What are risk detections?

Article • 08/21/2024 • 22 contributors

In this article

Microsoft Entra ID Protection provides organizations with information to suspicious activity in their tenant and allows them to respond quickly to prevent further risk occurring. Risk detections are a powerful resource that can include any suspicious or anomalous activity related to a user account in the directory. ID Protection risk detections can be linked to an individual user or sign-in event and contribute to the overall user risk score found in the Risky Users report.

User risk detections might flag a legitimate user account as at risk, when a potential threat actor gains access to an account by compromising their credentials or when they detect some type of anomalous user activity. Sign-in risk detections represent the probability that a given authentication request isn't the authorized owner of the account. Having the ability to identify risk at the user and sign-in level is critical for customers to be empowered to secure their tenant.

Risk levels

ID Protection categorizes risk into three tiers: low, medium, and high. Risk levels calculated by our machine learning algorithms and represent how confident Microsoft is that one or more of the user's credentials are known by an unauthorized entity.

- A risk detection with risk level **High** signifies that Microsoft is highly confident that the account is compromised.
- A risk detection with risk level **Low** signifies that there are anomalies present in the sign-in or a user's credential, but we're less confident that these anomalies mean the account is compromised.

Many detections can fire at more than one of our risk levels depending on the number or severity of the anomalies detected. For example, Unfamiliar sign-in properties might fire at high, medium, or low based on the confidence in the signals. Some detections, like Leaked Credentials and Verified Threat Actor IP are always delivered as high risk.

This risk level is important when deciding which detections to prioritize, investigate, and remediate. They also play a key role in configuring risk based Conditional Access policies as each policy can be set to trigger for low, medium, high, or no risk detected. Based on the risk tolerance of your organization, you can create policies that require MFA or password reset when ID Protection detects a certain risk level for one of your users. These policies can guide the user to self-remediate to resolve the risk.

> **Important**
>
> All "low" risk level detections and users will persist in the product for 6 months, after which they will be automatically aged out to provide a cleaner investigation experience. Medium and high risk levels will persist until remediated or dismissed.
>
> Based on the risk tolerance of your organization, you can create policies that require MFA or password reset when ID Protection detects a certain risk level. These policies might guide the user to self-remediate and resolve the risk or block depending on your tolerances.

Real-time and offline detections

ID Protection utilizes techniques to increase the precision of user and sign-in risk detections by calculating some risks in real-time or offline after authentication. Detecting risk in real-time at sign-in gives the advantage of identifying risk early so that customers can quickly investigate the potential compromise. On detections that calculate risk offline, they can provide more insight as to how the threat actor gained access to the account and the impact on the legitimate user. Some detections can be triggered both offline and during sign-in, which increases confidence in being precise on the compromise.

Detections triggered in real-time take 5-10 minutes to surface details in the reports. Offline detections take up to 48 hours to surface in the reports, as it takes time to evaluate properties of the potential risk.

> **Note**
>
> Our system might detect that the risk event that contributed to the risk user risk score was either:
>
> - A false positive
> - The user risk was remediated by policy by either:
>   - Completing multifactor authentication
>   - Secure password change
>
> Our system will dismiss the risk state and a risk detail of **AI confirmed sign-in safe** will show and no longer contribute to the user's overall risk.

On risk-detailed data, **Time Detection** records the exact moment a risk is identified during a user's sign-in, which allows for real-time risk assessment and immediate policy application to safeguard the user and organization. **Detection last updated** shows the latest update to a risk detection, which could be due to new information, risk level changes, or administrative actions, and ensures up-to-date risk management.

These fields are essential for real-time monitoring, threat response, and maintaining secure access to organizational resources.

Risk detections mapped to riskEventType

Expand table

Risk detection	Detection type	Type	riskEventType
Sign-in risk detections			
Activity from anonymous	Offline	Premium	riskyIPAddress

IP address			
Additional risk detected (sign-in)	Real-time or Offline	Nonpremium	generic = Premium detection classification for non-P2 tenants
Admin confirmed user compromised	Offline	Nonpremium	adminConfirmedUserCompromised
Anomalous Token	Real-time or Offline	Premium	anomalousToken
Anonymous IP address	Real-time	Nonpremium	anonymizedIPAddress
Atypical travel	Offline	Premium	unlikelyTravel
Impossible travel	Offline	Premium	mcasImpossibleTravel
Malicious IP address	Offline	Premium	maliciousIPAddress
Mass Access to Sensitive Files	Offline	Premium	mcasFinSuspiciousFileAccess
Microsoft Entra threat intelligence (sign-in)	Real-time or Offline	Nonpremium	investigationsThreatIntelligence
New country	Offline	Premium	newCountry
Password spray	Offline	Premium	passwordSpray
Suspicious browser	Offline	Premium	suspiciousBrowser
Suspicious inbox forwarding	Offline	Premium	suspiciousInboxForwarding

