A new approach to help you achieve least privilege at cloud scale.

Cloud Infrastructure Entitlements Management (CIEM)

White paper

CLOUDKNOX

One bad human identity. One poorly configured firewall. One machine identity with excessive high-risk permissions. Three commands.[1]

That's all it took to expose the personal data of more than 100,000 customers of a major US-based bank in a widely reported 2019 security breach. The breach, which led to fines of over $80 million to the institution, resulted from a hacker that took advantage of an over-permissioned AWS role, which in this case included the ability to discover and exfiltrate personal identifying information.

In the now famous response to the breach, AWS's CISO Stephen Schmitt stated that "even if a customer misconfigures a resource, if the customer properly implements "least privilege policy," there is relatively little an actor has access to once they are authenticated — significantly diminishing the customer's risk.[2]"

So, how can enterprises reduce their risk in the cloud? They need to understand the attack surface has changed and operate under the assumption that the number one risk to their hybrid and multi-cloud infrastructure is a trusted identity with excessive high-risk permissions. The only way to manage that risk is to implement the principle of least privilege across their cloud environment. If not, they run the risk of compromising every security system, policy, and procedure they've worked to put in place.

Risk Starts in the Enterprise

Recent breaches shed light on the fact that many enterprises don't realize that protecting their applications and data in the cloud is a shared responsibility between their organization and its cloud services providers (CSPs). In its simplest terms, the cloud shared responsibility model depicts the division of responsibilities between the cloud providers and their customers – placing the responsibility of security and availability of the cloud infrastructure on the CSP's and the responsibility for security in the cloud solely on the customer.

According to Gartner, it is estimated that at least 95% of cloud security failures through 2022 will be the fault of the enterprise.

There's a fundamental challenge with this model. For example, AWS is responsible for the security of its services and the infrastructure that runs the AWS cloud. However, the enterprise customer might be surprised to learn they are solely responsible for the security of the resource(s) they create in AWS. When an enterprise deploys an ec2 instance, they must manage the guest operating system, any applications they install on it, and the configuration of provided firewalls on these instances. They are also responsible for overseeing data, classifying assets, and implementing the proper permissions for identity and access management.

Moreover, each CSP employs different hardware and software security policies, methods, and mechanisms, creating a massive challenge for enterprises attempting to maintain standard policies and configurations across multiple cloud deployments. And CSPs generally only meet basic security standards for their platforms because they want to standardize how they monitor and mitigate threats across their entire customer base. That's why it's more important than ever before for enterprises to clearly understand the division of responsibilities between them and their cloud service providers.

Why Is It So Hard to Achieve the Principle of Least Privilege in the Cloud?

The principle of least privilege has been a fundamental tenet of the infosec world for as long as infosec has been around. However, what has changed is the complexity of implementing the principle and the severe consequences enterprises face by NOT enforcing it in the cloud.

In the past, organizations had fewer identities to manage, typically human employees – server admins, for example - or contractors. Fast-forward a few years, and it's not only server admins that have access to critical cloud infrastructure, but also developers, third-party contractors, and a host of non-human identities like service accounts, bots, API keys, and a variety of compute types like ec2 instances and serverless functions. These non-human identities have the same high-risk permissions and access to sensitive resources humans do.

In yesterday's legacy environments, it was easy to enforce both the separation of duties and the principle of least privilege. A server admin responsible for racking and stacking servers would rarely or never have the authority to perform actions on a network device or vice versa. Even if the server admin misused permissions, either accidentally or with malice, the damage would have been contained to one system.

But today, the number of identities accessing cloud infrastructure has increased by 50x, fueled by the exponential growth in non-human identities required for automation. These identities can now access more than 40,000 unique permissions across the four key cloud platforms. Nearly 50% of these permissions are classified as high-risk. If these permissions are used improperly – either accidentally or maliciously – the results can be catastrophic to an enterprise's infrastructure.

As a result, human and non-human identities with excessive high-risk permissions are the new attack surface because they wield enormous power to disrupt – often without their organizations' awareness.

