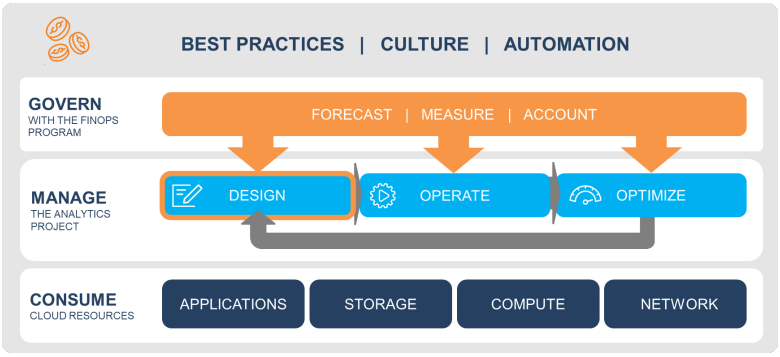
Cloud compute costs cause a lot of the economic pain given the bursty nature of some analytics workloads.

Managing the Analytics Project and Governing It with the FinOps Program

Data teams design, operate, and adjust analytics projects, periodically repeating those steps as part of an iterative lifecycle. The FinOps program helps govern, forecast, measure, and account for cloud costs at each step.

Design

The design phase has three stages: business planning, project design, and architectural design.



Business planning. First the lead data analyst or data scientist works with the business owner to build a business plan that defines their objectives and supporting use cases. They also agree on key performance indicators (KPIs). A business intelligence project might focus on dashboard timeliness and accuracy, while an ML project for customer recommendations might focus on sales win rates and average deal size. A cloud migration project, meanwhile, might focus on the KPIs of execution time, cost, and application performance on the target.

> ***Role of FinOps program:*** *the finance manager and FinOps manager review the business plan. They collaborate with the data and IT teams to build a rough* ***forecast*** *of the staff and cloud resources required and use that forecast to assess the return on investment (ROI). If the estimated ROI compares favorably with other potential projects, they approve the plan.*

Project planning. The project manager now builds the project plan. He scopes the necessary business and technical tasks, assigns owners, and sets milestones in collaboration with the data analysts and/or data scientists that will perform the actual analysis. The project plan also scopes the responsibilities of the data engineer that manages the data environment and CloudOps engineer that manages the cloud infrastructure.

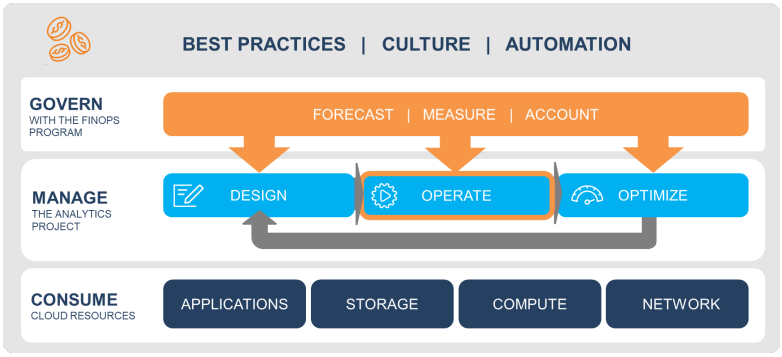
> ***Role of FinOps program:*** *the FinOps manager and project manager fine-tune the* ***forecast*** *based on a close review of the project plan in collaboration with the data and IT teams. They define granular KPIs that will* ***measure*** *the usage of cloud resources and application performance, then identify which stakeholders to hold* ***accountable*** *for those costs.*

Architectural design. The data engineer and CloudOps engineer now use the high level design from the business and project planning stages to produce a detailed data architecture that supports the analytics or migration project. This architecture includes data sources such as relational databases, SaaS applications, or IoT sensors; targets such as BI tools or AI/ML platforms; and pipelines that ingest and transform data between them. In many cases they can implement this design by integrating or reconfiguring architectural elements that already exist in their environment.

> ***Role of FinOps program:*** *the data engineer and CloudOps engineer help the project manager and FinOps manager fine-tune the* ***forecasts*** *further for each cloud resource, application, and accountable stakeholder. Together they perform a sensitivity analysis to understand the impact of different risk factors on estimated ROI for the project.*

FinOps programs help forecast, measure, and account for each stage of project design.

