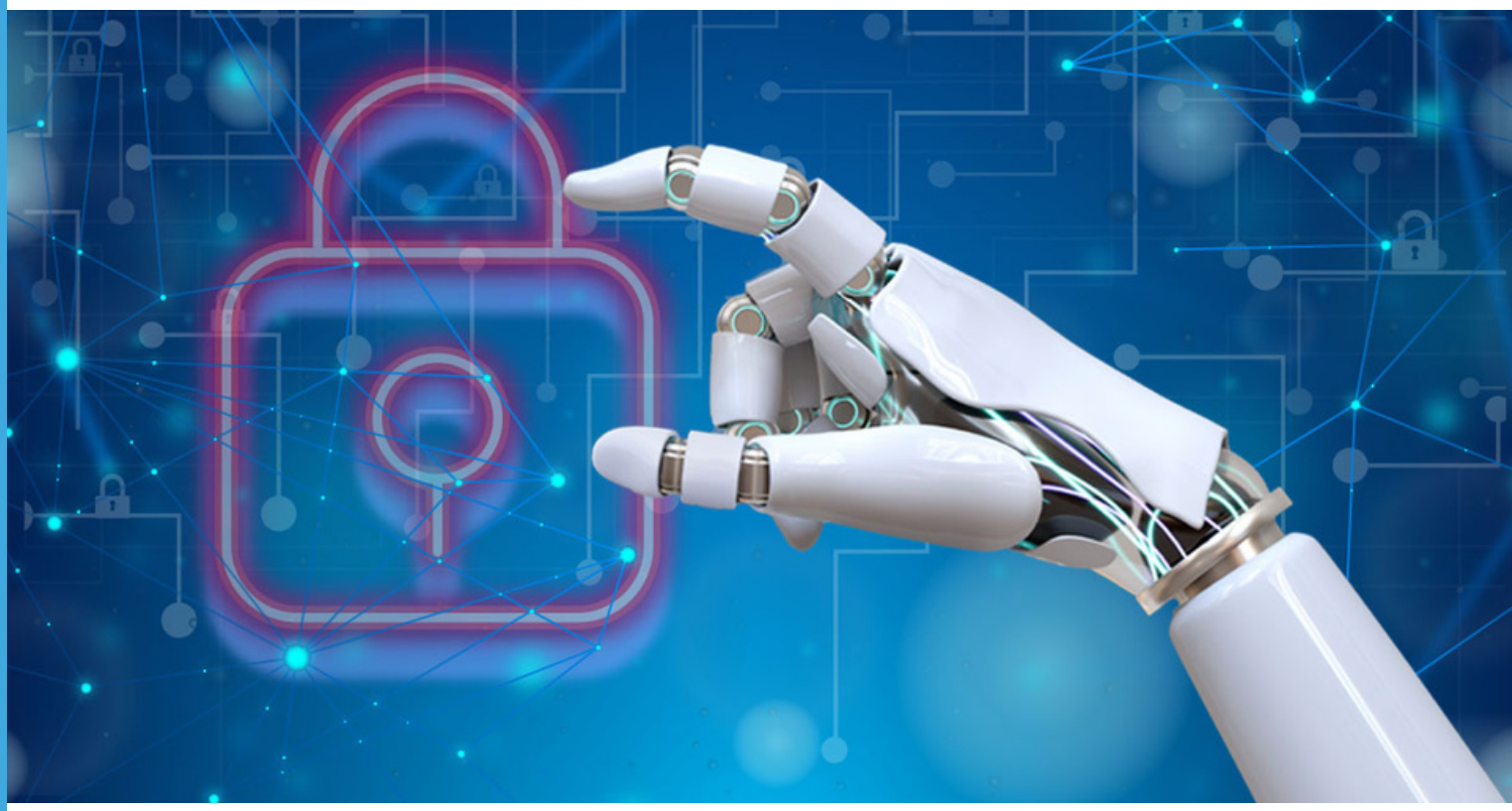
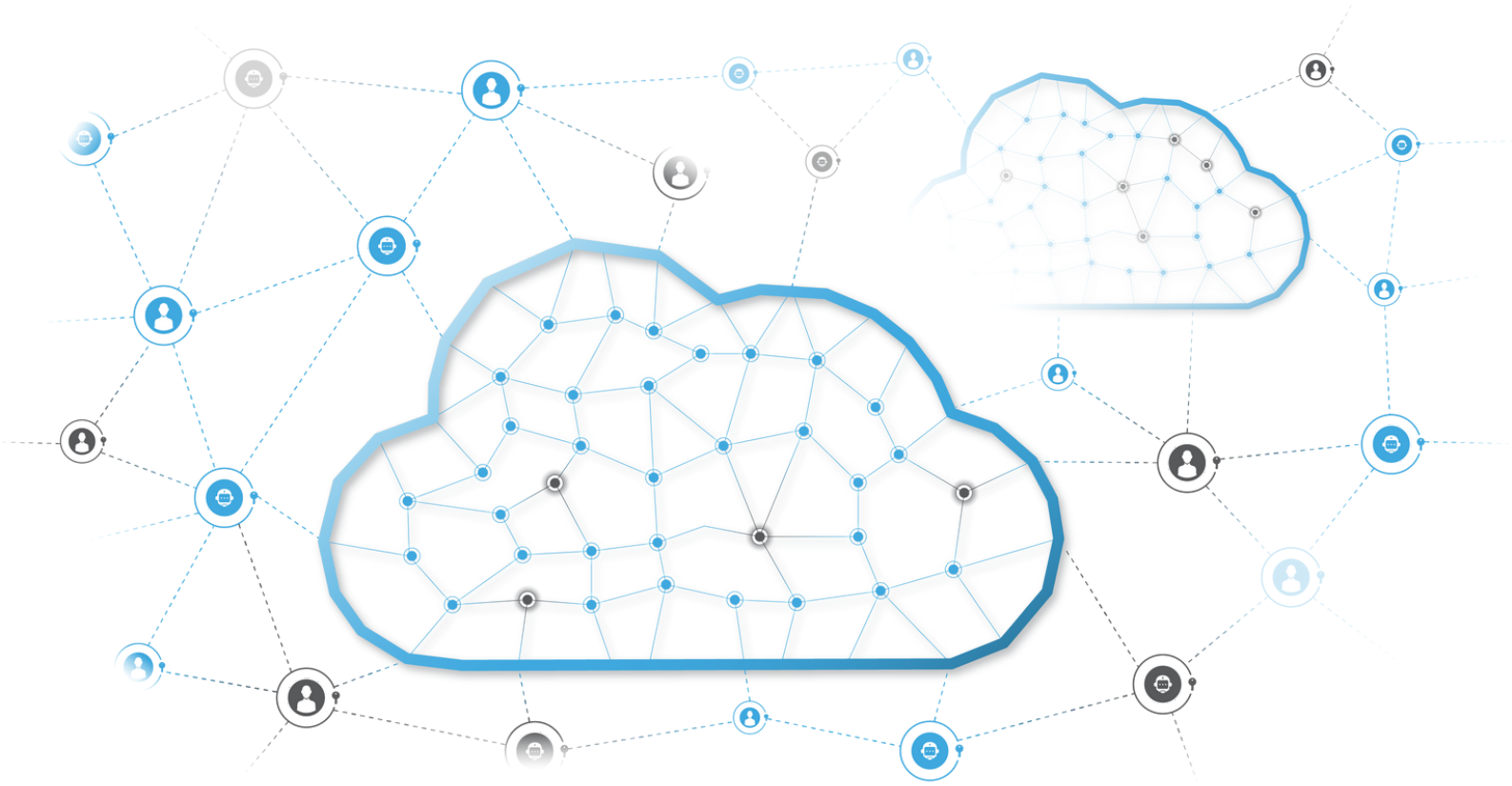
STACK IDENTITY

