# SIEM Integration - Use cases

# Overview

|  |  |
| --- | --- |
| **Owner** |  |
| **Contributors** |  |
| **ADO Board Epic / Feature** |  |
| **Sign Off** |  |
| **On this page** | * [Overview](#SIEMIntegration-Usecases-Overview) * [Document Version](#SIEMIntegration-Usecases-DocumentVersio) * [Summary](#SIEMIntegration-Usecases-Summary) * [Use cases & logs required to perform use case](#SIEMIntegration-Usecases-Usecases&logsr) |

# Document Version

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Author** | **Version** | **Description of Changes** |
| 02/02/2022 |  | 1.0 | Baselined |
| 07/02/2022 |  | 1.1 | Updates based on workshop with |
| 16/02/2022 |  | 1.2 | Update with updates to use cases and logs |
| 03/05/2023 |  | 1.3 | Updated for additional use case related to DB data exfiltration detection. See UC-SIEM-7. |

# Summary

As part of the requirement to detect, alert, analyse and respond to threats on the data within BDP4, Security, information and event management (SIEM) will be a solution to satisfy this requirement. SIEM technology collects event log data from a range of sources, identifies activity that deviates from the norm with real-time analysis, and takes appropriate action.

For further context on SIEM, please refer to the following link on the integration between SIEM and BDP4 - [SIEM Integration - High level design](https://bupaau-my.sharepoint.com/wiki/spaces/HDP/pages/2915237955/SIEM+Integration+-+High+level+design) .

The purpose of this document will be to outline the specific use cases for SIEM so that it can be implemented as part of robust alerting and monitoring practices within BDP4.

# Use cases & logs required to perform use case

Outlined below are the use cases for SIEM across four categories: **Collect, Correlate, Identify & Alert**.

This list of use cases is expected to grow over time, however, due to the need for an interim solution to be in place for the Program’s goal to deliver priority reports to the business by end Q1 2023, a set of MVP use cases have been agreed upon and will be the focus of current effort. Subsequent use cases will be prioritised as needed with agreement by security, SIEM SMEs and the Data Program PMs.

Note, Sentinel currently has logs enabled for the following services:

* **AAD**
* **Defender for identity**
* **Cyberarc and Sailpoint for IAM -** For RBAC - onboard system onto Sailpoint for RBAC, access provisioning and ongoing access reviews

