Microsoft Security

A new approach to achieve least privilege at cloud scale

Cloud Infrastructure Entitlement Management (CIEM)

One Compromised User Identity. One Poorly Configured Firewall. One Workload Identity with Excessive High-Risk Permissions. Three Commands.[1]

That's all it took to expose the personal data of millions of 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 organizations reduce their permission risks in the cloud? The attack surface has changed and the biggest risk to their multi-cloud infrastructure is trusted identities 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.

[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

The Shared Responsibility Model

Recent breaches shed light on the fact that many organizations 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.

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 customer might be surprised to learn they are solely responsible for the security of the resource(s) they create in AWS. When an organization 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 organizations 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 organizations 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 is a fundamental tenet of Zero Trust security. Fast-forward a few years, and it's not only server admins that have access to critical cloud infrastructure, but also developers, external contractors, and a host of workload identities like virtual machines, web apps, scripts, and containers. These workload identities have the same high-risk permissions and access to sensitive resources as users.

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.

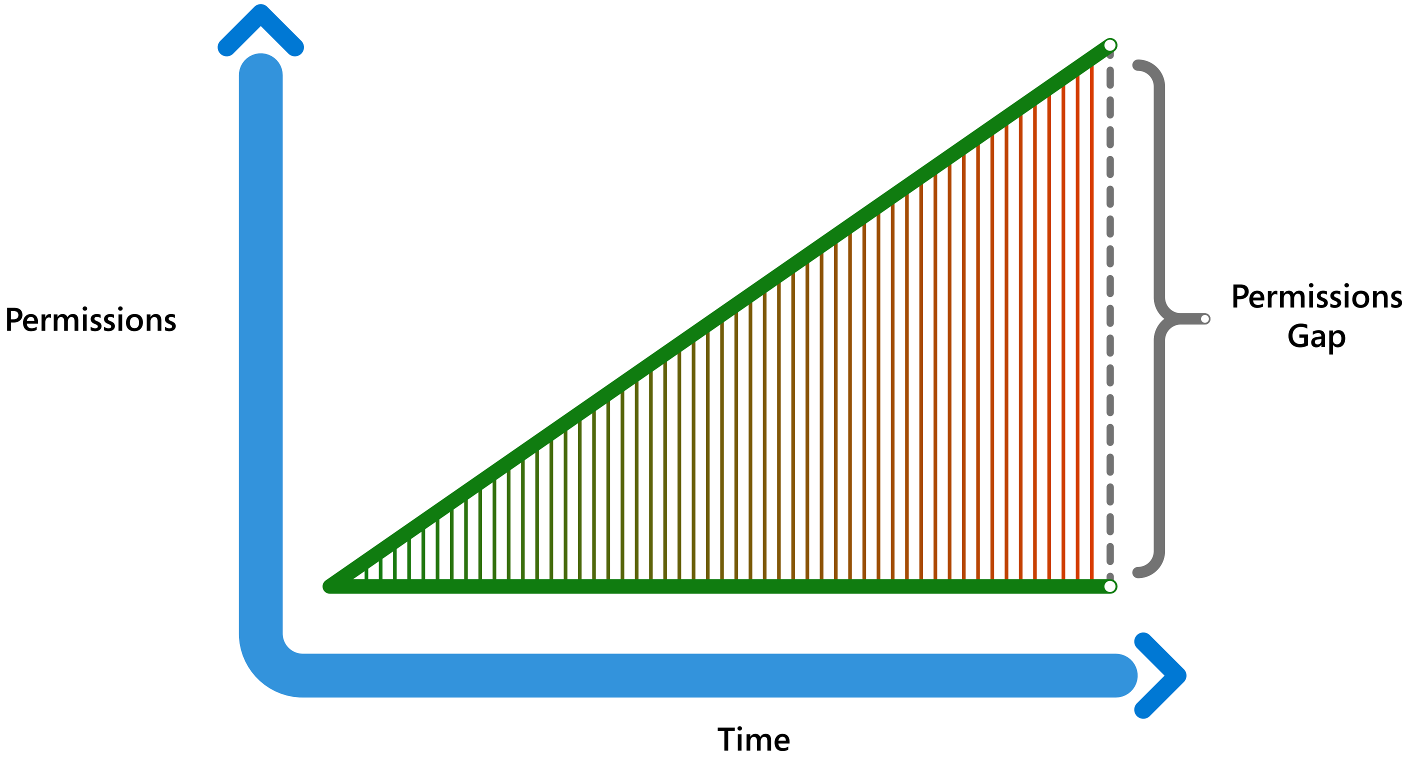
According to Gartner, at least 95% of cloud security failures through 2022 will be the fault of the enterprise

But today, the number of identities accessing cloud infrastructure has increased by 7x, fueled by the exponential growth in workload identities required for automation. These identities can now access more than 40,000 unique permissions across the main cloud platforms. More than 50% of these permissions are classified as high-risk. If these permissions are used improperly – either accidentally or maliciously – the results can be catastrophic for security and business. User and workload identities with unused and excessive high-risk permissions have expanded the attack surface as they wield enormous power to disrupt business – often without their organization's awareness.