Suspicious inbox manipulation rules	Offline	Premium	mcasSuspiciousInboxManipulationRules
Token issuer anomaly	Offline	Premium	tokenIssuerAnomaly
Unfamiliar sign-in properties	Real-time	Premium	unfamiliarFeatures
Verified threat actor IP	Real-time	Premium	nationStateIP
User risk detections			
Additional risk detected (user)	Real-time or Offline	Nonpremium	generic = Premium detection classification for non-P2 tenants
Anomalous user activity	Offline	Premium	anomalousUserActivity
Attacker in the Middle	Offline	Premium	attackerinTheMiddle
Leaked credentials	Offline	Nonpremium	leakedCredentials
Microsoft Entra threat intelligence (user)	Real-time or Offline	Nonpremium	investigationsThreatIntelligence
Possible attempt to access Primary Refresh Token (PRT)	Offline	Premium	attemptedPrtAccess

Suspicious API Traffic	Offline	Premium	suspiciousAPITraffic
Suspicious sending patterns	Offline	Premium	suspiciousSendingPatterns
User reported suspicious activity	Offline	Premium	userReportedSuspiciousActivity

Premium detections

The following premium detections are visible only to Microsoft Entra ID P2 customers.

Premium sign-in risk detections

Activity from anonymous IP address

Calculated offline. This detection is discovered using information provided by Microsoft Defender for Cloud Apps. This detection identifies that users were active from an IP address identified as an anonymous proxy IP address.

Anomalous token

Calculated in real-time or offline. This detection indicates abnormal characteristics in the token, such as an unusual lifetime or a token played from an unfamiliar location. This detection covers Session Tokens and Refresh Tokens.

Anomalous token is tuned to incur more noise than other detections at the same risk level. This tradeoff is chosen to increase the likelihood of detecting replayed tokens that might otherwise go unnoticed. There's a higher than normal chance that some of the sessions flagged by this

detection are false positives. We recommend investigating the sessions flagged by this detection in the context of other sign-ins from the user. If the location, application, IP address, User Agent, or other characteristics are unexpected for the user, the administrator should consider this risk as an indicator of potential token replay.

Tips for investigating anomalous token detections.

Atypical travel

Calculated offline. This risk detection type identifies two sign-ins originating from geographically distant locations, where at least one of the locations might also be atypical for the user, given past behavior. The algorithm takes into account multiple factors including the time between the two sign-ins and the time it would take for the user to travel from the first location to the second. This risk might indicate that a different user is using the same credentials.

The algorithm ignores obvious "false positives" contributing to the impossible travel conditions, such as VPNs and locations regularly used by other users in the organization. The system has an initial learning period of the earliest of 14 days or 10 logins, during which it learns a new user's sign-in behavior.

Tips for investigating atypical travel detections.

Impossible travel

Calculated offline. This detection is discovered using information provided by Microsoft Defender for Cloud Apps. This detection identifies user activities (in a single or multiple sessions) originating from geographically distant locations within a time period shorter than the time it takes to travel from the first location to the second. This risk might indicate that a different user is using the same credentials.

Malicious IP address

Calculated offline. This detection indicates sign-in from a malicious IP address. An IP address is considered malicious based on high failure rates because of invalid credentials received from the IP address or other IP reputation sources. In some instances, this detection triggers on previous malicious activity.

Tips for investigating malicious IP address detections.

Mass access to sensitive files

Calculated offline. This detection is discovered using information provided by Microsoft Defender for Cloud Apps. This detection looks at your environment and triggers alerts when users access multiple files from Microsoft SharePoint Online or Microsoft OneDrive. An alert is triggered only if the number of accessed files is uncommon for the user and the files might contain sensitive information.

New country

Calculated offline. This detection is discovered using information provided by Microsoft Defender for Cloud Apps. This detection considers past activity locations to determine new and infrequent locations. The anomaly detection engine stores information about previous locations used by users in the organization.

Password spray

Calculated offline. A password spray attack is where multiple identities are attacked using common passwords in a unified brute force manner. The risk detection is triggered when an account's password is valid and has an attempted sign in. This detection signals that the user's password has correctly been identified through a password spray attack, not that the attacker was able to access any resources.

Tips for investigating malicious IP address detections.

