Cost Measurement and Optimization of Cloud Computing While Enabling Virtualization of IP Network Functions

A Technical Paper prepared for SCTE by

Ajay Gavagal
Data Solutions Architect
Comcast
1800 Arch Street, Philadelphia, PA 19103
303.658.7826
ajay_gavagal@cable.comcast.com

Mehul Patel
Distinguished Architect
Comcast
183 Inverness Dr. W, Englewood, CO 80112
303-658-7826
mehul_patel@cable.comcast.com

Sinan Onder
Vice President, Software Architecture
Comcast
1800 Arch Street, Philadelphia, PA 19103
267-260-0964
sinan_onder@comcast.com

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

List of Tables

Title **Page Number**

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

1. Introduction

IP networks are designed to support a wide range of services, including video streaming, high-speed data transmission, voice communication, and more. However, any advancements or updates in either software or hardware necessitate the costly and time-consuming process of upgrading vendor-specific appliances deployed in these networks, posing significant challenges for network operators.

In recent years, network operators have critically reassessed the conventional approach of relying on vendor-specific appliances that tightly integrate hardware and software to deploy and operate their networks. As a result, the concept of virtualizing network functions has gained widespread acceptance among network operators. Virtualization allows specific functions of the network to be decoupled from proprietary hardware and instead implemented as software-based on virtual instances. These virtual instances utilize commercial off-the-shelf (COTS) hardware and open new possibilities for network operators. The virtual instances could either be in a private or a public cloud (consider on-demand scaling, disaster recovery scenarios). Management and operation of networks requires a whole host of software applications that can be run in a private or a public cloud as well.

Network operators have gained several advantages while virtualizing network functions. The advantages of this approach encompass both hardware and software deployments. However, in this paper, our focus will solely be on the software component, specifically on its cloud cost measurement and optimization. The authors have decided to reserve an in-depth examination of the hardware component for a separate study.

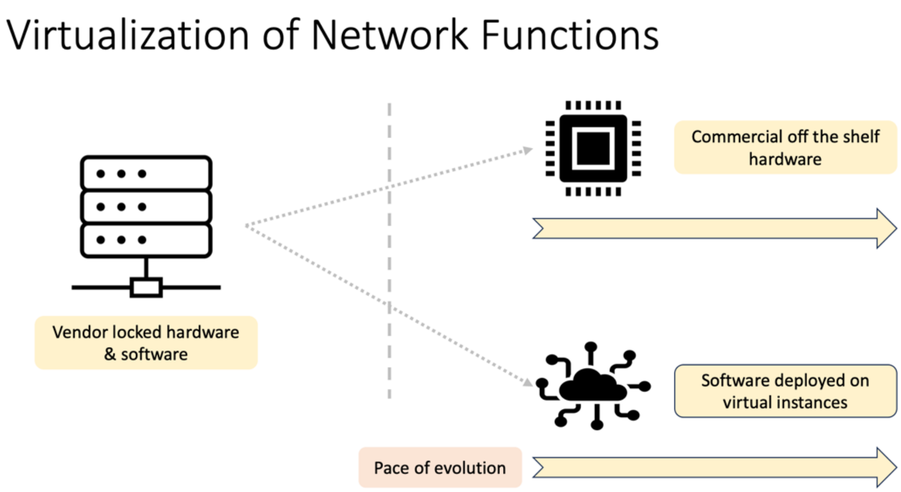

Figure 1 - Overview of Virtualization of Network Functions

Virtualizing network functions provides network operators with enhanced agility and flexibility. Operators can now rapidly deploy and scale software by creating or removing virtual instances as per demand. On the other hand, hardware deployment is also faster due to the use of the COTS model. Network operators use simple and multipurpose hardware. This approach of scaling software and hardware individually significantly improves the responsiveness of network operators to quickly adapt to changing customer needs and market demands.

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

With the virtualization of network functions, network operators can fine-tune their networks to specific business demands and capabilities allowing for diversification among operators. Examples of diversification include scalability and elasticity, vendor choices for hardware, and flexibility to manage software release rate without the rigidity of hardware refresh cycles. This diversification also enables the pace of innovation and empowers operators to respond to market trends and customer preferences in a more nuanced approach, ensuring they can stay competitive in the dynamic telecommunications landscape. Imagine if all network operators were stuck with only vendor-dictated options for both hardware and software, this would impact their level of fine-tuning and the level of diversification in the market they can achieve.

