Implementing a Next Gen CASB for Data Loss Prevention

MSc Research Project
MSc in Cybersecurity

Dharani Muruga Prasad
Student ID: 22208593

School of Computing
National College of Ireland

Supervisor: Prof. Vikas Sahni

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Dharani Muruga Prasad		
Student ID:	22208593		
Programme:	MSc in Cybersecurity	**Year:**	2023-2024
Module:	Practicum		
Supervisor:	Prof. Vikas Sahni		
Submission Due Date:	12/08/2024		
Project Title:	Implementing a Next Gen CASB for Data Loss Prevention		
Word Count:	6491	**Page Count:**	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Dharani Muruga Prasad

Date: 11/08/2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Implementing a Next Gen CASB for Data Loss Prevention

Dharani Muruga Prasad
22208593

Abstract

There is hardly any organization that can handle a data breach, which is why Data Loss Prevention (DLP) has become a crucial concern. Data breaches have gained global attention as organizations become more dependent on digital data, cloud services, and remote connectivity. Although cloud applications reduce costs and increase productivity, guarding and monitoring data in this environment is equally important. As more users connect directly to cloud services, there is an increased need to secure data stored in the cloud to meet compliance, privacy, and security requirements. Traditional DLP solutions largely fail to secure cloud data due to a lack of visibility and control over cloud activities which makes it essential for organizations to implement a Cloud Access Security Broker (CASB). This study explores the benefits and challenges of implementing DLP in CASB, emphasizing the importance of a unified approach to data security in the cloud era. It further assesses whether Defender for Cloud Apps is effective in protecting organizations from data threats in the cloud.

1 Introduction

Organizations are shifting to cloud through various forms like SaaS, IaaS, and PaaS at a rapid rate. By 2025, 85% of organizations will be “cloud first”, and the cloud market is projected to grow up to 1.6 trillion US Dollars by 2030 (Jayaraman Soundarya, 2024).

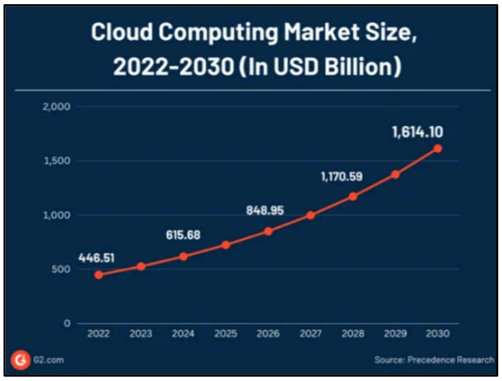


Figure 1: Cloud Computing Market

A lot of users and devices are categorized into IoT, remote, shadow IT, user-managed devices (BYOD). In such diverse and changing environments, legacy security solutions are not efficient. It is highly necessary to monitor and control the usage of cloud applications to ensure enterprise security. Instead of entirely deleting cloud services and potentially affecting the organization’s productivity, companies can adopt a new solution to mitigate this gap and secure their infrastructure. Cloud Access Security Broker (CASB) emerges as an advanced security technology customized to tackle threats in the cloud.

"Cloud access security brokers (CASBs) are on-premises, or cloud-based security policy enforcement points, placed between cloud service consumers and cloud service providers to combine and interject enterprise security policies as the cloud-based resources are accessed" is the definition by Gartner. (Gartner, n.d.)[1] A CASB whether hosted in the cloud or implemented on-premises, acts as an intermediary between cloud service providers and users, enabling organizations to securely access the cloud while protecting sensitive data. It provides an oversight and manages risks and data through auto-discovery to find all third-party cloud applications, assess the risk level associated with each of them and monitor their usage. This data is then used to create and enforce policies. By implementing a CASB, organizations can improve data protection, compliance, threat detection and visibility across all their cloud services. A CASB solution can be deployed on-premises or in the cloud using Forward Proxy, Reverse Proxy, or API-Control configurations.

