Three Moves for CIOs to Lower Business Costs With Cloud

Published: 24 March 2017 **ID:** G00326262

Analyst(s): Ron Blair

Summary

The proliferation of business unit cloud use is primarily motivated by digital transformation's need for innovation and agility in big data and real-time analytics, rather than cost optimization, which causes overspending. CIOs should use three moves to gain business influence and optimize costs.

Overview

Key Challenges

Organizations that manage cloud costs from an IT perspective only have little visibility as to the degree of the cost impact on the overall business.

The cloud has helped businesses improve costs in customer acquisition and customer service, reach new markets, push out new products, and minimize procurement and supply chain costs. However, without cost optimization, they could be overspending by more than 70%.

Innovation enablement and agility are important reasons cloud use is growing; however, central IT in most organizations has made little investment in learning how to extract this value.

If the promises of innovation and the cloud come to fruition, then organizations' costs will go up; however, they are ill-prepared to explain "why their cloud costs went up."

Lack of skill sets and experience with planning structures, poor automation tool choices, and cultural resistance have caused organizations to be dissatisfied with automation capabilities.

Recommendations

CIOs tasked with managing cost optimization should:

Create a cloud cost center of excellence as a bimodal bridge aimed at driving down IT costs and business-impacting margins, if their cloud spending is expected to exceed $750,000 next year. And, at minimum, engage a managed service provider.

Fund automation engineers in a CCoE with cash savings, and overcome the cultural challenges of change by focusing on discrete business process margins, resiliency and downtime costs.

Hire the right automation skill sets and have the CCoE work in agile approaches familiar to application developers, employing a leader who interfaces with the business and sets priorities. Demand that automation engineers spend at least 50% of their time on automation.

Become bimodal in cloud cost management by improving average IT costs and developing customer-facing digital business offerings with different service levels and rightsized margins.

Introduction

This document was revised on 26 April 2017. For more information, see the Corrections page (http://www.gartner.com/technology/about/policies/current_corrections.jsp) .

Imagine scenarios in which your IT staff's idea of cost optimization is less about server-to-full-time equivalent (FTE) ratios and, instead, is more about how they are architecting and automating to optimize cost, develop rightsized resiliency and deliver differentiated service levels that maximize business margins.

For example, a business unit could use Amazon Web Services (AWS). If it has a web or mobile application, ideally, it would offer three different service levels with respect to response time. This would involve how long it will take for it to upload, view or perform analytics on a file. In fact, the business unit might strongly feel it could have different service levels that would involve different price points: free (within 10 minutes), standard service (within five minutes) and premium service (nearly instantaneous).

In traditional IT, this would be difficult to bring to fruition. The data center, IT staff, storage and other cost components could not be rightsized in a material fashion that would allow costs to be commensurate with the desired service-level price points. Part of this is cultural, but part of it is just a fact. If one already paid for a storage area network (SAN), doing something different for one business group could be an improvement for that group, but detrimental to the organization as a whole.

However, if your team was already prepared for this, then it's well beyond traditional thinking. It's well-versed in cloud architectural patterns and has an opportunistic mindset. It's "battle tested" and able to handle architectural options that exploit cloud economics in ways that best serve the business. Instead of a one-size-fits-all approach, it's clever. It realizes that, on AWS, it can take advantage of the fact that pricing for certain databases is rooted in how often writes are made. So it can use queueing services (SQS) to set how often the writes are made, depending on the customer's subscription level.

Rather than immediate writes for each request, the flow can be automated as to whether the writes will happen immediately or occur every five to 10 minutes. This has a significant impact on the business bottom line. Beyond that, it has deployed Spot Instances (which cost as much as 80% less than AWS on-demand instances).

However, the team is aware that Spot Instances may be terminated by the cloud provider with only a two-minute notice. It is agreed this is fine for free subscribers, but the team has created spot fleets and has automated via auto-scaling and serverless (Lambda) functions to ensure lower costs and improved resiliency for paying silver and gold subscribers.

Speaking of resiliency, the team is acutely aware of the need to have redundancy between two availability zones to mitigate the risk of an outage, should an availability zone go down. It knows that, at a minimum, the application will have 30 instances running at any time. Normally, this would require having 30 instances in availability Zone A and another 30 instances in availability Zone B —