Furthermore, the virtualization of network functions offers substantial cost efficiencies since bulk cashflow outlays can be avoided. By decoupling the network functions from dedicated hardware, operators can speed up or slow down software and hardware upgrades separately and at their own comfortable pace. Also, avoiding vendor lock-in, giving network operators a broader range of selection for both software and hardware vendors while avoiding costly upgrade cycles.

Network operators also gain a lot of speed and agility specifically in software development and deployment. Virtual instances can be provisioned dynamically, which automatically reduces the overall time for software releases introducing new features or fixing bugs. This does not mean we are excluding other contributing factors to the increase in speeds of software deployments such as the adoption of agile methodologies, the introduction of various new technologies, and hyperscaler's introducing higher order services.

Now that we have seen the benefits of virtualizing network functions, let's focus on some of the challenges that have been brought in through this model for network operators, specifically in the software component when using virtual instances in either a private or public cloud.

The introduction of virtualization brought about a shift in the pace at which software could be deployed when using cloud. However, this advancement also resulted in the removal of certain advantageous constraints that were previously imposed on software deployment due to financial considerations. To gain a more comprehensive understanding of this matter, let us delve further into the traditional approach that was prevalent before the use of the cloud.

In the past, software development teams, typically centralized within an IT department, engaged in extensive negotiations with the finance department to obtain the necessary funding for developing new software or upgrading existing releases. These resources, once approved, went through a typically long cycle of identification, purchase, and deployment prior to any software development activities. Software development teams were required to present a compelling case to the finance department, justifying the introduction or modification of a business capability and the requirement of resources for the same. This intricate process, although time-consuming, ensured that the deployment of resources was subject to meticulous evaluation, considering questions like "how does this addition or change in business capability impact the overall strategic value?"

However, with the advent of provisioning virtual instances in a cloud for development and deployment purposes, software development teams, and more specifically engineers, gained a considerable degree of autonomy away from finance and justification. They were now empowered to freely provision instances as needed to meet their software requirements, without giving the same degree of consideration to the value associated with each deployment with respect to the business capability or eventually the business activity it is supporting.

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

This predicament was further exacerbated for finance. Suddenly, the traditional capital expenditure associated with resource provisioning began to transform into operational expenditure, requiring the finance department to adapt and prepare for this notable shift. Additionally, the visibility and control that finance teams once had, enabling them to identify the purpose of a provisioned instance for a specific project, became exceedingly challenging to attain from private or public cloud environments. Consequently, when the strategic viability of a project diminishes, finance teams encounter difficulties in enforcing rigorous clean-up or repurposing measures for the associated instances. This lack of visibility and control is prompting finance teams to seek answers regarding where to initiate their efforts, sparking the rise of a new discipline known as FinOps, or financial operations, which has gained substantial traction within the industry. Aptly so, FinOps Foundation executive director J. R. Storment says, noting that the "cloud removes finance from the buying process and hands the credit card to cloud engineers."

This paper partly concentrates on the impact of transitioning from a capital expenditure model to operational expenditure model on finance during software operations, neglecting any discussion regarding the impact on finance during software development. However, it is worth noting that the problem of transitioning from capital expenditures (CapEx) to operational expenditures (OpEx) due to the adoption of cloud also persists during software development, necessitating a comprehensive study on its own.

When considering the convergence of two significant trends, the challenges faced by network operators become evident. In a landscape where virtualizing network functions and embracing cloud deployments for enhanced speed, adaptability, and innovation are standard practices, the questions arise:

1. How can a network operator effectively manage and control the software operational expenses that accumulate?
2. What strategic value can network operators add to their portfolio by being on top of such operational expenses?
3. Is all this effort to go after these expenses even worth the time & effort?

In this paper, the authors dedicate some amount of time to discuss all the issues there are in cloud cost measurement, interpretation, and optimization in a cloud environment and suggest some strategies on how best to get a hold on the situation to avoid overrun of operational expense.

2. Death by a thousand cuts

Before delving into the crux of this white paper on cloud cost measurement and optimization, it is imperative that we draw a compelling parallel between the financial dynamics of cloud expenditure and the purchasing tendencies exhibited by an average consumer when acquiring software applications through subscription models, often resulting in underutilization or complete non-usage of resources.

The new landscape of consumer behavior has witnessed a significant shift, with consumer product companies opting to adopt the subscription-based paradigm over the conventional approach of selling software products at a one-time price which is not affordable to a sect of consumers. This change can be observed across a spectrum of software apps, spanning from operating systems tailored for smartphones, laptops, and tablets to a vast array of applications encompassing video streaming platforms, instant messaging services, video conferencing tools, gaming applications, and an extensive repository across other categories.