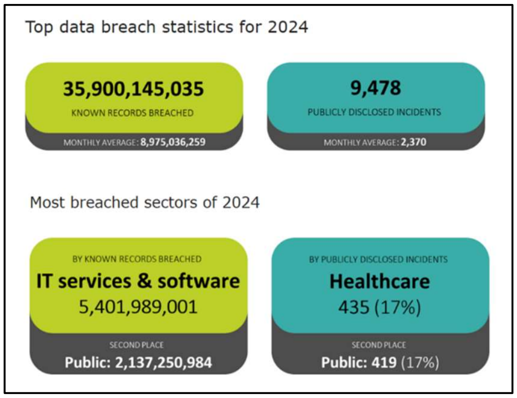


Figure 2: Data Breaches (IT Governance, 2024)

One major issue these days is that businesses lose confidential data. The risk has grown much larger as more and more companies store their data in the cloud. This data can be lost in many ways; when an employee makes a mistake, data gets deleted by accident, or hackers attack. Data loss can cost a business significantly and affect the company's reputation. These risks make it extremely important for companies to be aware of and try everything possible to prevent the loss of their data when using cloud services. The companies must take strong measures to protect their information. This includes applying proper security methods, checking systems often, and training workers on how to keep data safe. Besides this, companies should always have a backup plan in case they lose data so that they can get back on their feet as fast as possible in the event of a data breach.

Data loss prevention (DLP) is a comprehensive approach aimed to protect a company's valuable data from potential threats from either internal or external sources. This approach makes use of several services and procedures that collaborate to identify and protect an organization's data assets in accordance with a company-established data handling policy. Data handling policies are created by organizations specifically for the kinds of data they handle and retain. These regulations, which may include adhering to PCI-DSS or HIPAA requirements, must be in accordance with the business's operational demands. DLP solutions identify policy

[1] Gartner - https://www.gartner.com/en/information-technology/glossary/cloud-access-security-brokers-casbs

breaches, address and resolve the issue by triggering alerts, notifying users and adhering to compliance by implementing safety measures such as clearing the clipboard after a user copies unauthorized data.

One of the key benefits of implementing a CASB is its enhanced visibility and centralised management of data in the cloud. With the help of a CASB, organizations can monitor users' activities, prevent unauthorized access to sensitive data and enforce data loss prevention policies. For example, a CASB can regulate the sharing of private information externally or block uploading specific file types to the cloud. It provides security against different data threats by proactively identifying and remediating any suspicious activity. CASBs can detect unauthorized access, malicious file extensions and other cyber threats that might compromise the data. Additionally, administrators can be notified about any unusual activity or policy violations, ensuring timely to action avoid potential damage. The key element of any organization's data loss prevention strategy for the cloud is the utilization of a CASB solution. By ensuring regulatory compliance, threat protection, data control and visibility, and compatibility with other security tools, CASBs play an important role in helping businesses in protecting their data in the cloud and preventing financial damage of data breaches.

Organizations are forced to put strict measures in place to protect their sensitive data from unauthorized disclosure due to the rising number of data breaches. CASBs offer a highly effective solution to protect data from unwanted access. A CASB solution essentially acts as an intermediary between businesses and cloud providers. Organizations can lower the risk of data loss by improving their monitoring and management of cloud apps with the use of a CASB solution.

1.1 Research Question:
How can the implementation of Data Loss Prevention (DLP) in Cloud Access Security Broker (CASB) platforms enhance data protection and security in cloud environments?

1.2 Structure of the report:
This study describes a DLP solution implemented in a Next Gen CASB to secure data on the cloud. Section 2 comprises detailed research in this field including subsections for CASB, DLP, Cloud security, related work that led to this project and justification for the choice of platform. The research methodology used is detailed in Section 3. Section 4 covers the design specifications and Section 5 details the implementation. Evaluation of the results is explained in detail in Section 6. Section 7 covers the conclusion and addresses future work.