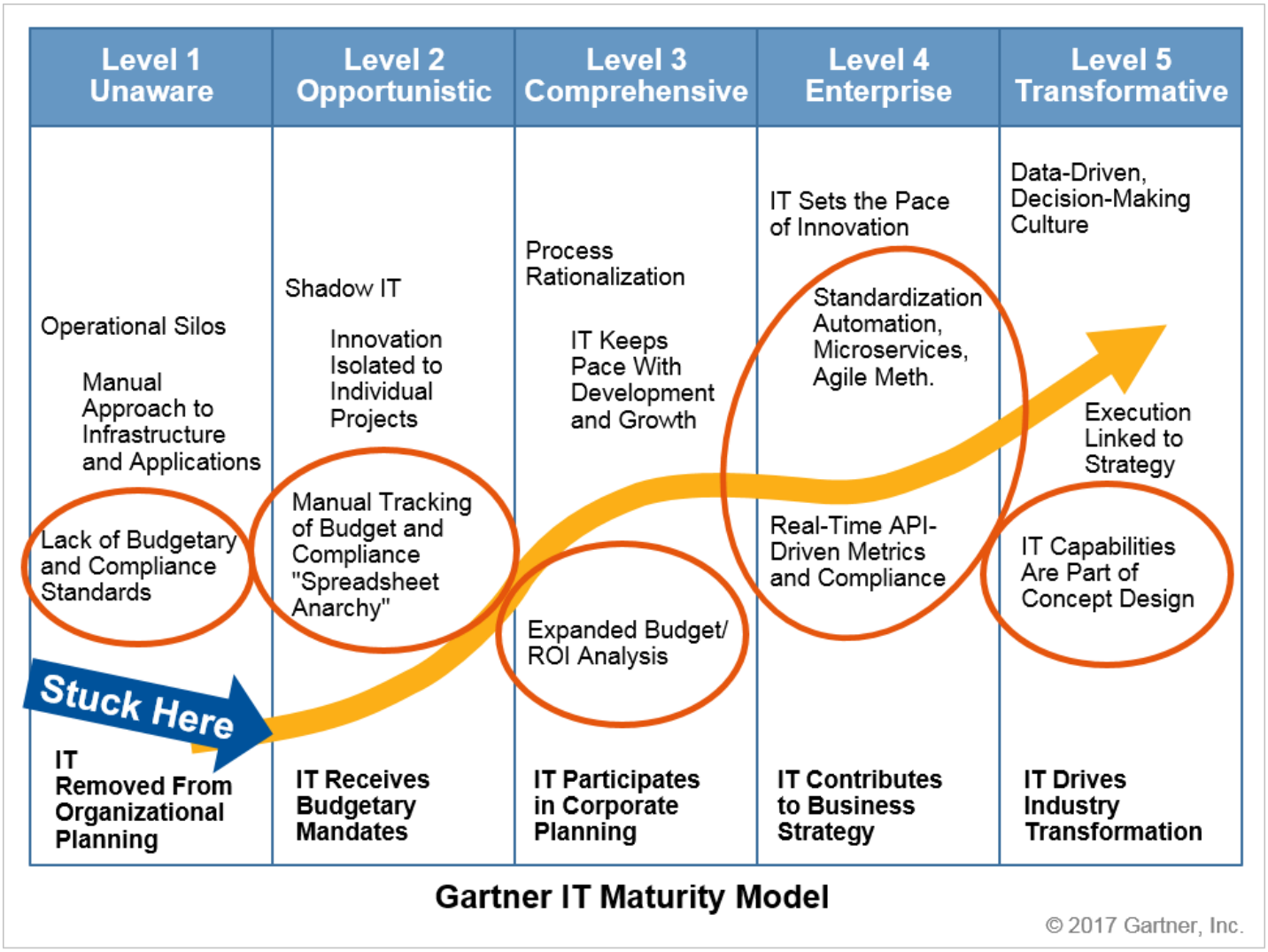
for a total of 60 instances. However, because your team is uber-focused on cost optimization, it would go beyond the ordinary and deploy between three availability zones —15 instances in each, for a total of 45 instances.

This one step alone produces 25% savings with equal resiliency. Meanwhile, your financial operations (FinOps) have captured the average cost per free and paying subscriber and the cost per transaction on the cost optimization dashboard, enabling the team to trumpet successes and to track continuous improvements.

In short, with no additional capital investment, long procurement cycles or agonizingly slow proofs-of-concept (POCs), IT has helped the business create three differentiated service-level offerings, ensured resilience and made a significantly positive impact on the business' bottom line. Moreover, IT has contributed to business strategy and transformation in a big way, which can be presented in a meaningful way to the business, with relatable fiscal context.

Is that a pipe dream? For many, yes. In fact, most organizations are "stuck in the weeds" (see Figure 1). This is particularly true for the infrastructure and operations (I&O) segments of CIO organizations, which have suffered average budget decreases of 7.6% during the past two years.

Figure 1. Many Organizations Are Stuck in the Weeds



Source: Gartner (March 2017)

Organizations list agility, innovation and lower costs as the top three reasons for adopting the cloud (see "Survey Analysis: How Cloud Adoption Trends Differ by Organization Size" for more details). However, many organizations do not put forth the investment required toward preparation, skill sets and organizational structure to abstract value out of these three areas for IT cost benefits — let alone for business impact benefits. The remainder of this research will work through the three main steps needed for CIOs to get to the level of maturity described above.

Analysis

Create a Cost Center of Excellence for the Cloud

At a minimum, a cost center of excellence (CCoE) should initially consist of a cloud architect, two FinOps staff members with ties back to the CFO, and two or three automation engineers with cloud knowledge and DevOps experience related to RESTful API and continuous configuration automation. Security should have a prominent role in selling the business on the vision, and the automation engineers should pay for themselves with the hard cloud cost savings alone. However, they will initially work on the basics, such as consolidating the invoice for savings, tagging, reporting and eliminating waste alongside the FinOps team.