Under a software subscription-based model, consumers are frequently offered an enticing assortment of software applications to cater to their needs. In this environment, the cost of individual applications seems inconsequential when contrasted with the recurring stream of modest subscription fees. However, this cost-

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

effective shopping spree can quickly lead to an accumulation of unwanted, duplicate applications that are either utilized scarcely or, in certain instances, not used at all.

Drawing a direct parallel to the world of cloud computing and its costs, organizations often find themselves with similar purchasing habits, except on a large scale. The benefit of the cloud, and its promise of scalability, flexibility, and pay-as-you-go pricing, make organizations provision a multitude of resources, services, and infrastructure without extensive thought into their strategic value, actual utilization, or the full cycle from provisioning to decommission. Just as an average consumer may accumulate a whole host of unused or underused software apps, organizations can find themselves with an unwieldy cloud bill with no insight into redundancy, inefficiency, and financial wastage.

The procurement process for cloud services is equally intricate compared to the traditional software and hardware purchase model. While the pricing structure may appear transparent and based solely on operational expenditures (OpEx), cloud providers offer diverse options to assist consumers, which can nonetheless result in cost opacity. For instance, consumers may be able to opt to make an upfront payment for a subscription fee, enabling them to secure a significant discount on the service. Or bulk purchases may be available, and which may result in substantial discounts. Furthermore, consumers frequently have a wide range of choices when it comes to purchasing licenses on a subscription model or accessing higher-order services for a higher price to meet their specific requirements. All these factors significantly complicate the decision-making process of consumers during both software development and operations. The decision-making process of choosing the right resources falls entirely on the software engineering team due to the autonomy of cloud provisioning.

The authors call the problem of having to deal with a thousand details of choosing and using the right resources as "death by a thousand cuts" to an average consumer or enterprise.

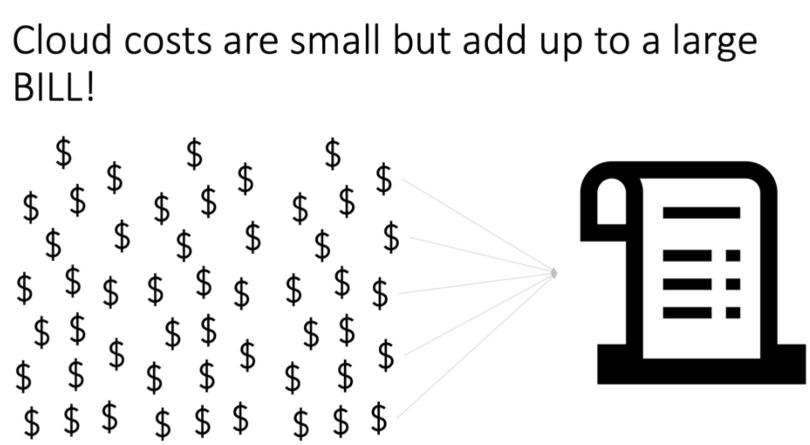

Figure 2 - Small Charges Leading to a Large Bill

Looking at this striking similarity between consumer app subscriptions and cloud costs, it becomes evident that organizations must exercise extreme caution and adopt a proactive approach toward cloud cost measurement and optimization. Failure to do so can result in substantial financial resources that are squandered on underutilized or unnecessary cloud resources, negating the benefits and advantages offered by the cloud.

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

3. Strategy Outline

Addressing cloud expenditure shares similarities with other strategies for cost optimization. Organizations can pursue several key initiatives to regain control over their cloud costs:

1. Measure cloud costs per business capability and eventually per activity
2. Minimize wastage, including the mitigation of underutilization.
3. Engage in comprehensive evaluations with cloud providers, which involve:
    a. Negotiating pricing for services and infrastructure.
        i. Capitalizing on bulk purchase discounts.
    b. Engaging in best practice discussions.
    c. Tracking the roadmap to prevent redundant efforts.
    d. Submitting feature requests that facilitate cost reduction.
4. Re-architect applications and their components to diminish redundant computing and storage.
5. Perform accurate forecasting to anticipate future requirements.
6. Embrace a proactive optimization approach for future needs.

When examining the above overarching strategies, it becomes evident that **“Measurement and Observability”** of cloud costs are fundamental prerequisites for successfully implementing any of these initiatives. Another crucial factor for achieving success in a substantial undertaking like this is to ensure unwavering commitment from every individual involved throughout the organization. Furthermore, it is important to emphasize that this commitment should not be treated as a one-time event, but rather as a continuous and ongoing effort.