The "Cloud Permissions Gap" Emerges as the New Attack Surface

A Cloud Permissions Gap occurs when an enterprise has a dangerous delta between permissions granted and permissions used. While an identity should have only the permissions needed to do its job, in a review of responses from more than 100 global enterprises, it was discovered in most organizations, over 95% of privileged identities were grossly over-permissioned, a state that could leave an organization's cloud infrastructure significantly exposed.

Further analysis revealed identities used less than 10% of the cloud permissions granted to perform their daily tasks, leaving more than 90% of unused permissions unnecessarily wide open to accidental misuse or malicious exploitation.

The Cloud Permissions Gap occurs not through malfeasance but because organizations simply don't have the protocols and capabilities in place to correctly assign, manage, and monitor the exponential growth of permissions across their growing cloud footprints.

Existing Solutions and Strategies Have Fallen Short

How are enterprises working to manage the permission levels placing their critical cloud resources at risk? The short answer: not well enough. Most enterprises today turn to traditional identity and access management (IAM) tools. However, it quickly becomes apparent these solutions are neither granular nor dynamic enough to keep up with the highly automated and digitized environments that define modern infrastructure.

Despite these shortcomings, some enterprises try to manage permissions with the native cloud IAM tools that come with each cloud platform, most of which offer only basic functionality. Moreover, these built-in mechanisms come with their dedicated toolsets, management screens, and workflows that simply won't work for enterprises running multi-cloud and hybrid cloud deployments.

Many enterprises try to control and manage privileged access with traditional strategies, such as RBAC, or role-based access control. RBAC is an older, static method involving the creation of standard roles with pre-defined and broad sets of permissions based on job descriptions and functions within an organization and assigning identities to these roles.

Of course, in today's dynamic environments — even with the most disciplined use of RBAC — organizations can't keep up with managing all the new permissions made available for each cloud service in use. As a result, most roles are seldom updated, if at all. And where they are, the temptation is always to add more permissions to existing roles rather than redesign the roles entirely.

Moreover, once an identity is assigned a role, it is rarely reviewed again. More often than not, the identity is never removed from a role even if it no longer performs the job function. Take the example of a contractor who left a project or a DevOps engineer who moved to another team yet still retains access to his/her original role.

While legacy RBAC helped organizations in previous eras assign proper access to systems, the complexity level in today's dynamic cloud environments has made it ineffective to control the number of roles and tradeoffs required to secure the growing number of permissions across an organization.

The limitations of existing solutions create a new market need: one for a cloud-native, scalable, and extensible way to automate the continuous management of permissions in the cloud. With this need in mind, industry analyst firm Gartner recently created a new research category, Cloud Infrastructure Entitlements Management (CIEM). Key to CIEM is its description of the next generation of solutions for managing access to permissions and enforcing least privilege in the cloud.

We examine CIEM, drawing from the June 2020 Gartner report, Managing Privileged Access in Cloud Infrastructure, below.[3]

A Paradigm Shift: Cloud Infrastructure Entitlements Management (CIEM)

Gartner analyst Paul Mezzera has made a strong case for the need for a new approach to identity, access, and permissions management in the cloud. In the research note mentioned above, Mezzera describes the core requirements of secure cloud infrastructure. This need is pressing; the report states because Gartner advises cloud users to, "Consider the following Gartner Strategic Planning Assumptions:

- Through 2023, at least 99% of cloud security failures will be the customer's fault, up from 95% in 2017
- By 2023, 75% of security failures will result from inadequate management of identities, access, and privileges, up from 50% in 2020"

Mezzera notes that "… to determine a user's effective access privileges, one needs to understand all the different policies attached to the user."[4]

The pillars of CIEM are designed to help users evaluate and implement the best solutions for their cloud identity and permissions journeys. They include, according to Gartner's Mezzera, the following attributes, taken from his report:

CIEM Requirement	Description
Account and Entitlements Discovery	"… an inventory of identities and entitlements across an enterprise's cloud infrastructure." Characteristics, according to Gartner, include continuous, event-based discovery; identification and tracking of all identity types; analyzing all access policies, and discovery of any federated and native cloud identities, including those from CSP accounts, identity providers, and traditional directories, e.g. Active Directory.
Cross-cloud Entitlements Correlation	"Organizations need a method by which accounts and entitlements across clouds can be correlated and normalized into a unified access model."
Entitlements Visualization	"Given the large number of entitlements that organizations need to manage, traditional table-driven visualization methods for viewing and analyzing this information is not feasible. The following characteristics are essential for effectively visualizing cloud infrastructure entitlement data within and across cloud platforms: – Graph identity and entitlement view – Natural language query capabilities – Metrics dashboard"
Entitlements Optimization	"Usage data generated by privileged operations across cloud infrastructure combined with entitlement data is essential in determining least-privileged entitlement assignments."
Entitlements Protection	"An important control for ensuring the overall integrity of the cloud infrastructure is the ability to detect changes within all managed cloud infrastructure environments and to remediate changes made outside of policy."
Entitlements Detection	"The analysis process should detect changes made outside of sanctioned processes or changes that are deemed anomalous due to external factors, are atypical, or considered high-risk."
Entitlements Remediation	"Changes are often required as a result of entitlement optimization or the change analysis process. In either case, an organization may prefer that security tools not make changes directly, but rather trigger a change event containing the updated policy or entitlement assignment. … The ability to detect cloud infrastructure threats and respond by generating events and performing mitigation operations is a required security function."

Taken as a group, the pillars of CIEM are daunting in scope. Nevertheless, to continuously protect critical cloud resources from accidental misuse or malicious exploitation of permissions and achieve a true state of least privilege across clouds, enterprises must move forward on all axes.

A Lifecycle Approach to CIEM

Tackling CIEM through the lens of a lifecycle framework enables enterprises to continuously discover, manage and monitor the activity of every unique human and non-human identity operating in the cloud, alerting security and infrastructure teams to areas of unexpected or excessive risk.

The lifecycle approach also acknowledges the reality of today's operations:

- Organizations will continue to move workloads to the cloud
- CSPs will continue to add new capabilities and services that will breed tens of thousands of permissions
- The number of identities, specifically non-human, will grow exponentially

Critical aspects of a CIEM lifecycle include the ability to:

- Discover risk by uncovering who (identities) is doing what (actions), where (resources), and when across your cloud infrastructure;
- Manage risk by ensuring identities have the least number of permissions needed to perform daily tasks – and no more; and,
- Monitor risk by continuously tracking and measuring changes in identity activity (behavior) and prioritizing alerts based on pre-defined risk criteria.

Discover Risk

You can't fix what you can't see, which is why granular visibility is the first step in the discovery phase of the lifecycle. It starts by uncovering all unique human and non-human identities that can touch an enterprise's cloud infrastructure, what operations (or actions) they are authorized to execute, what actions they have historically performed, and which cloud resources they have accessed.

In hybrid and multi-cloud environments, this level of visibility requires a CIEM solution that can abstract, collect, normalize, and present both real-time and historical identity activity in a single, unified, consumable format. Only with this depth of visibility and insight can organizations understand and mitigate the risk related to the threat that over-permissioned identities pose to the enterprise.

Establishing a Baseline: The right solution will determine this risk by calculating the delta between permissions granted and permissions used over a specific period. From an identity perspective, security teams need this data to build "activity profiles" for each unique human and non-human identity in their cloud environment. These profiles can then be used as a baseline to measure risk and the organization's ability to enforce and maintain a state of least privilege over time. Activity profiles can also be used to detect anomalous or suspicious behavior, such as an identity that suddenly performs a high-risk action for the first time on a critical or sensitive resource they have never accessed before.

Manage Risk

A CIEM solution should combine the visibility of real-time and historical activity data with a simple, automated remediation mechanism and offer enterprises multiple right-sizing tactics. For example, organizations should have the option to either create (or design) custom least-privilege roles based on the historical activity of one or more identities or remove unused or risky permissions directly from a high-risk identity profile.

As CIEM solutions evolve, the ability to "auto-remediate" will become critical, especially as the complexity of managing multiple cloud operating models grows. Essentially, this "auto-pilot" type of functionality is about ensuring continuous "hygiene" and enforcement of least privilege policies across an enterprise's environment without ongoing involvement from the security and cloud infrastructure teams. For example, with an auto-remediation feature, a periodic search for inactive identities can be generated to automatically remove all permissions.