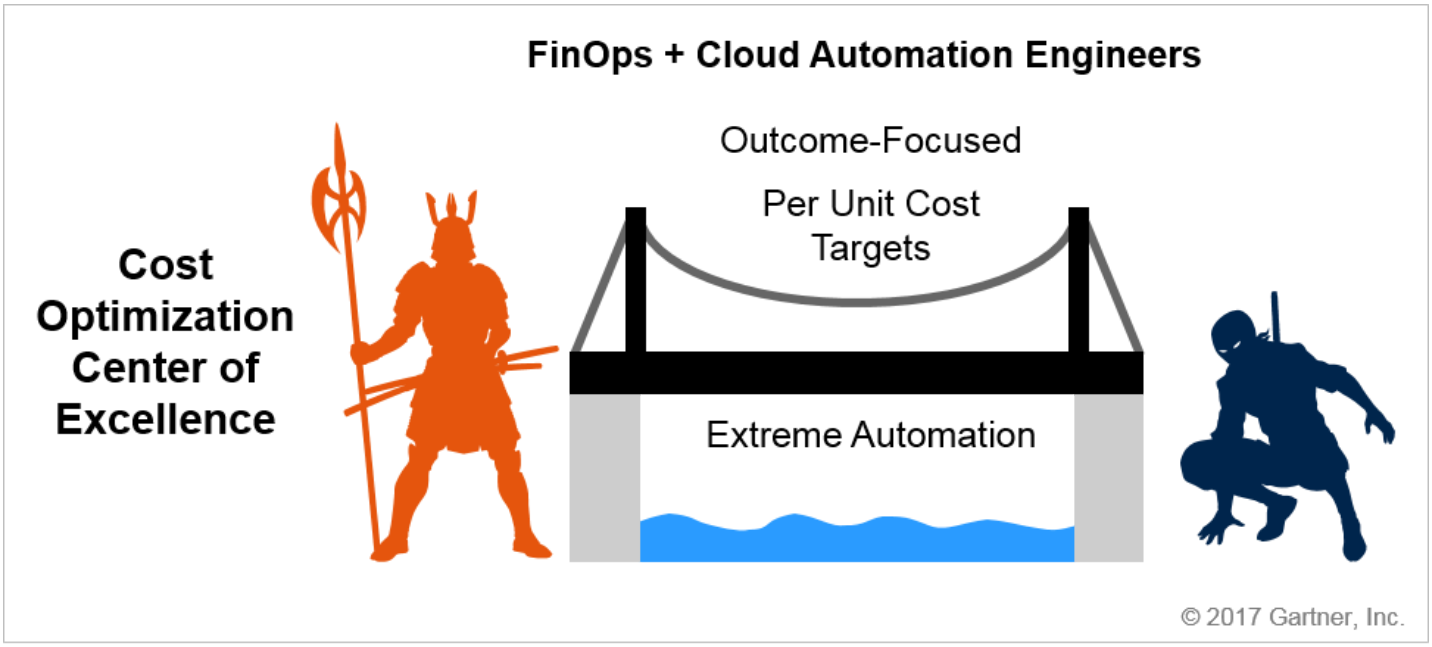
Why is this team needed? Because, like IT as a whole, cost optimization must be bimodal. However, more specifically, there are three reasons:

The magnitude of the cloud options.

Older methods of sourcing and cost optimization are too slow, reactive and constrained.

A CCoE will serve as a bridge for strategic buy-in and influence with the CFO and business units (see Figure 2).

