Microsoft

Secure Identities:

Strengthening identity protection in the face of highly sophisticated attacks

Digital identity security has never been so crucial. In the early days of the cloud, digital identity primarily protected our files and email. But today, it also protects applications, IoT, and OT systems connected to the physical world. In support of critical systems ranging from defense to finance to transportation to healthcare and more, today's digital identity systems must support more collaboration models, more remote work scenarios, more application identities, and more compliance requirements than ever before.

The cloud provides so many opportunities for innovation and productivity. As the benefits of digital transformation have become clear and digital identity has evolved to embrace more of them, virtually all of the systems and processes which impact our daily lives have moved into the cloud—each protected by digital identity.

Unfortunately, the incredible value of the systems protected by digital identity has also attracted more sophisticated and persistent attackers than ever before—criminals who seek to subvert digital identity systems for fraud, extortion, espionage, or direct harm. Advanced nation state and criminal syndicate actors have engaged in attacks against identity systems, and collaboration and code sharing tools have allowed copycat attackers to massively increase the scale of their sophisticated attacks.

The scale, speed, and sophistication of today's identity attacks are unprecedented. According to the Microsoft Digital Defense Report, 2023, the number of attempted attacks on passwords detected and blocked by Microsoft increased more than tenfold compared to the same period in 2022, from around 3 billion per month to over 30 billion. This translates to an average of 4,000 password attacks per second targeting Microsoft cloud identities this year.

At Microsoft, our investment and commitment to security runs like a current under everything we build, everything we protect, and everything we touch. This isn't limited to identity—we build hardware, devices, and accessories, and the software that helps organizations run their businesses and consumers run their lives. Whether facilitating users interacting with systems, or digital workloads interacting with each other, identity security plays a critical role in all these domains. Identity is the "front door" to all our diverse offerings. Through single sign-on (SSO) and all the resources our customers care about, our identity security surface area is vast and virtually unmatched.

Securing our vast surface area gives us a singular, unique view into the threats we all face in today's cyber world. We process over 750 million signals per second, synthesizing them with sophisticated data analytics and AI algorithms to understand and protect against criminal and nation state cyberactivity. The public rarely hears about the 4,000 identity authentication threats we block every second of every day. But each attack we intercept helps us hone our security systems and skills. We have increased our cybersecurity investment of $20 billion over five years and made significant strides in technology, but there is always a need to continuously stay ahead. It's our responsibility to respond, inform, innovate, and educate our customers *and* the industry on how we deliver solutions that meet the highest possible standard of security in the new era of AI and beyond. Responding to this responsibility, Microsoft has introduced the Secure Future Initiative—a multi-year commitment to advance the way we design, build, test, and operate our technology to meet the highest possible standard of security.

In this paper, we will explore the tangible actions we're taking to continually improve our identity platforms and explain our engineering advances to strengthen identity protection, focusing on key management and identity.

Strengthening identity by making keys inaccessible to advanced adversaries

In the past year, cyberattacks have impacted 120 countries with an alarming increase in state-run, government-sponsored espionage. More than 40 percent of attacks in the past year touched government or private-sector organizations that build and maintain critical infrastructure. In response, we are accelerating our investments to protect customers against escalating identity attacks by nation state sponsored actors. These actions build on learnings from Midnight Blizzard (NOBELIUM)'s use of stolen login credentials in attacks against SolarWinds customers and our investigation into how a China-based adversary we track as Storm-0558 used forged authentication tokens to access customers' email.

Storm-0558 and Midnight Blizzard are examples of advanced and nation state actors that have demonstrated interest in key systems to gain illicit system access. To address these targeted attacks by the most sophisticated actors, we are implementing several advancements in our key management systems to be resilient against such attacks. Broadly, these fall into five categories:

- Enhanced automation for key management (zero touch)
- Storing and managing keys in secure hardware (HSM)
- Ensuring keys are protected in memory (confidential compute service)
- Increasing key rotation frequency (rapid key rotation)
- Monitoring key usage for suspicious activity (built in telemetry)

Full automation for key management (zero touch)

When the key impacted by Storm-0558 was created in 2016, key creation and management was a manual process. An engineer would call a key creation API from a secure workstation to generate a keypair, then propagate the private key into secure, encrypted storage. Any manual process creates opportunities for error and exploitation; the key could be incorrectly deleted, propagated, or saved.