Navigating the Complexity of Machine Identities in the Gen AI and LLM Era

White Paper





Executive Summary

Only 4% identities are human

STACK IDENTITY SHADOW ACCESS REPORT

The significant increase in Cloud Services, Data, and API-based access has exploded the number of identities. Stack Identity' Shadow Access Impact Report reveals that only 4% of identities as human, while the remaining are non-human identities. Cyber Ark 2022 Identity Security Threat Landscape reports that 68% of machines have access to sensitive data. To make things more complicated, Gen AI and LLMs have accelerated this trend, necessitating a fundamental rethink of how enterprises manage the risk of machine identities.

The rise of Generative AI (Gen AI) and Large Language Models (LLMs), is reshaping the digital landscape, presenting both unprecedented opportunities and unforeseen risks. At the core of this transformation are machine identities and digital entities that are associated with service accounts, workloads, and cloud services.

The significant increase in Cloud Services, Data, and API-based access has exploded the number of identities. Stack Identity' Shadow Access Impact Report reveals that only 4% of identities as human, while the remaining are non-human identities.

Cyber Ark 2022 Identity Security Threat Landscape reports that 68% of machines have access to sensitive data. To make things more complicated, Gen AI and LLMs have accelerated this trend, necessitating a fundamental rethink of how enterprises manage the risk of machine identities.

IAM and Identity teams have to evolve quickly evolve their practices to address the exploding attack surface of machine identities and answer three fundamental questions.

WHAT ARE MACHINE IDENTITIES?

Gartner defines machine identities as separate from human identities (employees, partners, vendors, customers, consultants, etc.), and also groups machine identities into two subgroups:

• Devices – Mobile devices, IoT/OT devices, desktop computers, code signing, etc.