Figure 2. CCoE Is a Bridge



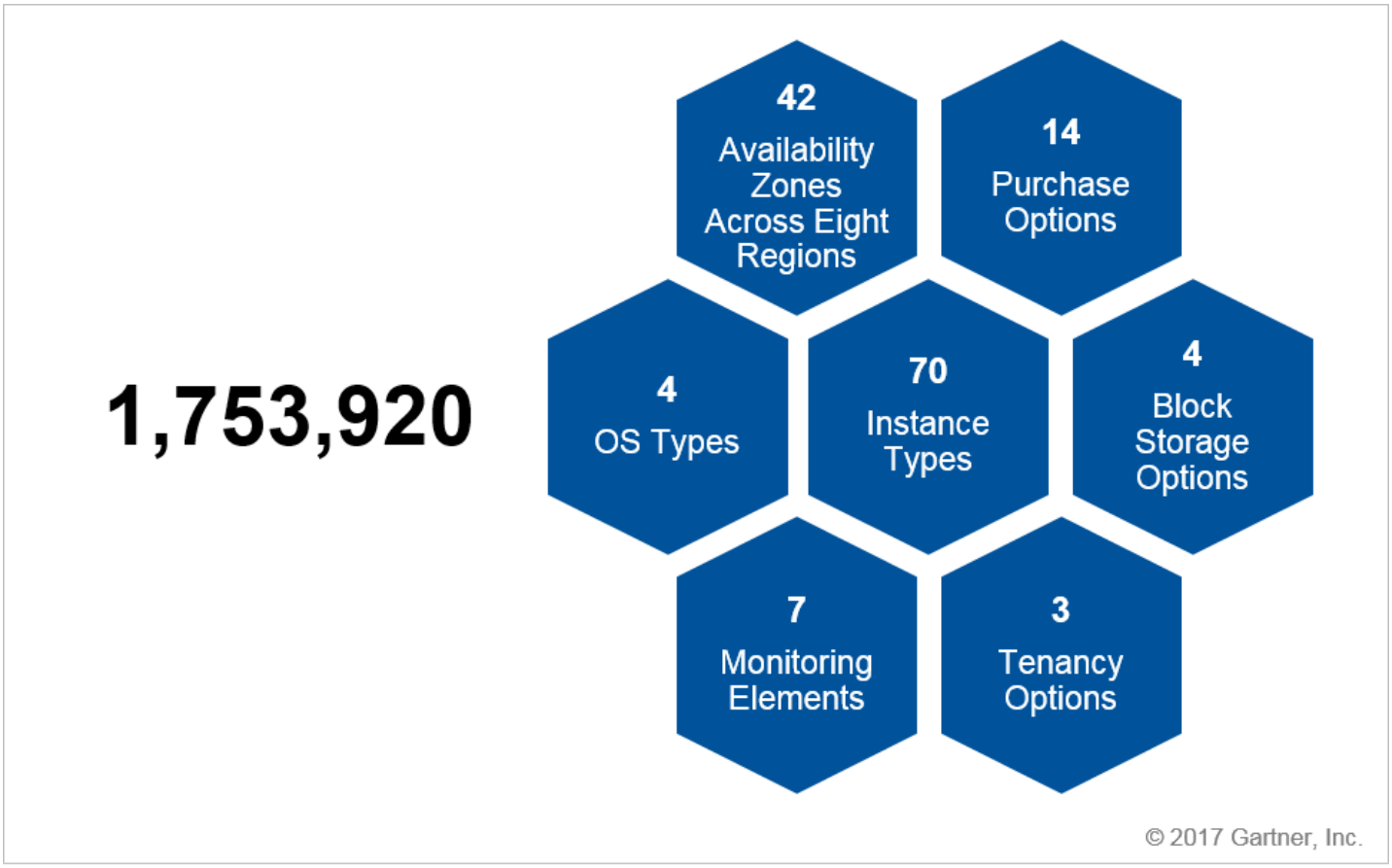
Source: Gartner (March 2017)

We have already touched on organizations that are stuck with regard to cost optimization per the Gartner IT Maturity chart and the fact that current cost optimization approaches will not bring the scenario in the introduction to life for your organization. However, what do we mean by "magnitude of cloud options"? Most would think that, because it's simpler — no data center to manage, no power, no cooling, no patching, cabling, etc. — it should be more straightforward? Unfortunately it isn't.

Gartner estimates that organizations that have done little or no cloud cost optimization are overspending by 70% or more. In fact, it is not uncommon for Gartner to receive inquiries from clients who are spending two to three times more than what was initially budgeted. One reason is an underappreciation of the complexity of cloud architectures and commercials. For example, AWS has eight regions with 43 availability zones, 70 instance types, 32 images types, 14 purchase options, etc. This comes out to more than 1.7 million different combinations from which a system operator or a developer can choose. And that is just for Elastic Cloud Compute (EC2). That doesn't begin to touch serverless computing, containers, storage, when data transfers are charged and when not, content delivery networks, data analytic options, database configurations, or message queuing (see "Ten Moves to Lower Your AWS IaaS Cost").

Organizations are already overwhelmed with optimizing costs for on-premises workloads, and I&O is merely trying to keep up with budget cuts. The result is many organizations overspend and never get to a point at which they can separate "the forest from the trees" and truly enable the business by reducing business costs (see Figure 3).

Figure 3. AWS EC2 Instance



Source: Gartner (March 2017)

As touched on earlier, well before the cloud, IT organizations as a whole have been stuck well short of contributing to business strategy and driving transformation. By using cloud usage as the "compelling event" and by moving forward with the three items below, CIOs can take their organizations to greater levels of impactfulness.

This move is strategically the foundational bridge toward bringing I&O, application teams, the businesses, the CFO, and security together toward common goals that focus on financial performance and automation excellence.