Today, our enterprise key management is done entirely through automation, heavily audited and monitored to avoid the introduction of human error or exploitation. And in the near future, our consumer keys will use this same system. No human touch for keys means keys are not manually managed by humans at all, but are instead generated, stored, and used automatically by systems. They will be managed with no human access to keys by host infrastructure, at rest, or in transit. This ensures the key was propagated correctly, to the right place, and deleted correctly. Our commitment improves the security of keys by removing the risk of human-introduced errors or human compromise.

Identity token signing key management operations will be executed only by signed code through heavily instrumented commands. The key management systems represent the "inner ring" of our defenses, will be fully isolated from all other systems in our production environment, including changes to prevent backup, automated crash dump captures, or other access to these systems. This will ensure the integrity of the identity token signing systems and prevent any data leaving the signing environment, even if surrounding infrastructure is compromised.

Storing and managing keys in secure hardware (HSM)

We now aim to have all signing keys stored in Hardware Security Modules (HSM). Storing keys in these elements makes the keys invulnerable to accidental or intentional storage access, including services which backup our systems or manage incidents. Our new key management system will use Microsoft's Managed HSM to provide secure key storage in HSM hardware, preventing any memory or disk-based access either by accident or intent. While current systems are implemented to prevent access to storage, moving the key storage into FIPS 140-2 Level 3 validated Hardware Security Modules (HSM) makes such events impossible.

Ensuring keys are protected in memory (confidential compute service)

We will also ensure keys cannot be exfiltrated—even if the underlying processes become compromised, or host processes attempt to directly inspect memory—by using Microsoft Azure's confidential compute service to manage our signing processes.

Confidential compute architecture can help:

- Prevent unauthorized access: Run sensitive data in the cloud with the best data protection possible, with little to no change from what happens today.
- Meet regulatory compliance: Keep full control of data while migrating workloads to the cloud to satisfy government regulations for protecting personal information and secure organizational IP.
- Enable secure and untrusted collaboration: Tackle industry-wide work-scale problems by combing data across organizations—even competitors—to unlock broad data analytics and deeper insights.
- Isolate processing: Offer a new wave of products that remove liability on private data with blind processing. User data can't even be retrieved by the service provider.

By combining HSM storage with confidential compute, identity token signing **keys are protected from access in storage and in use.**

Rapid Key Rotation

While the protections above vastly reduce or eliminate the possibility of a compromised key, we will also regularly and more rapidly retire and rotate keys in the identity infrastructure, so in the unlikely event a key is acquired again, attackers will have little time to use it. This is an example of "defense in depth"—building multiple overlapping mitigations to maximize key security. Soon, all keys will be rotated at least every 4 weeks, so that even if there is an operational or system error that results in a key leaving secure storage, the window for discovery and exploitation will be limited by design.

Monitoring key usage for suspicious activity

Even with all the measures above, we have to assume the worst can happen (the Zero-Trust principle of "Assume Breach"). So, for each aspect of the system, we are defining security invariants—things that must hold—and then explicitly building system logging, detections, and alerting to make sure we know instantly that something is behaving outside of our expectations and that we can respond quickly. Where possible, we are automating the response so the system is able to contain issues as soon as they are detected. Despite the fact that our systems design should make actual issues—and thus alerting—impossible, we are also building in system drills to ensure the systems and team are exercising these "defense in depth" muscles for detection, alerting, and mitigation of any issues.

Practice governance and utilized standard identity libraries to eliminate errors

Improved governance of services across Microsoft will ensure all Microsoft services and infrastructure benefit from and adhere to modern best practices for identity security. A major component of this will be improved and expanded **standard identity libraries**, which Microsoft services will be required to use. These libraries will also be freely available to our customers who want to take advantage of these cutting-edge identity security standards.