4. Cloud cost measurement

Before delving into the measurement of cloud costs, let us first take a step back and examine the types of observabilities we can establish during software design, development, and deployment. The authors propose that there are primarily two types of observabilities that can be established:

1. Observability during the design phase: This approach takes a proactive stance, treating cloud cost observability as a design principle and a fundamental pillar of sound software design. By incorporating observability into the design stage, it becomes possible to comprehend software behavior, identify areas for optimization and debugging, determine where and how to scale, and more. Employing approaches such as enhancing telemetry to help finance measure activities during development vs. operations and their respective cloud costs is one way to achieve this. Cost tagging is another approach. However, achieving this proactive observability approach is challenging in reality due to factors such as priorities, escalating requirements, tight deadlines, and limited budgets. Consequently, the authors defer a comprehensive examination of implementing proactive observability for cloud cost measurement to a future study.

2. Observability after development and deployment: This approach is more reactive in nature, as it involves responding to cloud costs after they have already been incurred rather than anticipating and addressing them proactively. Unfortunately, many enterprises still lack a solid foundational approach to measuring and observing cloud costs even in a reactive approach, resulting in uncoordinated responses. Nevertheless, there is good news as establishing such an observability platform is relatively straightforward to establish with existing tools and processes. Therefore, the authors choose to focus on establishing a robust measurement and observability platform—in a

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

reactive approach focusing only on software operations within this paper. Additionally, the authors plan to conduct a more detailed study on how to react to observed cloud costs in the future.

Now let's explore the intricacies of cloud cost measurement. Through the upcoming sections, we will examine:

1. The different sources of information available for cloud cost measurement.
2. Challenges that arise during the data collection process.
3. Analyzing the extent to which this information is digitized.
4. How this information is interpreted within the organization.
5. Finally, we will delve into the incentives that drive participation in cost measurement and optimization efforts, as well as the importance of effective communication channels and the frequency of communication.

4.1. Measurement & Observability End State

The authors have a vision for the end state of the measurement and observability platform, which is depicted below. However, before we present the desired end state, it is crucial to establish a clear understanding of all the terminologies employed in the accompanying diagram. We recommend linking these terminologies with the ones familiar to your enterprise, as the terminology might differ in wording but ultimately converge on the same meaning.

Terminologies & their meaning:

a. Business Activity: Any activity that helps make money for the business.
b. Financial Budget: Spending plan established based on income and expenses of the business.
c. Business Capability: Capabilities or functions required by a business to achieve a business activity.
d. Service: One or more software services established to achieve a business capability.

The remaining terminologies are straightforward and will not be elaborated upon. Now that we have introduced the terminologies, we dive into what the measurement and observability platform end state might look like:

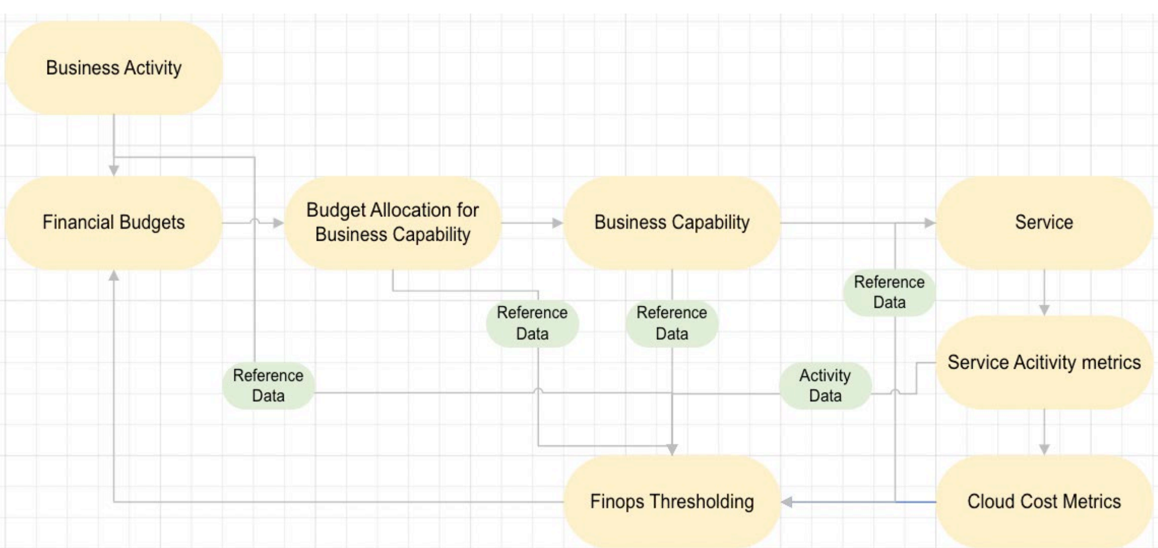

Figure 3 - Vision for Cloud Cost Measurement and Observability End State

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

4.2. Sources of cloud cost data & its related elements

Looking at different sources of information available to enable a successful cloud cost measurement.