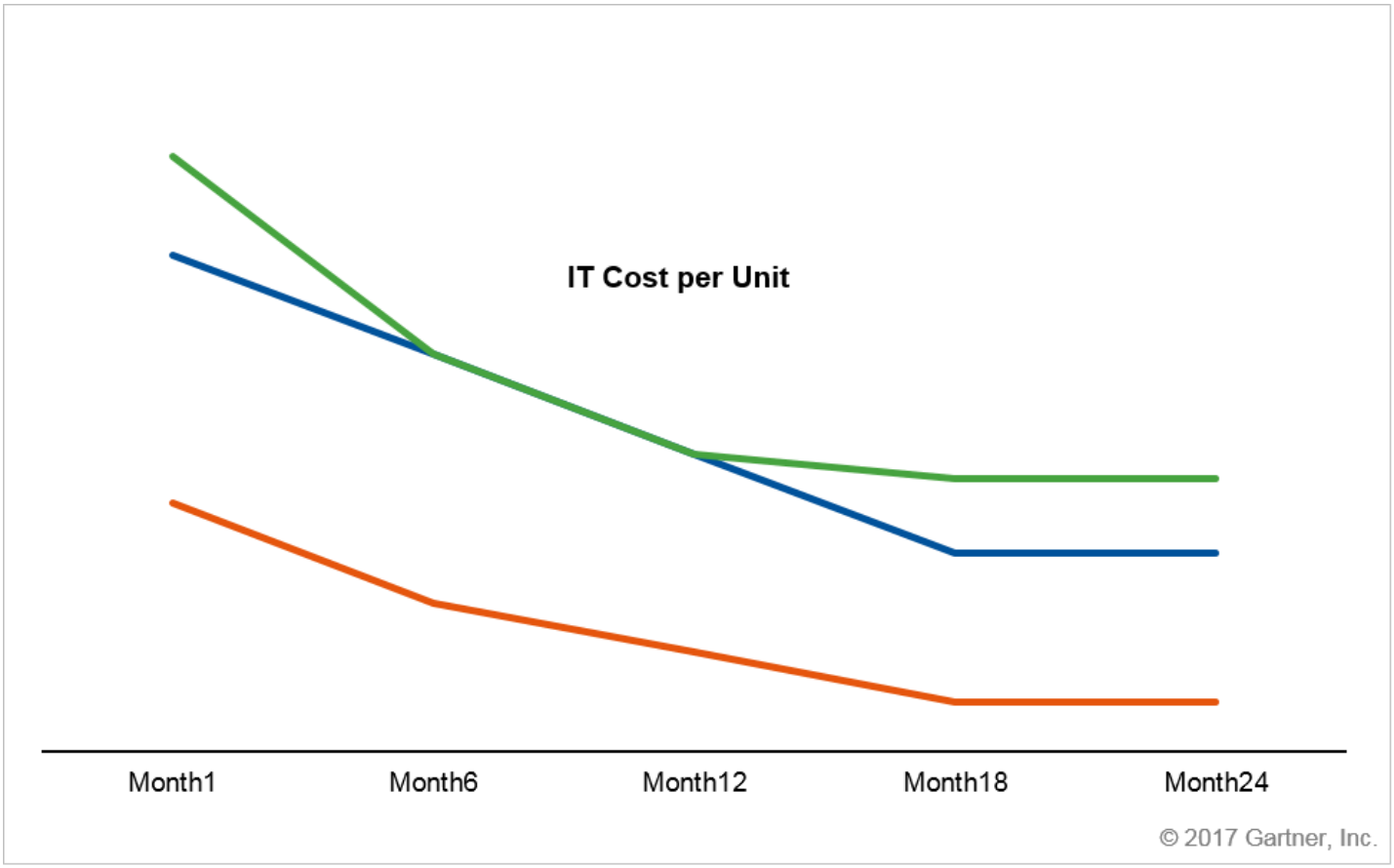
Get Stakeholder Buy-In With Your Initial Vision and Two Examples

Painting the initial storyline (or similar) is needed, so that the stakeholders can look beyond how things have always been done and toward the future. Here are two examples.

Example No. 1: Measure the Impact on Business per Unit Costs

Measuring IT cost efficiency is, of course, important for ensuring that the services provided back to the business are cost-effective and cost-competitive. The FinOps team (again an extension of the CFO) are responsible for tracking overall costs, as well as IT per unit cost efficiency (see Figure 4). With proper cost optimization, you should see steady declines.