• Workloads – Containers, virtual machines, applications, services, etc.

Machine Identities use cryptographic keys, tokens, passkeys, and claims for authentication and authorization.

EVOLUTION OF MACHINE IDENTITIES

The evolution of Machine Identities has seen tremendous traction over time, let's look at how machine identities have evolved over distinct phases:

1. Service Accounts were created to automate routine and repetitive tasks in the network without human intervention. Service accounts are often provided a level of privileged access similar to that of admin which may pose critical security risks if not monitored.

2. BYOD and SaaS led to a dramatic spike in the number of machine identities for authenticating and authorizing access to unmanaged devices.

3. RPAs and Automation - The automation of routine human tasks through Robotic Process Automation (RPAs) further triggered the number of machine identities within enterprises.

4. Cloud Native Era- In this era, every entity deployed in the cloud, including instances, services, containers, and APIs, required a trusted machine identity, marking the third phase of expansion.

5. LLMs and Generative AI are expected to witness an explosion in machine identity growth, posing unprecedented challenges.

RISK CHARACTERISTICS OF MACHINE IDENTITIES

According to the 2022 CyberArk Report on Massive Growth of Digital Identities, the number of human identities has remained relatively flat in contrast to the exponential rise in machine identities in recent years. The number of machine identities outweighs the number of human identities by as much as 45 times on average.

The threat that Machine identities pose is distinct when compared to human identities:

Governing Access: Machine identities often rely on static credentials with infrequent updates, making them challenging to manage compared to human identities with multi-factor authentication (MFA) and password resets.

Programmatic Access: In Cloud Native environments, machine identities can programmatically acquire and elevate their privileges, diverging from the static access traditionally associated with Service Accounts.

Delegated Access: The rise of automation enables users to delegate access to applications, allowing applications to acquire and expand their privileges, leading to potentially toxic combinations.

Reshared/Abusive Access: Instances where human users share their access with applications or abuse machine identities for personal access become more prevalent, introducing new governance challenges.

SCENARIOS ILLUSTRATING RISK

Three scenarios highlight the unique challenges associated with machine identities:

Data Scientist's ML Stack: A data scientist's ML stack accessing datasets using the data scientist's credentials poses a governance violation, emphasizing the need for improved access controls in the era of Gen AI.

EC2 Instance Profile Abuse: Human users abusing machine identities, such as an EC2 instance profile, for unauthorized access to cloud services, underscores the need for robust governance in Cloud Native environments.

API Key Vulnerabilities: API keys, intended for programmatic access, can be exploited by human users to gain unauthorized access to enterprise data, necessitating enhanced security measures.

Data Scientist's ML Stack: A data scientist's ML stack accessing datasets using the data scientist's credentials poses a governance violation, emphasizing the need for improved access controls in the era of Gen AI.

EC2 Instance Profile Abuse: Human users abusing machine identities, such as an EC2 instance profile, for unauthorized access to cloud services, underscores the need for robust governance in Cloud Native environments.

API Key Vulnerabilities: API keys, intended for programmatic access, can be exploited by human users to gain unauthorized access to enterprise data, necessitating enhanced security measures.

MITIGATING MACHINE IDENTITY RISKS

The new realities of the LLM and Gen AI era, only suggest that companies today have far more vulnerabilities than they can count. To address these challenges, organizations must adopt a comprehensive approach.

Inventory: Develop a comprehensive inventory of machine identities, establishing a source of truth for Machine identity management.

Build Machine Identity Baseline: Establish a new baseline for machine identities, that defines how machine identities need to operate in your environment.

Access and Authorization Map: Create a detailed map outlining active pathways between machine identities and their target access, identifying potential security risks, anomalous behaviors, vulnerabilities, and threats.

Policy Violations: Regularly assess and identify policy violations, including unauthorized or unapproved identities, shared identities, delegated access, and toxic combinations.

Implement Governance Processes: Develop and implement robust processes to secure, control, and govern machine identities against the desired baseline, ensuring compliance with security policies.

SUMMARY

As enterprises embrace the transformative power of Gen AI and LLMs, the surge in machine identities poses unparalleled risks to security, operations, compliance, and governance. Business and security leaders must recognize this paradigm shift and take immediate action to implement policies, guardrails, processes, and technologies to harness the potential of Cloud and AI technologies while mitigating the risks associated with machine identities. The future resilience of enterprises depends on their ability to navigate and govern the complex landscape of machine identities in the evolving digital era.

REFERENCES

1) For further reading on Machine Identities in the context of IAM, refer to the Cloud Security Alliance's insightful resource: Machine Identity in Cybersecurity and IAM.

2) Stack Identity Shadow Access Impact Report

3) Cyber Ark 2022 Identity Security Threat Landscape