2 Related Work

2.1 CASB

The main contributions of (Javed et al., 2024) are creating a way to ensure log integrity using blockchain, building a prototype, and providing a security analysis. However, the approach faces issues like performance slowdowns, scalability problems, integration difficulties, privacy concerns, regulatory compliance challenges, and limited use beyond remote health monitoring.

Main contributions of (Chimpiri, 2024) include real-time monitoring across cloud environments, granular policy enforcement, data loss prevention, compliance monitoring, and multi-cloud integration. The paper highlights CASB's role in enhancing visibility, control, and security consistency in multi-cloud setups. However, limitations of the study include lack of performance metrics, absence of comparisons with other CASB solutions.

A key contribution (Ramesha and Sahni, 2023) is an XDR (Extended Detection and Response) framework enhanced by machine learning, which can collaborate with CASB to supervise and identify threats dynamically. The research utilizes Principal Component Analysis (PCA) as an unsupervised learning method to minimize data loss and reveal patterns in user engagements with cloud resources.

(Reddy et al., 2023) explains major features, including user activity monitoring, access control, anomaly detection of UEBA and machine learning. Precautionary measures, including multi-factor authentication, are also a part of this research. False positives and the difficulty of moving against the evolving threat are some of the limitations. The paper concludes that CASBs secure the cloud but need continuous development to remain effective.

The authors in (Bs, Arvindhan and Kalimuthu, 2023) have proposed that CASBs are essential in safeguarding sensitive data against unauthorized access and ensuring compliance with regulatory mandates. However, the research is missing a detailed assessment of the efficiency of CASBs in boosting security within cloud computing infrastructures, with limited research data to validate the statements made by the authors.

In (Kharb and Chahal, 2023), authors discuss the advantages of integrating CASBs like improved cloud security, more visibility and better management of cloud operations, enforcement of compliance, and safeguarding against internal threats. They also argue that enterprises must consider obstacles like performance impact, issues related to data confidentiality, and the intricacy of integration during CASB deployment.

(Collier, 2023) highlights the contribution of CASBs in increasing security through visibility, compliance, and data protection in the cloud. It emphasizes the importance of understanding organizational needs and the regulatory landscape when selecting a CASB However, drawbacks are that the integration of such solutions with the existing systems of the organizations may be complex and that their effectiveness may differ; therefore, it may be challenging to achieve the maximum security.

(Bhattacharya et al., 2021) propose a dynamic CASB model, which integrated machine learning techniques for real-time monitoring and analysis of user behaviour for detecting security threats and anomalies. Key contributions include an adaptive framework to cut down time to respond to threats and low false positives in security alerts. On the other hand, heavy dependency on huge data sets for training the AI models and probable challenges in customizing the system to diverse cloud environments are some of the limitations of this study.

Key contributions of (Brouwer and Groenewegen, 2021) include proposing a definition of CASB, exploring the relationship between CASB and Shadow IT, analysing how leading CASB vendors address Shadow IT, and investigating ways to mitigate administrative overhead in CASB implementations. Limitations of the study include: a focus on the Dutch market which may not be applicable globally, reliance on a limited number of trial environments from vendors, and potential bias in the semi-structured interview process.

The main contribution of (Obregon, 2021) is a thorough analysis of the design and deployment methods of CASBs as well as important components including threat prevention, DLP, and access control. The significance of CASBs in resolving security issues in cloud environments, particularly those related to data protection and compliance are emphasized in the white paper. However, there are many drawbacks in the study, such as the absence of research data from actual CASB deployments, the lack of a comparative analysis between various CASB manufacturers, and the incomplete examination of potential challenges associated with CASB implementation.

The main contribution of (Ahmad, Mehfuz and Beg, 2021) is the development of an XDR model that provides dynamic monitoring of user behaviour through unsupervised learning techniques, including the PCA algorithm, to look for anomalies and reduce threats. This paper, however, has recognized some of the limitations, such as the challenges to real-time implementation of machine learning and the need for constant data inputs to maintain effectiveness against security threats that are continuously changing.