Figure 4. IT Cost per Unit

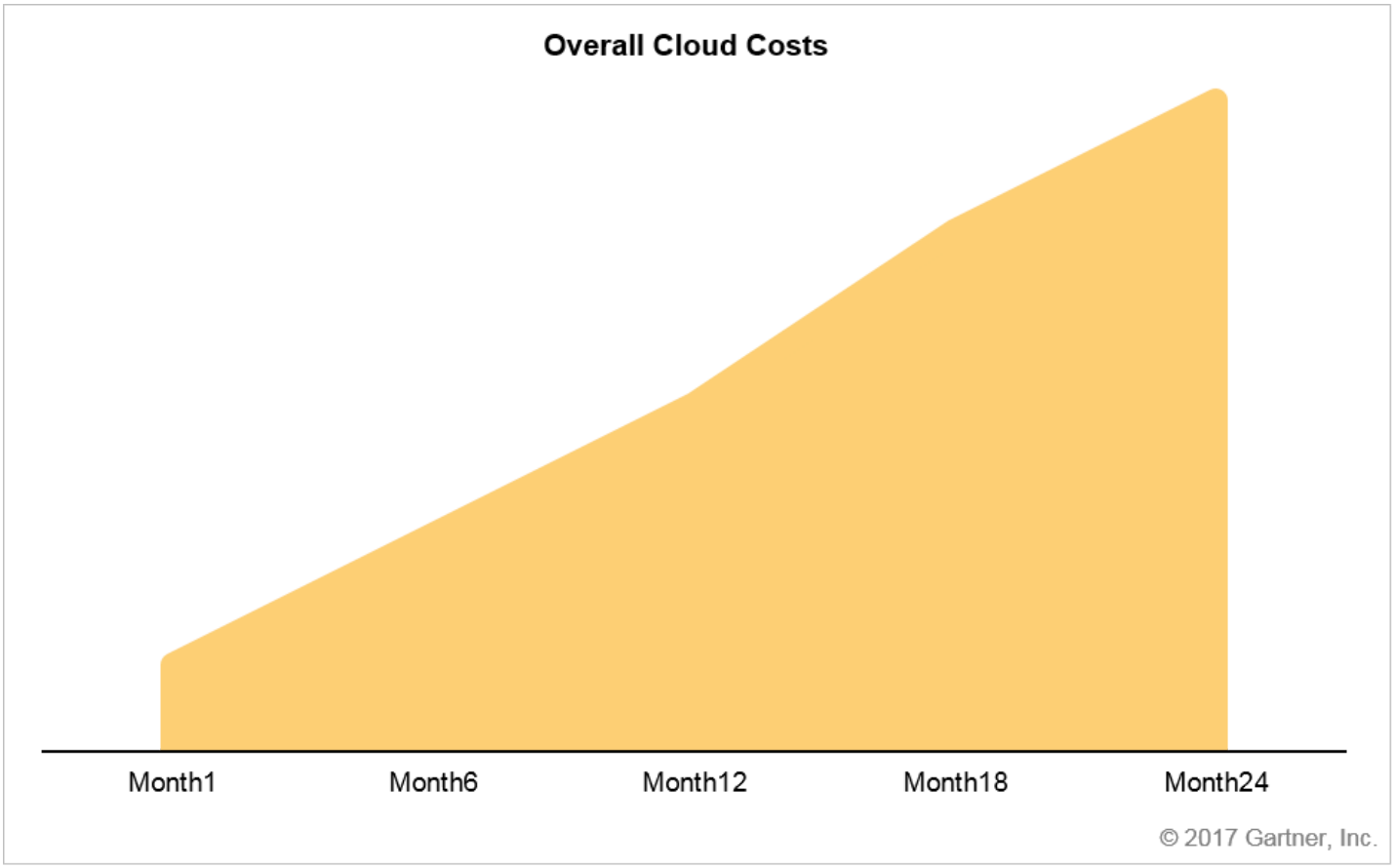


Source: Gartner (March 2017)

However, assuming the business reasons for going to cloud are successful (innovation and business agility), overall cloud costs will increase. Not only will this overshadow optimization efforts, but questions will remain as to whether the increase in volume is having a positive material effect or whether it is just being poorly managed.

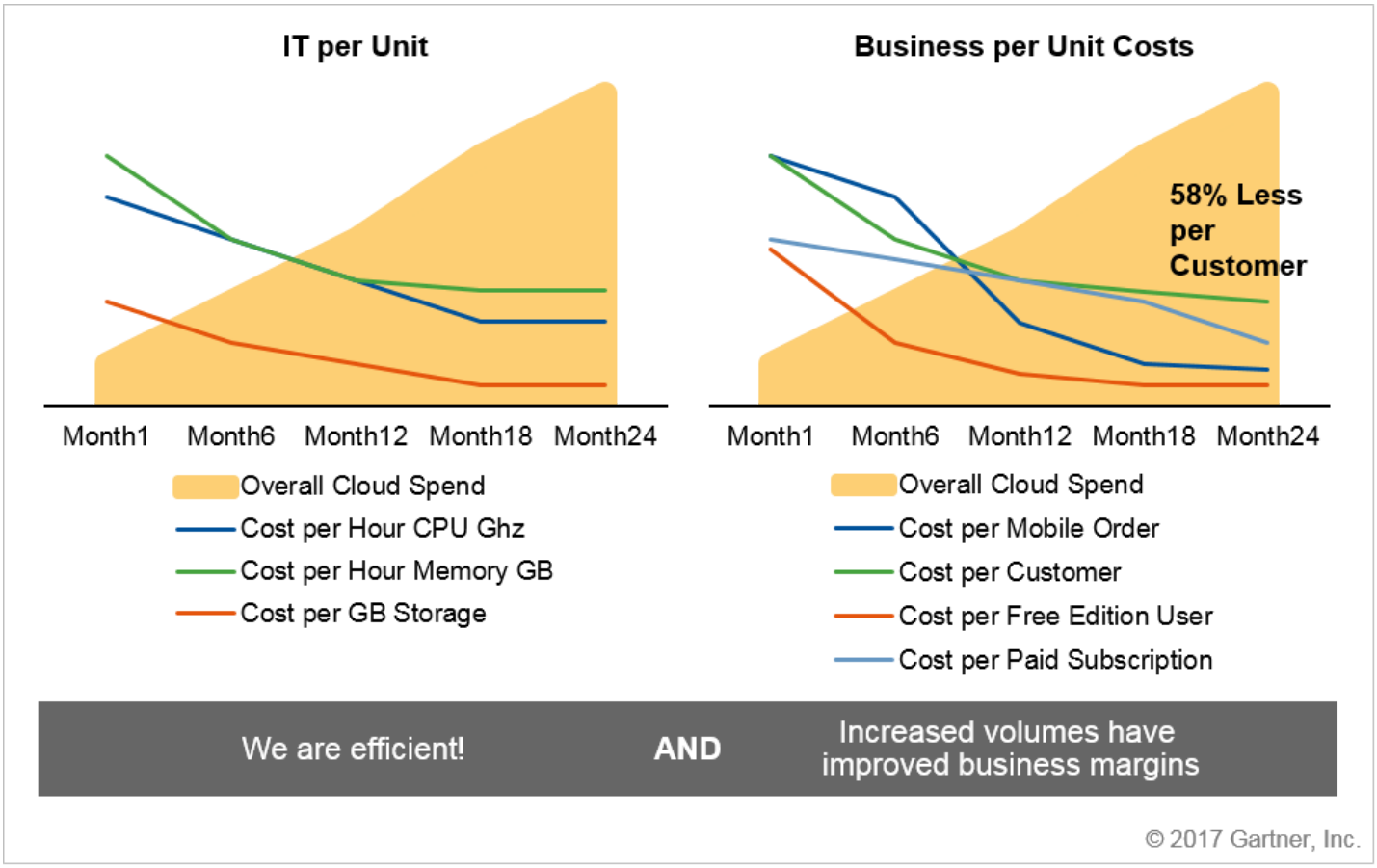
To that end, you need to measure costs on two planes (see Figure 5 and Figure 6).