Operate



Show time! The team puts the plan into operation. The data analyst might generate weekly business intelligence reports or maintain a real-time dashboard. The data scientist might build and test ML models for customer recommendations or fraud detection. The data engineer might kick off his migration project and start moving applications and data to the cloud. All the while they consume data and cloud resources that the data

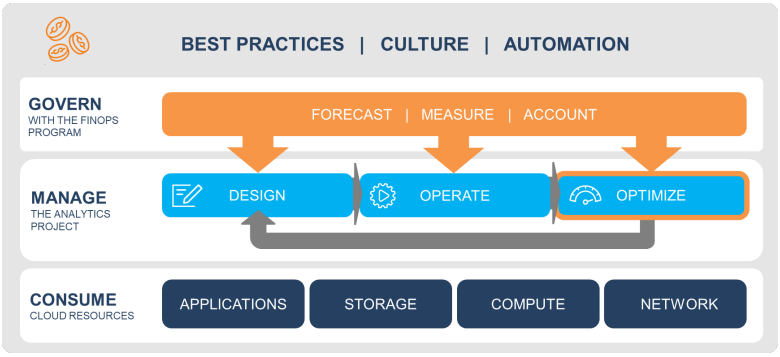
engineer and CloudOps engineer maintain. The project manager oversees team members as they collaborate to complete milestones and hand off tasks.

> ***Role of FinOps program:*** *the FinOps manager collaborates with each stakeholder to* ***measure*** *the consumption of cloud resources against* ***forecasts****. She helps the data engineer measure how data pipelines consume storage, compute, and network resources. She helps the CloudOps engineer track how applications and analytical tools consume those resources, as do the data analyst and data scientist. The FinOps manager* ***accounts*** *for all these resource costs in her updates back to the finance manager, flagging any risks or overruns.*

The FinOps manager helps the data engineer and CloudOps engineer control cloud consumption as they operate the analytics project.

Optimize

Team members optimize the project by monitoring overall progress and adjusting activities as they go. The data analyst invites feedback from business owners on his reports or dashboards, then adds or removes elements based on what they hear back. The data scientist observes production ML models for signs of data drift, then intervenes to re-train or swap out models. The data engineer tracks data pipeline performance and the CloudOps engineer tracks cloud resource utilization and performance, reconfiguring or tuning resources as needed. In all these cases stakeholders compare status and progress against KPIs, reporting back to the project manager.



> ***Role of FinOps program:*** *This is where the elasticity of cloud infrastructure becomes an advantage. FinOps team members use their newfound knowledge to optimize project costs. The data analyst and data scientist* ***forecast*** *the consumption of cloud resources so they can adjust projects to minimize risks and prevent overruns. They might consult with the data engineer to filter out unnecessary data and streamline queries, then* ***measure*** *the reduction in cloud compute cycles. They might consult with the CloudOps engineer to reschedule jobs or rebalance workloads, and again* ***measure*** *the impact on cloud compute. As before, the FinOps manager* ***accounts*** *for these cost adjustments in her updates back to the finance manager.*

Wrapping It Up

Best Practices

Each of the key FinOps actions described above should become a best practice. By documenting standardized and repeatable best practices, teams can make their FinOps programs more efficient and effective. Best practices reduce the risk of a FinOps program and enable managers to build FinOps responsibilities into their team members' job descriptions. They also help established FinOps teams "spread the gospel" by training new teams and contributors. Sound best practices hold consistent across industries, helping minimize waste and improve the bottom line.

Culture

In many ways, the success of a FinOps program hinges on culture: i.e., the norms, rituals, and stories that influence the actions of team members. For example, FinOps team members should follow the norms of always minimizing waste and never making decisions in isolation. A data scientist should adopt the ritual of always asking CloudOps engineers to validate their cloud-compute assumptions. A data engineer, meanwhile, should adopt the ritual of always forecasting cloud consumption before putting a pipeline into production.

All team members should tell stories that reinforce these norms and rituals. For example, executives should showcase and celebrate team "wins," such as an analytics project that avoided certain cloud costs or exceeded ROI forecasts. Norms, rituals, and stories like these can shape a successful FinOps culture and program.