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Use case ID** | **Category** | **Priority - MVP, Future** | **Description** | **How to leverage logs with Platform that either is Existing / Non Existing** |
| UC-SIEM-1 | Identify, Alert | MVP 1 | **Unauthorised access:** The solution should be able to generate alerts when there is an attempt to access sensitive data without the required permissions or credentials. This includes but is not limited to:   * Access outside of endorsed RBAC group - report on if an individual or a service has a permission that does not match to the endorsed position on RBAC. * Abnormal usage (Data access) - Thresholds have been set from a usage perspective. When a user, service or event occurs that falls outside of a predefined threshold (as an example, a service attempts to retrieve a large amount of data greater then 100% of the previous 3 days usage within a 10 minute period), an incident is raised within ServiceNow and is assigned to the operations team for triage and resolution.   + **Example scenario:** Person A is a data analyst and will typically work during business hours and run queries with an output size between 10-100MB. Person A runs a query at 11pm on Saturday with an output size of 10GB. This activity is determined to be anomalous. Resultant actions are;     - the query output is quarantined     - event data is sent to the SIEM with sufficient information describing the event (who/what/where/when)     - SIEM raises event/incident (as relevant in Bupa) in ServiceNow     - Operations team triages and resolves - SLAs to be confirmed | **For platform need to consider following in relation to unauthorised access relating to RBAC:**   * **PowerBI -** question on access needs to be authorised by Azure AD? Currently no logs received from PowerBI on Sentinel. Need to consider best way forward. Could be Azure AD or if its not Azure AD there would be a specific logging out of the static data base. If it uses Azure AD, then we can leverage exisiting logs into sentinel.   + **Action:** Confirm whether access needs to be authorised by Azure AD? * **Databricks SQL -** User needs to be part of Active directory for them to be authenticated into Databricks. Therefore covered through Azure. Can be AD or AAD and sentinel will covered. * **Service principles logging into ADLS -** Only privileged users can access ADLS and is time bound. Authentication is via Active Directory as they log in via Azure portal. If -priv accounts are being used and they are onboarded into CyberArc, then can use the CyberArc logs.   + **Action**: need to confirm whether we use -a accounts or -priv accounts. * **ADF -** Managed identity used to access various components - can use AAD or AD - sentinel covered.   **For platform need to consider following in relation to abnormal access:**   * Situation is data analyst has normal actions (i.e. pull x amount of data at certain time). Then on a Saturday, exfiltration of data occurs where its abnormal. * **Databricks** can pick up normalised behaviour. If there is abnormal behaviour, it can be sent to SIEM for alerting.   + How much data is being accessed in a query cannot be found in the log. We only know what query was executed - current gap.   + ADLS has functionality to log size of exfiltration but not currently enabled in Dev. Would need to confirm with Data Ops when it can be enabled.   + IBM guardian can monitor queries that are run - includes number of records and size of exfiltration. Need to confirm whether IBM guardian can cover ADF and ADLS.     - **Action:** Check if IBM guardian can cover ADLS and ADF.     - **Action:** Check if ADLS and Databricks can do normalised, abmonalised behaviour and the quarantining behaviour - then from here decide if tis log can be sent to Sentinel - to organize discussion with Dan   + Sentinel can do alerting - if we want quarantine a query then it would need to get all logs sent rather than just an abnormal query. |
| UC-SIEM-2 | Alert | MVP 1 | **Data Breach Alerts:** The solution should be able to generate alerts when there is an unauthorised access or data exfiltration of sensitive data stored in the BDP4 Platform. This may be referred to as ‘outlier analysis’ where the expectation that the solution display events out of the ordinary, this includes but not limited to the following examples:   * Excessive use * Attempted brute force on data sets without provisioned access * Failure reporting * Non-typical activities by privileged users.   + This includes anomaly detection in relation to a users typical data exfiltration exercises. For example, there will be instances where a business team may exfiltrate x rows of data as part of their business function but another business team may only typically exfiltrate x rows. In the instance the latter team exfiltrates x greater amount of rows of data, this should be picked up by Immuta and alert sent to the monitoring team. Refer to example scenario in UC-SIEM 1 | **For platform need to consider following in relation to excessive use:**   * See above   **For platform need to consider following in relation to attempted brute force use:**   * Authorization being blocked by AAD and AD identity protection. Log existing to Sentinel * Gap to consider:   + Anything not AD and AAD authenticated per previous use case?     - Immuta may give us fine grain controls. Logging to be built from Immuta to Sentinel to address this   **For platform need to consider following in relation to failure reporting use case:**   * Per use case above * Rolled up into authentication - Azure vs non azure. If non azure then logging to be pushed into platform.   **For platform need to consider following in relation to non-typical/anomaly detection use case:**   * Timeframe for onboarding:   + 1. Defining logging requirements - in progress now   + 2. Logging reqs defined - then Shannon R team involved to build technical integration - Shannon can answer this   + 3. If there are any bespoke / customer use cases - then we make changes as required via CyberCX - anywhere between 2 - 4 weeks.     - Custom use cases could be:       * Related to Privilaged users if we dont have non-standard -priv accounts via Cyberarc.       * Excessive usage - if OOTB using IBM guardian then all good, if anything else then may be custom. Ideally we would want a tool to monitor this     - Cost to develop use cases: Abdullah to share |