Figure 5. Overall Cloud Costs



Source: Gartner (March 2017)

Figure 6. Per Unit Costs



Source: Gartner (March 2017)

IT PER UNIT

You should track how you are doing from an IT perspective and be able to answer questions regarding average cost per GB storage and hourly costs associated with compute, and market your success. Although it can be complicated for the cloud, the concept is rather straightforward for organizations with shared services.

BUSINESS PER UNIT COSTS

Ultimately, you will need to account for the uptick in volume that is likely. Measure the positive impact made on business-specific items that affect margins. Success is being efficient from an IT cost standpoint, as well as positively affecting business services and associated margins — such as costs per mobile app purchase, which should be specifically tracked and should show reductions as well. This specific business-related impact is an example of achieving greater levels of IT maturity, as was shown earlier in Figure 1.

Example No. 2: Mitigate Cost of Downtime and Increased Innovation

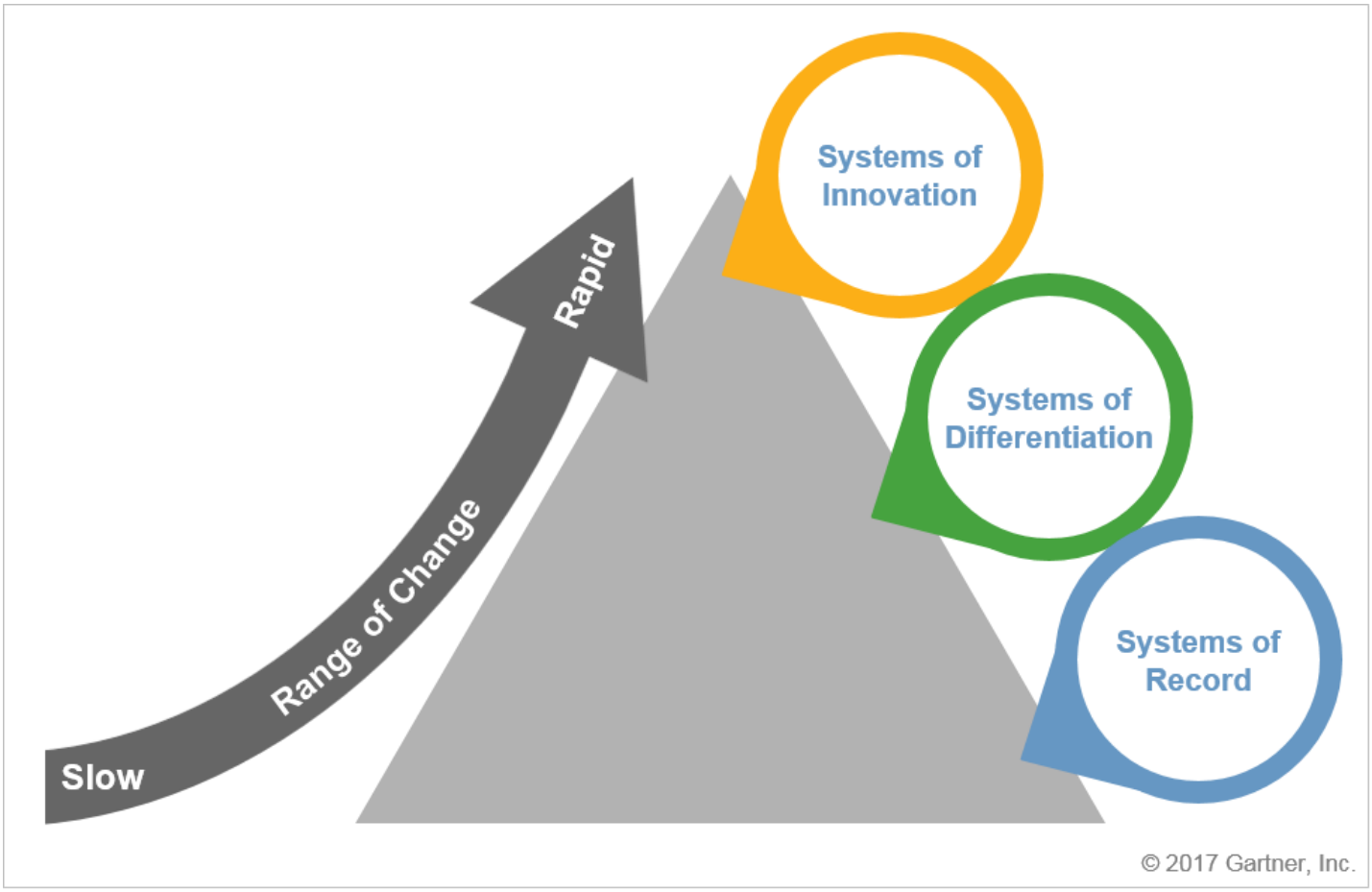
Outages can occur that can affect the bottom line to the business units. The fiscal damage can be the result of lost revenue, rework by both IT and business units, reputation impacts resulting in customer decline, or compliance and regulatory penalties.