Least Privilege, Just in Time: Gartner also recently advised security leaders to implement "a process for quick and easy requesting and granting of additional privileges with minimal disruption to an individual's workflow." This capability has also been referred to as "privilege-on-demand" (PoD), "just-in-time" (JIT) privileges, or "just-in-time" (JIT) access. CIEM takes the least privilege concept one step further by establishing that identities should not have standing permissions unless they need them for a specific task. The idea is that instead of granting always-on "standing permissions," organizations can use this feature to limit access to permission(s) and resource(s) for a pre-defined time, at which point permissions are rescinded.

This approach mitigates the risk of permission abuse by significantly reducing the amount of time a cyber attacker or malicious insider has to gain access to privileged credentials before moving laterally through a system and gaining unauthorized access to sensitive data.

Monitor Risk

To maintain control and security across clouds, enterprises need to know what is going on at all times. In the modern cloud environment, tens of thousands of identities may be active at any one time, making the task of monitoring them and looking for things that are not right an absolute nightmare. This is why it is critical that a CIEM solution provides robust monitoring and alerting capabilities that empower enterprises to continuously track the activity patterns of all unique human and non-human identities across multiple cloud deployments.

Ideally, enterprises should have the ability to monitor their cloud environments from a multi-dimensional perspective. For example, monitoring activity through the "identity" lens enables the security and cloud infrastructure teams to track changes based on the identity's activity profile. They can quickly ascertain which permissions an identity used, which permissions have not been used, and which resources they have accessed over time.

In the breach example cited earlier, anomalous activity on a cloud resource (e.g. S3 bucket) by an over-permissioned non-human identity went undetected, causing a massive loss of customer data and triggering one of the largest fines to date against an institution.

The ability to continually monitor activity data is critical because it provides the context necessary to detect anomalous behavior, such as an identity that suddenly uses a high-risk permission (e.g., aws s3 sync s3://sensitive_data_bucket) or accesses a sensitive resource (e.g., s3 bucket) for the first time. Monitoring activity from a resource perspective allows the team to track which identities are accessing a sensitive resource and what types of actions they have performed on it.

Most importantly, when something anomalous does happen, the CIEM solution should include the option to invoke an automated remediation response or notify the right team, either through email or third party SIEM or SOAR tools, to take immediate action. Because security teams are already overwhelmed by an avalanche of alerts, fixing security holes requires CIEM solutions to provide context that enables prioritization. To strengthen remediation capabilities, it is simply not enough for a CIEM solution to alert teams to potential areas of risks or threats; the CIEM must deliver an easy, automated way to prioritize those alerts and assess the threat in context.

Conclusion

Cloud Security is only as good as an enterprise's ability to continuously control the level of access privileged human and non-human identities have to their cloud infrastructure. Since the actions these identities can perform are dictated by the number and type of permissions granted them, preventing identities from accumulating unnecessary permissions and quickly responding when they are misused has become a critical capability and top priority for organizations.

The right CIEM solution can offer enterprises a practical, scalable, cloud-native alternative to existing IAM products and manual methods that don't work in the cloud by empowering organizations to continuously enforce the principle of least privilege principle at cloud scale.

About CloudKnox

CloudKnox Security is the leader in the emerging CIEM market. It delivers the only multi-cloud permissions management platform built from the ground up to protect enterprises from the number one risk to cloud infrastructure – human and non-human identities with excessive permissions. The scalable platform is powered by CloudKnox's patented Activity-based Authorization technology enabling automated and continuous implementation and enforcement of least privilege policies at cloud scale.

To learn more about how the CloudKnox Cloud Permissions Management Platform can help your organization reduce its risk and achieve a least-privilege state at cloud scale, please visit cloudknox.io.

May 17, 2021

[1] DAO Research white paper, Securing Data and Applications in the Cloud, July 2020, page 20
[2] https://www.wyden.senate.gov/imo/media/doc/081319%20Amazon%20Letter%20to%20Sen%20Wyden%20RE%20Consumer%20Data.pdf
[3] Mezzera, Paul. Managing Privileged Access in Cloud Infrastructure, 9 June 2020, Gartner, Inc.
[4] Ibid, page 3