Suspicious browser

Calculated offline. Suspicious browser detection indicates anomalous behavior based on suspicious sign-in activity across multiple tenants from different countries in the same browser.

Tips for investigating suspicious browser detections.

Suspicious inbox forwarding

Calculated offline. This detection is discovered using information provided by Microsoft Defender for Cloud Apps. This detection looks for suspicious email forwarding rules, for example, if a user created an inbox rule that forwards a copy of all emails to an external address.

Suspicious inbox manipulation rules

Calculated offline. This detection is discovered using information provided by Microsoft Defender for Cloud Apps. This detection looks at your environment and triggers alerts when suspicious rules that delete or move messages or folders are set on a user's inbox. This detection might indicate: a user's account is compromised, messages are being intentionally hidden, and the mailbox is being used to distribute spam or malware in your organization.

Token issuer anomaly

Calculated offline. This risk detection indicates the SAML token issuer for the associated SAML token is potentially compromised. The claims included in the token are unusual or match known attacker patterns.

Tips for investigating token issuer anomaly detections.

Unfamiliar sign-in properties

Calculated in real-time. This risk detection type considers past sign-in history to look for anomalous sign-ins. The system stores information

about previous sign-ins, and triggers a risk detection when a sign-in occurs with properties that are unfamiliar to the user. These properties can include IP, ASN, location, device, browser, and tenant IP subnet. Newly created users are in a "learning mode" period where the unfamiliar sign-in properties risk detection is turned off while our algorithms learn the user's behavior. The learning mode duration is dynamic and depends on how much time it takes the algorithm to gather enough information about the user's sign-in patterns. The minimum duration is five days. A user can go back into learning mode after a long period of inactivity.

We also run this detection for basic authentication (or legacy protocols). Because these protocols don't have modern properties such as client ID, there's limited data to reduce false positives. We recommend our customers to move to modern authentication.

Unfamiliar sign-in properties can be detected on both interactive and non-interactive sign-ins. When this detection is detected on non-interactive sign-ins, it deserves increased scrutiny due to the risk of token replay attacks.

Selecting an unfamiliar sign-in properties risk allows you to see more info showing more detail about why this risk triggered.

Verified threat actor IP

Calculated in real-time. This risk detection type indicates sign-in activity that is consistent with known IP addresses associated with nation state actors or cyber crime groups, based on data from the Microsoft Threat Intelligence Center (MSTIC).

Premium user risk detections

Anomalous user activity

Calculated offline. This risk detection baselines normal administrative user behavior in Microsoft Entra ID, and spots anomalous patterns of behavior like suspicious changes to the directory. The detection is triggered against the administrator making the change or the object that was changed.

Attacker in the Middle

Calculated offline. Also known as Adversary in the Middle, this high precision detection is triggered when an authentication session is linked to a malicious reverse proxy. In this kind of attack, the adversary can intercept the user's credentials, including tokens issued to the user. The Microsoft Security Research team uses Microsoft 365 Defender to capture the identified risk and raises the user to **High** risk. We recommend administrators manually investigate the user when this detection is triggered to ensure the risk is cleared. Clearing this risk might require secure password reset or revocation of existing sessions.

Possible attempt to access Primary Refresh Token (PRT)

Calculated offline. This risk detection type is discovered using information provided by Microsoft Defender for Endpoint (MDE). A Primary Refresh Token (PRT) is a key artifact of Microsoft Entra authentication on Windows 10, Windows Server 2016, and later versions, iOS, and Android devices. A PRT is a JSON Web Token (JWT) issued to Microsoft first-party token brokers to enable single sign-on (SSO) across the applications used on those devices. Attackers can attempt to access this resource to move laterally into an organization or perform credential theft. This detection moves users to high risk and only fires in organizations that deploy MDE. This detection is high risk and we recommend prompt remediation of these users. It appears infrequently in most organizations due to its low volume.

Suspicious API traffic

Calculated offline. This risk detection is reported when abnormal GraphAPI traffic or directory enumeration is observed. Suspicious API traffic might suggest that a user is compromised and conducting reconnaissance in the environment.

Suspicious sending patterns

Calculated offline. This risk detection type is discovered using information provided by Microsoft Defender for Office 365 (MDO). This alert is generated when someone in your organization sent suspicious email and is either at risk of being or is restricted from sending email. This detection moves users to medium risk and only fires in organizations that deploy MDO. This detection is low-volume and is seen infrequently in most organizations.

User reported suspicious activity

Calculated offline. This risk detection is reported when a user denies a multifactor authentication (MFA) prompt and reports it as suspicious activity. An MFA prompt not initiated by a user might mean their credentials are compromised.