Table 1 – Sources of Cloud Cost Data & Its Related Elements

Data Name	Source of Data	Usage of data in cloud cost measurement	Size of Data	Data available for consumption in digitized form?
Software Service Name, Owner, Contact, Purpose, etc.,	Service portfolio management systems	Categorization of cloud costs per service within an organization	Small	Only for software services that are mature
Business Capability Name, Owner, Contact, Purpose, Software application supporting the business capability, etc.,	Business Capability Maps	Categorization of cloud cost data per business capability	Medium	Only for mature business capabilities
Business Activity Name, Owner, Business Capabilities supporting the specific business activities, etc.,	Enterprise Resource Planning systems (ERP)	Categorization of cloud costs per business activity	small	Yes
Cloud costs in detail per service used	Cloud provider usually provides a tool that tracks costs per service, tag, etc.,	Total cloud cost reporting by service, tags, etc.,	Large	Yes
Financial budgets allocated for cloud spend per account and per organization	Some sort of ERP systems	Tracking actual spend vs. budget to influence cash flow requirements	Small	Yes
Thresholds on cloud spend per service, service	In-house built systems to store such information	To provide notifications in a reactive	Large	Maybe

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

limits, contractual agreements etc.,	or distributed across individuals systems	environment on cloud costs to finance, engineers etc.,		
Contact information in case of exceeding thresholds, escalation paths etc.,	Workflow management software	Notify the right contacts to resolve breaches in cloud service limits, spend etc.,	Medium	Yes
Cloud resource inventory	Systems that maintain all the cloud resources deployed within an organization	Cloud centers of excellence to understand cloud usage patterns across an organization	Large	Yes

By referring to the preceding table, it becomes evident that a wide range of services and information systems are essential for facilitating the collection, ingestion, and persistence of data pertaining to cloud costs.

4.3. Guidelines for establishing foundational datasets

If you are embarking on the journey of cloud cost measurement for the first time, laying the groundwork with foundational datasets can significantly impact your ability to analyze data and identify opportunities for optimization. The authors suggest a selection of foundational datasets presented in tabular form, which serves as a starting point rather than an exhaustive list.

4.4. Business Capability Dataset

Table 2 – Business Capability Dataset

Column Name	Description
Business Capability Identifier	Identifier for a Business Capability
Business Capability Name	Name of the Business Capability powering a business activity
Business Capability Description	A short description of the Business Capability
Business Capability Owner	Owner of the Business Capability
Business Capability Owner Contact Information	Contact information is meant to identify how a Business Capability owner should be notified

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

4.5. Software Service Dataset

Table 3 – Software Services Dataset

Column Name	Description
Service Identifier	Identifier for a software Service
Service Name	Name of the Service as identified and recognized by consumers
Service Description	A short description of the Service, its intent
Service Owner	Owner of the Service
Service Owner Contact Information	Contact information is meant to identify how a Service owner/s should be notified

4.6. Business Capability to Software Service Mapping Dataset

Table 4 – Business Capability to Software Service Mapping Dataset

Column Name	Description
Service Identifier	Identifier for a software Service
Business Capability Identifier	Identifier for a Business Capability

4.7. Service Component dataset

Table 5 – Service Component Dataset

Column Name	Description
Service Identifier	Identifier for the Service that component belongs to
Service Component Identifier	Identifier for the Service within the application
Service Component Name	Name of the Service component as identified and recognized by developers
Service Component Description	A short description of the Service Component, its intent

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

4.8. Cloud service dataset

Table 6 – Cloud Service Dataset

Column Name	Description
Cloud service Identifier	Identifier for the cloud service
Cloud service name	Name of the cloud service used
Cloud service description	A short description of the cloud service

4.9. Cloud service component cost dataset

Table 7 – Cloud Service Component Cost Dataset

Column Name	Description
Cloud service identifier	Identifier for the cloud service
Cloud service component Identifier	Components of the cloud service where cost is assigned
Unit cost for the component	Unit cost here refers to the kind of unit the cloud provider has assigned to measure the usage of a service component