Excessive Permissions are Expanding the Attack Surface

While an identity should have only the permissions needed to do its job, in a review of responses from more than 100 global organizations, it was discovered in most organizations, over 90% of privileged identities were grossly over-permissioned, a state that could leave an organization's cloud infrastructure significantly exposed. This dangerous delta between permissions granted and permissions used is called the **Permissions Gap**. The bigger the gap, the bigger the attack surface.

Further analysis revealed identities used less than 5% of the cloud permissions granted to perform their daily tasks, leaving more than 95% of unused permissions unnecessarily wide open to accidental misuse or malicious exploitation.[3]



The 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.

[3] 2021 State of Cloud Permissions Risks Report

A Paradigm Shift: Cloud Infrastructure Entitlement Management (CIEM)

How are organizations working to manage the permission levels placing their critical cloud resources at risk? The short answer: not well enough. In today's dynamic environments, managing thousands of new permissions across multiple cloud presents quite a challenge for organizations to do on their own.

When implementing manual processes or using legacy tools, permissions are granted based on job function and most roles are seldom updated, if at all. When 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.

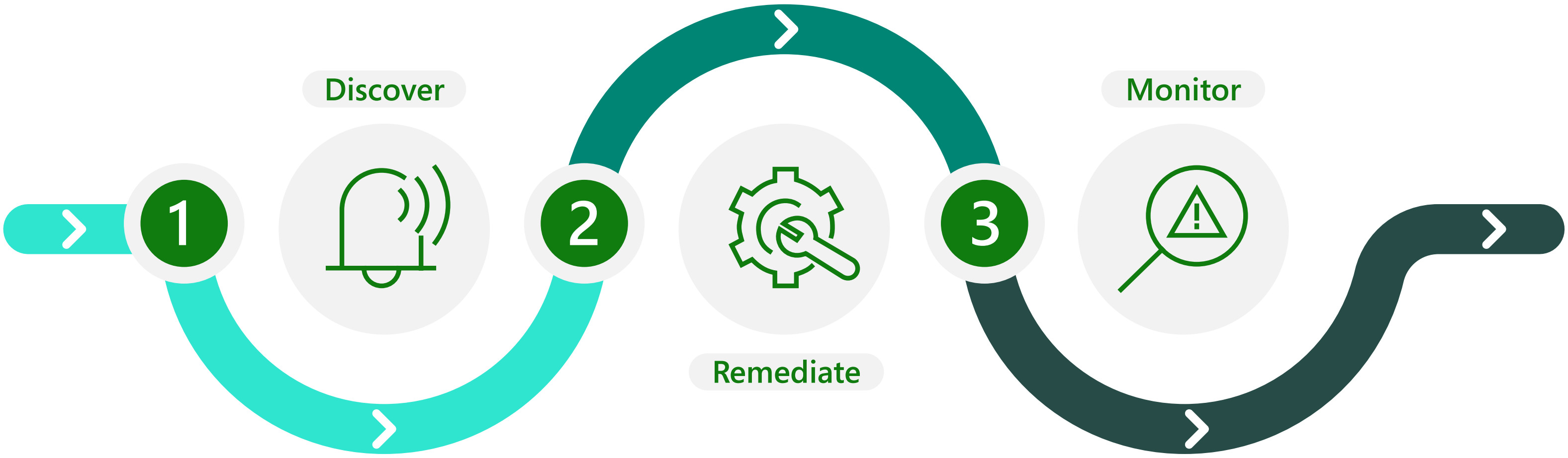
Take the example of a contractor who left a project or a DevOps engineer who moved to another team yet still retains access to his/her original role. Often, the identity is never removed from a role even if it no longer performs the job function. It quickly becomes apparent that legacy tools alone are neither granular nor dynamic enough to keep up with the highly automated and digitized environments that define modern infrastructure.

To effectively scale in the cloud an keep their environments secure, organizations must move from a **static, assumption-based model**, to a **continuous, activity-based** one. The limitations of existing manual processes 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 Entitlement 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. 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 analyst Paul 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, organizations must move forward on all axes.



A Lifecycle Approach to Cloud Permissions Management

Tackling cloud permissions through the lens of a lifecycle framework enables organizations to continuously discover, remediate and monitor the activity of every unique user and workload 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
- Cloud providers will continue to add new capabilities and services that will breed tens of thousands of permissions
- The number of identities, specifically for workloads, will grow exponentially

Critical aspects of a cloud permissions lifecycle include the ability to:

- **Discover risk** by uncovering who (identities) is doing what (actions), where (resources), and when across your cloud infrastructure
- **Remediate risk** by ensuring identities have the least number of permissions needed to be productive
- **Monitor risk** by continuously tracking and measuring changes in identity behavior and permission activity and prioritizing alerts based on pre-defined risk

1. 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 workload identities that can touch an organization'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 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 organization.