Nonpremium detections

Customers without Microsoft Entra ID P2 licenses receive detections titled **Additional risk detected** without the detailed information regarding the detection that customers with P2 licenses do. For more information, see the license requirements.

Nonpremium sign-in risk detections

Additional risk detected (sign-in)

Calculated in real-time or offline. This detection indicates that one of the premium detections was detected. Since the premium detections are

visible only to Microsoft Entra ID P2 customers, they're titled **Additional risk detected** for customers without Microsoft Entra ID P2 licenses.

Admin confirmed user compromised

Calculated offline. This detection indicates an administrator selected **Confirm user compromised** in the risky users UI or using riskyUsers API. To see which administrator confirmed this user compromised, check the user's risk history (via UI or API).

Anonymous IP address

Calculated in real-time. This risk detection type indicates sign-ins from an anonymous IP address (for example, Tor browser or anonymous VPN). These IP addresses are typically used by actors who want to hide their sign-in information (IP address, location, device, and so on) for potentially malicious intent.

Microsoft Entra threat intelligence (sign-in)

Calculated in real-time or offline. This risk detection type indicates user activity that is unusual for the user or consistent with known attack patterns. This detection is based on Microsoft's internal and external threat intelligence sources.

Tips for investigating Microsoft Entra threat intelligence detections.

Nonpremium user risk detections

Additional risk detected (user)

Calculated in real-time or offline. This detection indicates that one of the premium detections was detected. Since the premium detections are visible only to Microsoft Entra ID P2 customers, they're titled **Additional risk detected** for customers without Microsoft Entra ID P2 licenses.

Leaked credentials

Calculated offline. This risk detection type indicates that the user's valid credentials leaked. When cybercriminals compromise valid passwords of legitimate users, they often share these gathered credentials. This sharing is typically done by posting publicly on the dark web, paste sites, or by trading and selling the credentials on the black market. When the Microsoft leaked credentials service acquires user credentials from the dark web, paste sites, or other sources, they're checked against Microsoft Entra users' current valid credentials to find valid matches. For more information about leaked credentials, see common questions.

Tips for investigating leaked credentials detections.

Microsoft Entra threat intelligence (user)

Calculated offline. This risk detection type indicates user activity that is unusual for the user or consistent with known attack patterns. This detection is based on Microsoft's internal and external threat intelligence sources.

Tips for investigating Microsoft Entra threat intelligence detections.

Common questions

What if incorrect credentials were used to attempt to sign-in?

ID Protection generates risk detections only when the correct credentials are used. If incorrect credentials are used on a sign-in, it doesn't represent risk of credential compromise.

Is password hash synchronization required?

Risk detections like leaked credentials require the presence of password hashes for detection to occur. For more information about password

hash synchronization, see the article, Implement password hash synchronization with Microsoft Entra Connect Sync.

Why are risk detections generated for disabled accounts?

User accounts in a disabled state can be re-enabled. If the credentials of a disabled account are compromised, and the account gets re-enabled, bad actors might use those credentials to gain access. ID Protection generates risk detections for suspicious activities against these disabled accounts to alert customers about potential account compromise. If an account is no longer in use and won't be re-enabled, customers should consider deleting it to prevent compromise. No risk detections are generated for deleted accounts.

Common leaked credentials questions

Where does Microsoft find leaked credentials?

Microsoft finds leaked credentials in various places, including:

- Public paste sites where bad actors typically post such material.
- Law enforcement agencies.
- Other groups at Microsoft doing dark web research.

Why am I not seeing any leaked credentials?

Leaked credentials are processed anytime Microsoft finds a new, publicly available batch. Because of the sensitive nature, the leaked credentials are deleted shortly after processing. Only new leaked credentials found **after** you enable password hash synchronization (PHS) are processed against your tenant. Verifying against previously found credential pairs isn't done.

I don't see any leaked credential risk events

If you don't see any leaked credential risk events, it is because of the following reasons:

- You don't have PHS enabled for your tenant.
- Microsoft didn't find any leaked credential pairs that match your users.

How often does Microsoft process new credentials?

Credentials are processed immediately after they're found, normally in multiple batches per day.

Locations

Location in risk detections is determined using IP address lookup. Sign-ins from trusted named locations improve the accuracy of Microsoft Entra ID Protection's risk calculation, lowering a user's sign-in risk when they authenticate from a location marked as trusted.

Related content

- Learn about risk-based access policies
- Learn how to investigate risk

Feedback

Was this page helpful? Yes No

Provide product feedback