4.10. Financial budget

Table 8 – Financial Budget

Column Name	Description
Financial budget identifier	Identifier for the set budget
Business Capability Name	Name of the Business Capability
Business Capability financial budget	Set budget for the Business Capability
Maximum allowed deviation	Deviation allowed from budget beyond which triggers need to be kicked off
Finance owner	Owner from finance who is in charge of budget allocation
Finance owner contact information	Contact information for finance owner

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

4.11. Cloud Usage Dataset

Table 9 – Cloud Usage Dataset

Column Name	Description
Cloud Service Usage identifier	Self-explanatory
Cloud Service Usage Metric Name	Self-explanatory
Cloud Service Usage Metric	Brief definition of the cloud service usage metric
Free tier	Cloud service providers might offer free tier usage to promote their services
Minimum usage required	Minimum usage required by the cloud service provider to ensure the rates provided justifies their cost

5. Data transformation, storage & accessibility

In order to establish the foundational datasets, as mentioned previously, a critical step involves identifying the appropriate source systems within the enterprise. Once these systems are identified, the next requirement is to implement a robust big data processing and storage platform. It is important to note that when dealing with metrics related to cloud service usage and costs, the volume of data can reach staggering proportions, easily amassing billions of records. Given the enormity of these datasets, processing them effectively necessitates a substantial number of clusters and considerable processing power to handle the computational demands.

Below is an application-level view of the entire pipeline when set up:

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

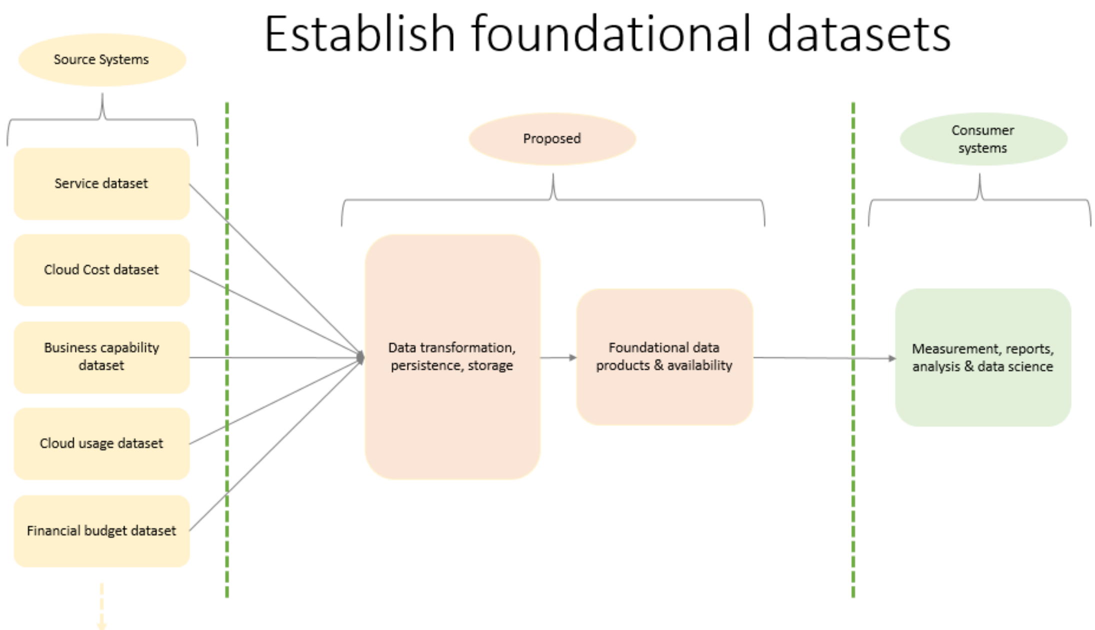

Figure 4 - A Look at End-to-End Datapipeline for Cloud Costs

Once such an environment is established, network operators gain the capability to measure cloud costs from multiple perspectives.

Let's explore some of the questions that can be addressed through this setup:

1. For Leaders/Finance:
    - What does the overall cloud cost look like across the organization?
        1. Breakdown of costs by business capability and business activity.
    - Which business capability requires focused attention due to escalating cloud costs?
        1. Identification of business capability owners
    - Determination of services associated with this business capability.
2. For Architects/Engineers:
    - Which components of applications are contributing to exceeding of cost thresholds?
    - Can we consolidate components from different applications to create higher-order services? (Cost serving as a justifiable factor for strategic discussions).
    - What are the most significant opportunities for optimization?
3. For Product Owners/Program Managers:
    - Which services within my product/portfolio account for the highest cost expenditure?
    - Can the cost breakdown be specified by individual application components, aiding in the formulation of user stories for further investigation?