(Ahmad, Mehfuz and Beg, 2020) emphasizes that CASB is critical for ensuring compliance and visibility over cloud applications used by remote workers. However, this short review has some drawbacks, it misses technical depth or details about specific case studies of the implementation of CASB. It does not investigate some of the problems that might be faced in deploying CASB solutions for remote work settings or compare CASB vendors.

(Khan, no date) highlights the significance of upholding a strong security stance in cloud environments to safeguard sensitive data, guarantee regulatory adherence, and maintain customer confidence is underscored. The paper also delves into the functionalities and advantages of CASBs, alongside their constraints and hurdles. Challenges discussed include the integration with diverse cloud services, performance impact, and intricate policy administration.

This paper (Choudhary et al., 2022) analyses the evolution and compares different CASBs in the market. It contributes by reviewing various CASB solutions, tracing their development, and comparing their features and capabilities. The study offers insights into CASB adaptation to emerging security challenges and provides a framework for evaluating CASB solutions. Limitations include the rapid pace of change in cloud security, potentially outdating some comparisons, and possible constraints in the selection of CASBs analysed.

The white paper (Tayouri et al., 2022) speaks about features like DLP and access control, it also mentions the challenges like possible performance degradation and complications with integration. The basic emphasis of (Yiliyaer and Kim, 2022) is that CASBs are needed as part of a suite of security tools aided with Software-Defined WAN and endpoint security to extend protection end to end. Both papers agree on the point that careful implementation to maximize the benefits of CASBs on one side, and to alleviate the existing limitations such as complexity and potential performance issues, is crucially important.

The major contributions of (Selvam, 2022) are current issues in cloud security, a model to integrate CASB, and implementation considerations. However, the paper has some limitations - it lacks empirical data or case studies to support its claims and doesn't thoroughly address potential drawbacks or challenges of the CASB approach. Additionally, more details on the technical implementation and performance impacts would have strengthened the paper.

(Ahmad et al., 2022) contributes by identifying key trends and gaps in CASB literature, focusing on implementation, security features, and cloud integration. The paper highlights CASBs' evolving role in data protection, access control, and compliance across cloud platforms. Limitations include the rapidly changing nature of cloud security, potential gaps in existing research, and possible selection bias in the review process. The study offers a comprehensive overview of CASB research, providing insights for both academic and practical applications in cloud security.

2.2 Cloud Security

(Bhansali, 2023) and (Kumar and Gupta, 2023) highlight various issues of cloud security that CASBs will protect against data breaches, unauthorized access, encryption, compliance, and threat detection. Lack of visibility and control in cloud environments is the biggest issue CASBs' address. Issues to be targeted include multi-cloud environments, data protection, and prospective security trends like AI and blockchain corresponding to functionalities and enhancements. In general, CASBs add a lot of value to the cloud security controls that has been talked about in these papers, through centralized visibility, control, and policy enforcement across multiple cloud services.

2.3 DLP

Different DLP strategies are strongly recommended by the author of (Kodurupati, 2022) like strict access control, monitoring the movement of data in real time, and educating employees. Conversely, (Shahzad and Machado de Sousa, 2021) sheds light on the challenges encountered by CSPs in protecting customer data from insider threats. The importance is reinforced with CSPs putting into place strong DLP measures, non-disclosure agreements and service-level agreements to ensure data privacy. Both papers focus on the multi-faceted nature of DLP, which clearly indicates that there is a need for solutions to be customized to vulnerabilities within an organization and specific cloud environments.

2.4 Research Niche

(Deloitte, 2021) highlights Deloitte's integrated approach to data discovery, classification, DLP/CASB, and data obfuscation. The document emphasizes their capabilities in addressing regulatory requirements, emerging threats, and business challenges through strategy, governance, process, and technology implementation. Limitations include resource-intensive implementation, integration challenges with existing systems, and the need for continuous updates to address evolving threats and regulations.