We're committed to making sure all developers have access to standard identity libraries and get help from tools like GitHub Copilot, so common human errors that occur in implementation of identity features essentially disappear. The Microsoft Authentication Library (MSAL) enables developers to acquire security tokens from the Microsoft identity platform to authenticate users and access secured web APIs. It can be used to provide secure access to Microsoft Graph, other Microsoft APIs, third-party web APIs, or your own web API. MSAL supports many different application architectures and platforms including .NET, JavaScript, Java, Python, Android, and iOS.

We are enforcing standard identity libraries across Microsoft including implementation of standardized token validation, advanced identity defenses like token binding, continuous access evaluation (CAE), advanced application attack detections, and standardized application-side identity and telemetry and logging to help rapidly catch anomalies. Microsoft will implement new classes of identity threat detection - including detection of forged tokens - for this telemetry from our own applications. For applications not developed by Microsoft, we will provide guidance and Sentinel integration to help customers maintain parity detection on other applications. applications.

Because these capabilities are critical for *all* applications our customers use, **we are making these advanced capabilities freely available to non-Microsoft application developers** through standard identity libraries. By using standard identity libraries, companies can reduce the risk of security vulnerabilities being introduced into the codebase, improve visibility into security risks, make it easier to control access to sensitive data and resources, and ensure that critical information for security investigations are captured.

Moving forward, together, to safeguard our customers in unprecedented times

Microsoft continues to invest heavily in the security and privacy of both our consumer (Microsoft Account) and enterprise (Microsoft Entra ID, formerly Azure Active Directory) identity solutions. And with the addition of Microsoft's Secure Future Initiative, we are making meaningful engineering advances across three areas, including secure identities. We have focused on the creation, implementation, and improvement of identity-related specifications that foster strong authentication, secure sign-on, sessions, API security, and other critical infrastructure tasks, as part of the community of standards experts within official standards bodies such as IETF, W3C, or the OpenID Foundation.

Our own engineering advances and industry standards are a good foundation, but safeguarding our information, our organizations, and all our digital lives will take a cooperative, combined effort that includes industry partnerships, government policy, and our entire security community coming together.

Our Microsoft Identity Bounty Program is one of the highest-paying in the industry, and an important way our community comes together to improve security. We invite researchers across the globe to identify vulnerabilities in identity products and services and share them with our team. Qualified submissions are eligible for bounty rewards up to $100,000 USD. And not only do researchers benefit from submitting bugs, Microsoft is also committed to transparency and disclosure, allowing researchers to publish their findings—both to benefit their own reputation and to benefit the community and industry as a whole. In 2022, Microsoft paid out a total of $13.8 million in bug bounties, with the highest payout being $80,000.

Additionally, Microsoft partners with our large, supportive security research community:

- Across a broad scope: Microsoft's bounty program covers a wide range of products and services, including Windows, Office, Azure, and Xbox. This gives security researchers a lot of opportunities to find and report vulnerabilities.
- With quick response: Microsoft has a dedicated team who quickly reviews and responds to bug reports. This means security researchers can be confident that their findings will be taken seriously and that they will be rewarded promptly.

At Microsoft, we believe security is a team sport. We also believe that the scale, speed, and sophistication of these attacks are unprecedented for the industry. We collectively need to come together to address these attacks and ensure that we deliver solutions that meet the highest possible standard of security. There is strength in numbers, so when we share what we're learning, collectively we have a greater chance of bringing down the threat actors and keeping our customers safe. In the case of Storm-0558, we are grateful for the serious safeguards our government customers have in place, and the cybersecurity analysts they employ that implemented alerts to quickly identify this intrusion so we could investigate and mitigate it promptly.

In addition to this commitment to transparency, we are committed to continuous improvement with our rigorous, formal post-incident response process. This means reviewing every finding and rectifying flaws not only in affected systems but across our product portfolio, as well as sharing our findings with the larger security community.

These engineering advances in identity cover just one area in which we are improving the security of our technology. Read about our other engineering advances in EVP of Microsoft Security Charlie Bell's memo to all Microsoft engineers. For more information about the broader Microsoft's Secure Future Initiative, read Microsoft President, Brad Smith's post on how we're further safeguarding our customers in unprecedented times.

› Learn more about how Microsoft builds security into everything we design, develop, and deliver.