These questions not only pave the way for individual optimization efforts but also open avenues for optimization from various angles.

6. Issues arising while consolidating cloud cost data

Having familiarized ourselves with the wide array of diverse information accessible for establishing a robust framework for measuring cloud costs, it is now imperative to shift our focus towards the potential

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

challenges that may arise when consolidating such data to derive meaningful insights, which can aid in optimizing efforts and conducting benefit analysis.

1. Determining the appropriate level of resolution in data is consistently a challenging problem to solve. This challenge becomes particularly evident when comparing the resolutions between business capabilities and the underlying software that enables them, as opposed to the measurable cloud cost per software.
2. Digitized and readily accessible mappings between software services and the products they support are rarely available.
3. Data sources may be dispersed across different organizations, teams, silos, and geographical locations. This inherent distribution complicates the process of establishing a robust measurement framework.
4. Data sources can exist in diverse systems with varying storage patterns, including legacy systems with limited data accessibility and raw data formats.
5. The availability of contextual information alongside the data may be limited or entirely absent. For example, when a cloud cost for utilizing a serverless function is reported in "X" units, the interpretation and meaning of those "X" units might be unclear or undisclosed.
6. Cloud providers frequently alter their pricing models and the metrics they employ for charging. Failure to account for these modifications in the measurement process, especially when using non-digitized input methods, can lead to a significant increase in the effort required for cloud cost measurement.

7. Optimizing Cloud Costs

Cloud cost optimization is a vast and intricate topic that necessitates thorough examination. Although this paper only scratches the surface, the authors have tried to explore key concepts related to the cloud cost optimization process. They discuss the challenges that may be encountered during such efforts and offer a high-level insight into automating the entire process—from source to measurement to optimization. Without further ado, let us delve into the topics.

The optimal approach for conducting an optimization exercise involves leveraging measured data and the insights it offers. But after leveraging the measured data, there is a whole hosting of convincing the engineers to act. To illustrate this point further, the authors present two examples:

7.1. Scenario - 1

In one typical scenario, software engineers are assigned user stories to develop features for Project X. These engineers promptly set up the necessary infrastructure to deploy their workloads. As they progress through the development process and conduct load testing, they realize that using larger compute resources with increased storage capacity enhances workload performance. The cost difference between the two environments, in terms of compute and storage, is merely $0.02 per day. Considering the marginal cost increase against the significant boost in workload processing speed, the engineer decides to opt for the higher compute and storage options.

Let's assume there are one hundred engineers in this organization, each developing and deploying one workload. Furthermore, let's assume that 20% of the engineers do not perceive this slight cost increase as problematic. By performing some calculations:

Total amounts increase per year = (% of engineers with larger compute/storage) * (number of engineers) * (cost increase per workload per day) * 365

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

This results in a grand total of $146, which is approximately 7,300 times $0.02. At first glance, this may not appear significant. However, several factors have not been considered to make a full assessment:

a. Does the increase in compute/storage truly justify the enhanced output speed?
b. Are the consumers of these workloads satisfied with lower frequency outputs?
c. Who has the authority to make decisions regarding increased compute/storage requirements?
d. Has any cost modeling been conducted beforehand to assist engineers in making such decisions?

When engineers solely focus on individual workloads and their cost wastage, the problem seems small enough to even address.

On the other hand, when leadership or finance teams solely focus on the aggregate numbers, they may fail to recognize a crucial aspect: the diminishing value of the product relative to the amount of money being spent. In other words, as time progresses, the organization is receiving a progressively decreasing level of output in relation to the financial resources invested.

The above scenario is one of the many situations where wastage becomes a crucial issue to tackle for optimization efforts. Let's review another scenario.

7.2. Scenario - 2

In this scenario, software engineering teams have made the decision to deploy an application that requires a certain amount of compute resources (X) and storage (Y) to generate a specific output. As multiple teams within the organization develop similar applications to perform comparable functions (a common situation in large enterprises with isolated structures), the total compute and storage needed becomes dependent on the level of duplication.

The total expenditure for these applications can be calculated as: Total $ spent = N (X Compute $'s) + N (Y Storage $'s), where N represents the number of applications with similar functionality.

At this point, organizational leadership steps in to address the issue of silos and initiates a consolidation effort. However, during such consolidation processes, it is often inevitable that some loss in workforce is expected. The newly onboarded engineers are now faced with the challenge of dealing with numerous compute instances and storage setups as they undertake the consolidation effort.