Key contributions of (Wason, Aghili and Zavarsky, 2020) include a comprehensive framework combining CASB functionalities and deployment strategies. The model addresses visibility, data protection, threat prevention, and compliance across multiple cloud services. Limitations include lack of real-world validation, possible oversimplification of complex environments, and insufficient consideration of implementation challenges or costs. It does not adequately address adaptability to evolving technologies or compare the model with other CASB approaches.

Key contributions of (Kaur and Gupta, 2019) include increased visibility into cloud application usage, effective threat protection, and access control, all of which work together to improve DLP policies. The paper has some drawbacks, including the difficulty of integrating CASBs with the current IT infrastructure, the possibility of false positives in threat detection, and the requirement for user training to ensure compliance with security guidelines.

2.5 Microsoft Defender for Cloud Apps

Microsoft Defender for Cloud Apps has advanced from a CASB to a comprehensive SaaS security platform. Customers experience new problems in app protection, and as new attack routes emerge in the kill chain, they require current methods to defend their SaaS apps. Defender for Cloud Apps combines key CASB concepts with new SaaS app-protection technologies, ensuring customers have complete app coverage. It adds extra capabilities that go beyond the boundaries of a typical cloud access security broker (CASB) to improve app security and protect against malicious cloud apps. Provides comprehensive app coverage by combining SaaS security posture management, data loss prevention, app-to-app protection, and integrated threat protection.

3 Research Methodology

Defender for Cloud Apps detects sensitive information, helps organizations protect it with data loss prevention (DLP) features, and helps in responding to sensitivity labels on content that is identified. The integration of Defender for Cloud with Microsoft Purview facilitates security teams in leveraging out-of-the-box data classification types within their information security rules. No matter where the information is accessed from, Microsoft offers a wide range of data loss prevention capabilities to make sure it is protected. To find sensitive data files and determine where and by whom it is being accessed, it connects with SaaS apps. The following controls can be put in place by organizations to protect this data:

- Put a sensitivity label on it
- Prevent downloads to unmanaged devices
- Remove external contributors on sensitive files.

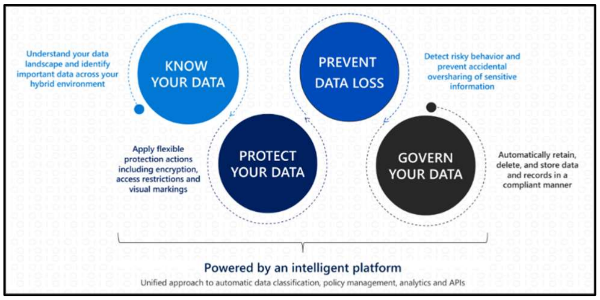

Understand your data landscape and identify important data across your hybrid environment
KNOW YOUR DATA
PREVENT DATA LOSS
Detect risky behavior and prevent accidental oversharing of sensitive information
Apply flexible protection actions including encryption, access restrictions and visual markings
PROTECT YOUR DATA
GOVERN YOUR DATA
Automatically retain, delete, and store data and records in a compliant manner
Powered by an intelligent platform
Unified approach to automatic data classification, policy management, analytics and APIs

Figure 3: Microsoft Purview (Carol Bailey et al., 2024)

Phase 1: Data Discovery

Connecting cloud apps to Defender for Cloud Apps is the first step towards discovering which data is being used in organizations. Once linked, Defender for Cloud Apps can scan data, apply classifications, and enforce policies and controls. The way apps are connected determines how and when scans and controls are implemented. One of the following methods can be used to connect apps:

- App connector: Microsoft's app connectors rely on APIs provided by app vendors. They give more visibility and control over the apps utilized in the organization. Scans are performed on a regular basis (every 12 hours) and in real time (when a change is identified).
- Conditional Access App Control: Apply controls to any app with Microsoft's Conditional Access app control solution, which leverages a reverse proxy architecture that is connected to Microsoft Entra Conditional Access.