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 workload 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.

2. Remediate Risk

A CIEM solution should combine the visibility of real-time and historical activity data with a simple, automated remediation mechanism and offer organizations 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 organization'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.

3. Monitor Risk

To maintain control and security across clouds, organizations 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 organizations to continuously track the activity patterns of all unique human and workload identities across multiple cloud deployments.

Ideally, organizations 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.

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.

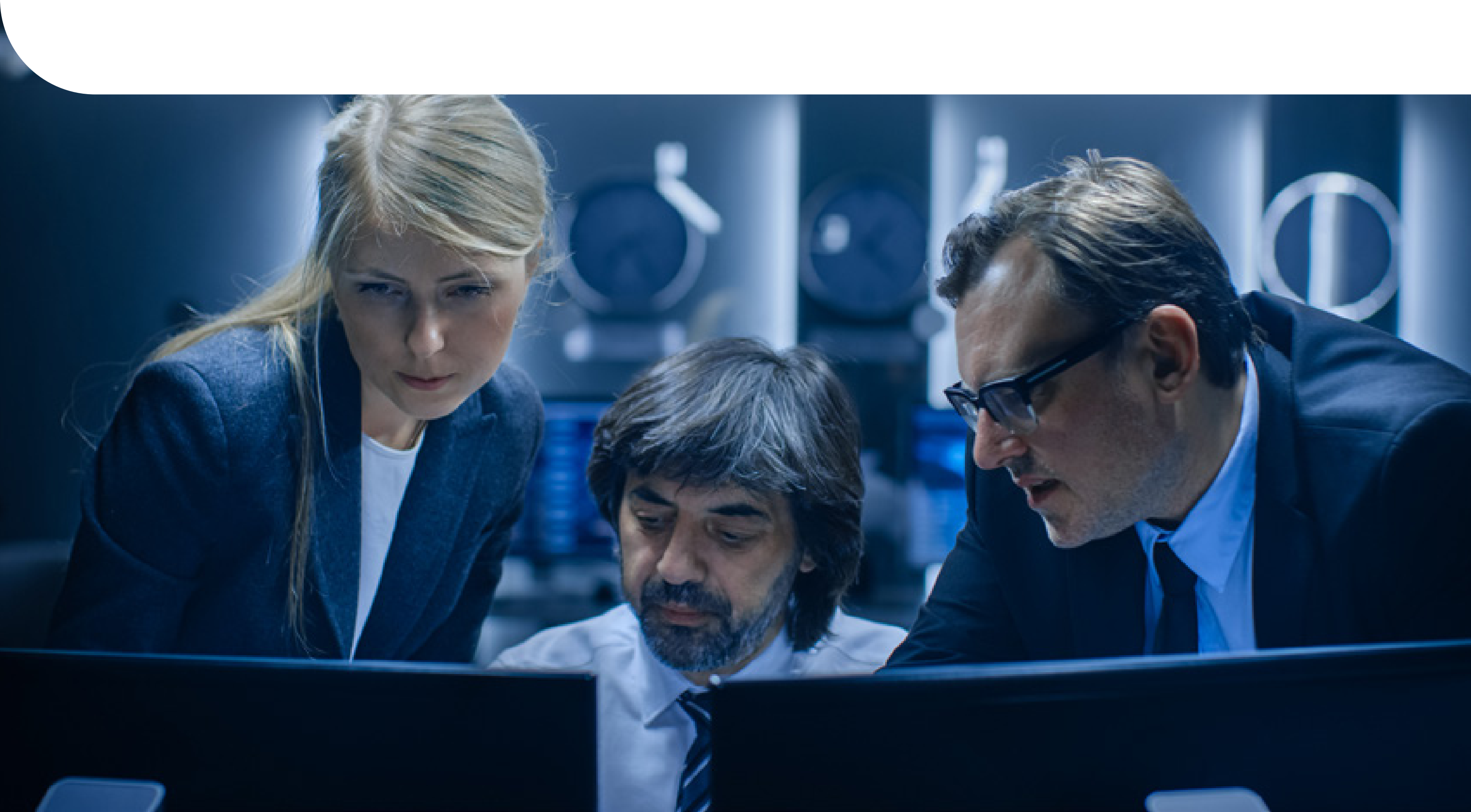
Conclusion

Cloud Security is only as good as an organization's ability to continuously control the level of access privileged human and workload 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 organizations a practical, scalable, cloud-native alternative to existing IAM products and manual methods that don't work in the cloud by empowering them to continuously enforce the principle of least privilege principle at cloud scale.

About CloudKnox Permissions Management

CloudKnox Permissions Management provides a single, unified platform to manage permissions for all identities – users and workloads – across all major cloud infrastructures. It allows organizations to discover, monitor, and remediate permissions risks and achieve Zero Trust security by implementing least privilege access in their entire digital estate.

If you are interested in trying out CloudKnox Permissions Management and would like to join our Public Preview please fill out the form at: aka.ms/CloudKnoxPublicPreview.



Microsoft Security