Automation

FinOps teams also can reduce the manual effort of controlling cloud consumption via automation. A variety of commercial tools can assist here, including data pipeline and orchestration tools such as **Fivetran** or **dbt**, workflow tools such as **Airflow**, and work notification systems such as **Slack**.

In addition, data observability tools can automatically predict issues, assess root causes, and recommend actions to prevent or remediate those issues. They also can implement automated governance processes and guardrails. In each of these cases data engineers, analysts, scientists, or operations teams play the role of overseer. They kick off automated tasks, receive the results, then review AI-driven recommendations about what to do next. As described earlier, these recommendations can get specific, for example by prescribing the fastest or least expensive way to run a pipeline and predicting the likely impact on KPIs such as compute cost or response times.

Data observability tools help FinOps teams reduce the manual effort of controlling cloud consumption.

Conclusion: Guiding Principles

Implemented well, FinOps programs help enterprises turn cloud providers usage-based pricing models to their advantage. They only pay for what they use, and better yet they only use what they need. This improves profitability, stability, and confidence in new strategic initiatives. Cross-functional teams should adopt the following guiding principles to succeed with their FinOps program.

- **Start with business requirements—then rinse and repeat.** Analytics leaders should design their analytics projects to stay flexible and meet business requirements that might change, even before project completion. Strategic initiatives, business priorities, and budgets might evolve slowly or change on a dime. Analytics and FinOps teams must stand ready to adjust their tools and processes when that happens.
- **Take small bites.** Like all technology programs, a FinOps program works best when implemented in phases. A FinOps team should start with one business unit or set of analytics projects and set modest, achievable objectives. If they demonstrate success with the first phase, they can gain the executive support needed to tackle bigger challenges and set more ambitious goals in subsequent phases.
- **Continuously improve.** FinOps programs depend on continuous improvement. Business leaders, analysts, and engineers all must find ways to sharpen best practices and reinforce a culture of accountability and vigilance. They must use data observability tools to drive this process with insights and automation.
- **Consider adopting a maturity model.** The FinOps Foundation offers a **FinOps Maturity Model** that can help cross-functional teams track program performance as they proceed along a "crawl, walk, run" development curve. Stakeholders can tweak tools, techniques, or processes to improve certain metrics–perhaps percentage of cloud cost allocation or spending variance vs. forecasts–and measure the results of their change right away. A model like this helps FinOps teams accelerate their journey to cloud cost governance.

About Eckerson Group

Wayne Eckerson, a globally-known author, speaker, and consultant, formed **Eckerson Group** to help organizations get more value from data and analytics. His goal is to provide organizations with expert guidance during every step of their data and analytics journey.

Eckerson Group helps organizations in three ways:

- **Our thought leaders** publish practical, compelling content that keeps data analytics leaders abreast of the latest trends, techniques, and tools in the field.
- **Our consultants** listen carefully, think deeply, and craft tailored solutions that translate business requirements into compelling strategies and solutions.
- **Our advisors** provide one-on-one coaching and mentoring to data leaders and help software vendors develop go-to-market strategies.

Eckerson Group is a global research and consulting firm that focuses solely on data and analytics. Our experts specialize in data governance, self-service analytics, data architecture, data science, data management, and business intelligence.

Our clients say we are hard-working, insightful, and humble. It all stems from our love of data and our desire to help organizations turn insights into action. We are a family of continuous learners, interpreting the world of data and analytics for you.

Get more value from your data. Put an expert on your side. **Learn what Eckerson Group can do for you!**

Eckerson Group

RESEARCH CONSULTING ADVISORY

GET • MORE • VALUE • FROM • YOUR • DATA

About the Sponsor

Unravel Data radically transforms the way businesses understand and optimize the performance and cost of their modern data applications—and the complex data pipelines that power those applications. Providing a unified view across the entire stack, Unravel's market-leading Data Observability platform leverages AI, machine learning, and advanced analytics to provide modern data teams with the actionable recommendations they need to turn data into insights. Some of the world's most recognized brands like Adobe, 84.51˚ (a Kroger company), and Deutsche Bank rely on Unravel Data to unlock data-driven insights and deliver new innovations to market.

To learn more, visit **www.unraveldata.com**.