| UC-SIEM-3 | Alert | MVP 2 | **Data masking failure alerts:** The solution should be able to generate alerts when there is a failure in the data masking process such as incorrect masking or masking failure. | **For platform need to consider following in relation to data masking failure:**   * Logging required out of Immuta into Sentinel when Data masking failure has occured. * There may be multiple sub-cases of data masking failures. e.g. data masking failed on x data set, data masking failed on this date etc. * From Cyber perspective, when data masking fails, what is the Cyber risk and does it fail open or closed:   + Closed: Security perspective ok   + Open: Unlikely   + E.g. - data masking fail occurs, does the process stop?   + What type of failures are of Cyber risk?   + If it always fails closed - then there may be a case not required to have logging in sentinel. However, there may need to be an operational consideration here.   + If there is a failure - its a service owners situation to deal with - consider operational considertation.     - Needs to be built into Immuta Design |
| UC-SIEM-4 | Alert | MVP 2 | **Encryption failure Alerts:** The solution should be able to generate alerts when there is a failure in the data encryption process, such as encryption key loss or encryption failure. | **For platform need to consider following in relation to encryption failure:**   * What tool detects encryption failure?   + in platform currently, we are not additionally encrypting data. ADLS by default offers encryption which is being used, on top of this no customized encryption   + Data in transit, we make use of TLS encryption but again nothing ontop of this   + We have no specific encryption keys   + Most systems used here we rely on vendor enabled encryption - e.g. emails rely on Microsoft encryption   + SRA team reviews encryption for systems and identifies gaps to address as required.   + PowerBI?     - Because microsoft service it would use TLS.     - Imported data set is also encrypted by PowerBI.   + NAS drives     - Users would use this to upload files     - Question is do we want to log this as its being encrypted?       * What situations are there when there is a failure of envryption of NAS drives?       * Is encryption going to be related to NAS drives we care about?         + **Action**: to confirm |
| UC-SIEM-5 | Alert | MVP 2 | **Data Policy violation alerts:** The solution should be able to generate alerts when there is a violation of the data security policies, such as unauthorised access to sensitive data or data retention violations. | **For platform need to consider following in relation to data policy violation failure:**   * Data security policies?   + If there is a DLP report that pertains to the data from Platform - who needs to be notified of this?   + There is Azure Information Protection AIP / Microsoft Information Protection MIP - all excel sheets, word docs etc are marked as business use only     - Given PowerBI main platform to access, depends if data classification has been enabled for PowerBI. Its on roadmap but would need to confirm if the workspaces have data classification enabled and the log in sentinel could be received for when there is a change from business use only to confidential and vice versa. AIP/MIP can cover it off if data classification is enabled.       * **Action**:   + AIP / MIP would not cover data retention     - Need to get details of data retention policy. Parked for now. |
| UC-SIEM-6 | Correlate | MVP 3 | **Investigate:** The solution should be used to investigate security incidents and provide a comprehensive view of the incident by analysing data from multiple sources in the data platform. | **For platform need to consider following in relation to investigate:**   * Once we have excessive and access use case covered, then we can understand what is covered by security tooling / sentinel.   + **Action**: to find best practice documentation of logging. Vendors may be able to provide guidance. |
| UC-SIEM-7 | Identify & Alert | MVP 3 | **Additional data exfiltration detection via Azure Databricks:** In addition to data exfiltration detection identified in UC-SIEM-1, additional scenarios should be considered for implementation. Idea would be to leverage Kusto Data Exfilitration script via Log Analytics. The scenarios are outlined below:   * **Scenario 1:** Control should be in place to check that data is not being stored on DBFS to avoid exfiltration via DBFS of data. Alert in place to check where data is being exfiltrated out of DBFS. * **Scenario 2:** Control in place to block saving of data to local drives. Alert should be generated for attempts to download data sets. Users should be able to have dataframes displayed and visualised all via Databricks. * **Scenario 3:** Control in place regarding displaying data multiple times in notebooks. When the display() command is used, only a limited set of data is printed to a browser. | Refer to Kusto Data exfiltration script run via Log Analytics. [azure-monitor-detect-exfiltration-databricks/databricksnotebook.kusto at main · rebremer/azure-monitor-detect-exfiltration-databricks · GitHub](https://github.com/rebremer/azure-monitor-detect-exfiltration-databricks/blob/main/detection/databricksnotebook.kusto) |

* Access Logs: Logs that track user access to sensitive data stored, including login attempts, data access events, and data access failures.
* Audit Logs: Logs that track changed to sensitive data, including the type of masking applied the data masked, and the user responsible for the masking.
* Data Masking Logs: Log that track the masking of sensitive data, including the type of masking applied, the data masked, and the user responsible for the masking.
* Encryption Logs: Logs that track the encryption of sensitive data, including the type of encryption, and the user responsible for the encryption.
* Alert Logs: Logs that track security alerts generated by the DLP system, including alerts for unauthorised access to sensitive data, data breaches, and data masking failures