In this context, instances and storage layers that remain unused during the consolidation process and incur relatively lower expenses, such as a few hundred dollars, often go unnoticed.

Now, let's expand the scope of this scenario to encompass the entire enterprise and consider the cumulative impact of the ongoing transformations, as well as the accumulation of unused cloud infrastructure capacity over time. While the overall cost is substantial when viewed as a whole, individual engineers may perceive the individualized resource cost as relatively insignificant or fragmented for the problem of underutilization.

The authors have emphasized only two scenarios, yet the issue of optimization becomes unmistakably apparent. Specifically, it raises the question of how to persuade engineering teams that seemingly insignificant expenses resulting from the unused or underutilized cloud infrastructure can accumulate into a significant sum if left unchecked.

The authors propose a suggestion to tackle this issue by implementing a set of measures. Please note that the following topics are mentioned briefly, but they require a comprehensive study on their own. These include establishing robust measurement mechanisms, defining thresholds using micro budgets, implementing alarm systems, and creating a feedback loop for engineers. Additionally, the authors

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

recommend that engineers who take the initiative to optimize and address the problem should be duly rewarded for their efforts.

The proposed end-to-end flow is outlined below:

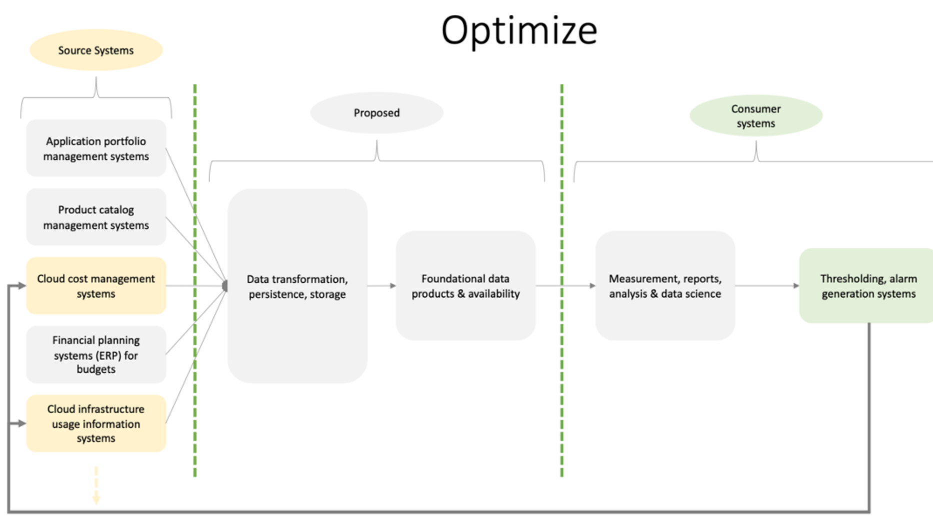

Figure 5 - End State Application View

In a planned subsequent paper, the authors deliberately defer the exploration of several crucial aspects related to establishing an automated feedback loop, implementing gamification techniques involving rewards for engineers, setting precise thresholds and alarms, as well as employing proactive approaches to cloud cost management. However, it is anticipated that the concepts presented herein serve as a catalyst for the generation of innovative ideas and inspire further investigations into the diverse range of possibilities in constructing a comprehensive and resilient framework encompassing the entirety of cloud cost measurement, optimization, and actionable strategies, all while leveraging the potential benefits derived from gamification principles.

8. Conclusion

Network operators have fully embraced the virtualization of network functions, recognizing its benefits in terms of speed, flexibility, and innovation. Simultaneously, the adoption of cloud technologies for deploying software through virtual instances is gaining momentum. These two trends are interconnected but also pose new challenges for network operators. One significant challenge is effectively managing cloud costs during software development and deployment. Organizations now confront complex cloud bills comprising billions of individual line items with small amounts. Finance and leadership teams face the daunting task of comprehending how these various monetary amounts contribute to business capabilities while minimizing waste and redundancy.

This paper addresses measures required to tackle such a situation, emphasizing the importance of a robust measurement and observability platform as a foundation. Additionally, the authors offer insights into potential issues that may arise in such a setup. Finally, the paper concludes by outlining the future of cloud cost management, advocating a proactive approach over a reactive one.

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.

Abbreviations

COTS	Commercial off the shelf
VNF	Virtualization of network functions

Bibliography & References

J. R. Storment, executive director of the FinOps Foundation says, "Cloud removes finance from the buying process and hands the credit card to cloud engineers."

© 2023, SCTE® CableLabs® and NCTA. All rights reserved.