Although there's no magic answer on how much impact downtime will have on your business, current industry surveys have shown that the average enterprise estimates an impact of approximately $5,600 for every minute of unplanned downtime in its primary computing environment

(see "Ensure Cost Balances Out With Risk in High-Availability Data Centers").

At the highest level, this is no different in the cloud than it is on-premises. However, when it comes to expectations of "always on" with cloud, the ways resiliency is created with cloud and what it says about organizations' abilities for innovative churn are significant (see Figure 7).

Figure 7. Range of Change



Source: Gartner (March 2017)

According to recent Puppet Labs surveys, organizations that have instituted massive automation with regard to deployment and operations have a significant positive impact on the cost of downtime, while significantly enabling more-frequent updates. This is due to having smaller change failure rates and significantly faster mean time to recovery (MTTR).

Table 1. The Effect of Automation

Deployment frequency per year*	
High Performers	1,460
Medium Performers	32

Low Performers	7
Change failure rate*	
High Performers	7.50%
Medium Performers	38%
Low Performers	23.50%
MTTR (hours)*	
High Performers	1
Medium Performers	24
Low Performers	24
Outage cost/hour**	
High Performers	$336,000
Medium Performers	$336,000
Low Performers	$336,000
Total cost of downtime per year	
High Performers	36,792,000
Medium Performers	98,058,240
Low Performers	13,265,280
Downtime cost per deployment	
High Performers	$25,200
Medium Performers	$3,064,320

Low Performers	$1,895,040
*** Source: Deployment, failure rate and MTTR statistics for critical applications from 2016 State of DevOps Report by Puppet Labs. (On average, low performers take longer to recover, at a statistically significant level; however, they had the same MTTR as the medium performers.) ** $5,600 per minute x 60 minutes per hour**	

Source: Adopted From Puppet Labs/Gartner (March 2017)

Those with a high degree of automation are able to take advantage of the agility and innovation cloud provides and turn it into business value by deploying more features at lower costs of downtime per average deployment. This is accomplished on two levels:

For new deployments, it is enabled by continuous integration and continuous development.

If your environments do not need or require such high frequency of change, even stagnant workloads can benefit by approaches often dubbed as DevOps or site-reliable engineering (SRE) — a combination of monitoring, autoscaling and automation that minimizes the frequency of downtime and how quickly it recovers. This often goes unnoticed by end customers.

If there is interlock regarding the costs of downtime by business process, automation opportunities can be prioritized accordingly.

Deliver on the Vision by Investing in Automation Skill Sets, Setting Targets and Being Firm on the Ground Rules

Vision and examples are good. However, as previously mentioned, automation initiatives often underachieve if they are approached with antiquated methods, use traditional automation techniques and lack cultural buy-in.

People and Tools — There are many different forms of automation tools. None are "automagically" going to do it all. However, ensure that the automation engineers have a DevOps background and experience with REST API and continuous configuration automation tools.

Productivity Mandates and Stakeholder Alignment — Ensure that the resources brought on spend at least 50% of their time on what the automation engineering team is actually automating. Anything less means they are firefighting and mired in historical I&O practices. This may be a sign to hire more people. That stated, establish an arrangement that, once thresholds are surpassed, members from the application development team come over to support. This serves as backup support when there is work overflow, gets everyone on the same page regarding goals (versus throwing it over the fence for someone else to deal with) and rightsizes application feature throughput with resiliency by signaling when change is too fast.

Approach for Prioritization — There should be a team leader with an agile process background. This leader should understand the business well. Engage with the other groups and determine the prioritization of automation activities.

Set Minor Goals and Expand — Start with a couple of business groups with specific applications that use the cloud. Measure and track current costs related to IT, but also the costs per transaction. Celebrate and market quick wins, and expand the program.

Get Help — For organizations that are uncertain of future cloud onboarding and spending, or are unsure whether such investments are justifiable, use a managed service provider (MSP), at least for the first couple years, to establish the foundational elements required (see "Magic Quadrant for Public Cloud Infrastructure Managed Service Providers, Worldwide").

The Bottom Line

The cloud presents the perfect opportunity to set the organization in the right direction with respect to enabling the business to differentiate in the marketplace and do so cost-effectively. However, CIOs need to invest in a CCoE for the cloud that has an automation approach toward cost optimization. Otherwise, the organization as a whole will spend too much, and teams will never make headway toward the truly impactful examples.