Defender for Cloud Apps checks every file a program uses after it is connected to it via API connection. To get a summary of the files shared by the cloud apps, their accessibility, and their current state, access Files under Cloud Apps in the Microsoft Defender Portal.

Phase 2: Classification of sensitive information

1. Organizations must first determine what is classified as sensitive information for them before scanning through files for it. Microsoft provides over 100 pre-defined sensitive information types as part of its data classification service, or businesses can design their own to meet their specific policies. The same sensitive types and labels are accessible across Defender for Cloud Apps and Microsoft Purview Information Protection because to their native integration. Therefore, companies who wish to specify sensitive information can do so by visiting the Microsoft Purview Information Protection portal. Once developed, the information will be accessible in Defender for Cloud Apps. Advanced categorization kinds like fingerprint and exact data match (EDM) are also available.
2. Information protection is enabled
3. Once the types of information to be protected are identified, different policies can be created to monitor the data

Phase 3: Protect the data

The actual need is to protect sensitive data from possible threats now that it can be identified in files. When an event is discovered, it can be manually resolved, or the files can be secured by using one of Defender for Cloud Apps' automatic governance steps. Among actions are real-time monitoring, actions provided by APIs, and native controls for Microsoft Purview Information Protection, among others. The type of policy being configured determines the type of governance that is used, like:

1. File policy governance actions: Uses Microsoft's native integrations and the API of the cloud app provider to secure files
2. Session policy controls: Employs reverse proxy features to secure files

Phase 4: Monitoring and Reporting

All the necessary policies are in place to identify and secure the data. The dashboard can be monitored daily to check if there are any new alerts. It's a good location to track the state of organization's cloud environment. The dashboard assists in understanding the events and, if required, start an investigation.(Batami Gold et al., 2024)

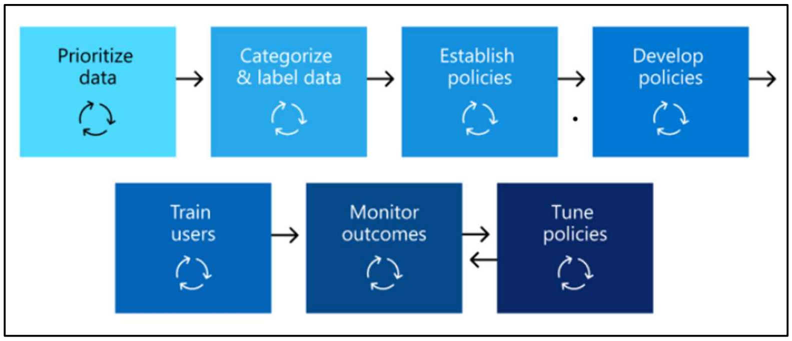


Figure 4: Methodology (Chris Fox MSFT et al., 2024)

Defender for Cloud Apps operates by using a reverse proxy to connect the user and the cloud application securely, enabling DLP and session controls in real time on the cloud app. Since all data passes over the proxied connection, it has access to all data that is sent between the cloud app and the user device. The change in the domain is evidence of this.

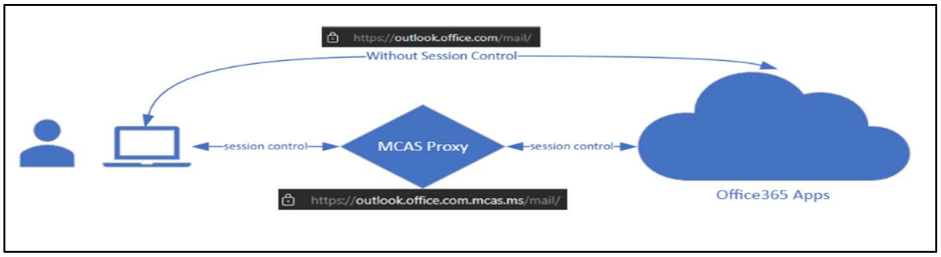


Figure 5: Data flows through the MDCA proxy between the device and the cloud app (Arjun Ramakrishnan, 2023)

4 Design Specification

Organizations require a mechanism that helps in preventing their users from incorrectly sharing sensitive data with people who shouldn't have it to protect this sensitive data and reduce the risk from oversharing. This approach is called data loss prevention (DLP). By creating and implementing DLP policies, businesses can prevent data loss in Microsoft Purview. A DLP policy allows companies to detect, monitor, and automatically secure confidential data across all applications.

DLP policies allow organizations to monitor user behaviour when it comes to sensitive data in use, transit, or rest and then take appropriate action. For instance, when a user tries to perform an unauthorized action, such as copying sensitive data to a forbidden location or sending medical information in an email, DLP can:

- Display a pop-up policy tip alerting user the possibility that they might be trying to inappropriately share classified data.
- Block the sharing and use a policy tip to allow overriding of block by user to record user justification.
- For data at rest - block the sharing without override option
- Lock and transfer sensitive items to a secure quarantine location.

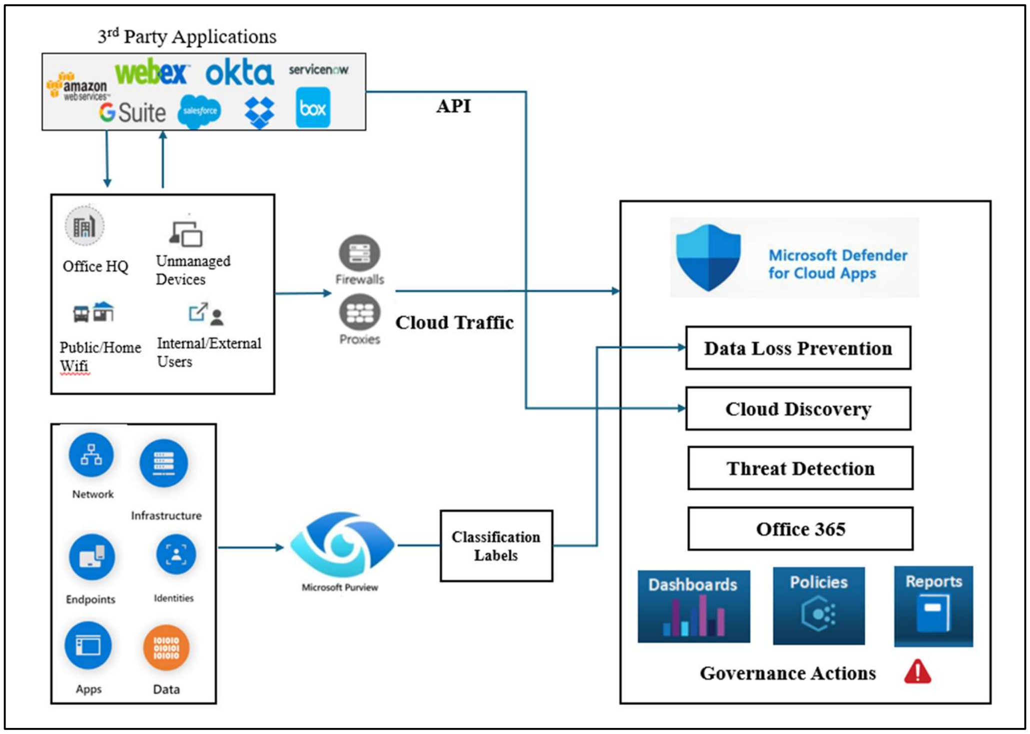


Figure 8: Architecture Diagram

4.1 Tools Used and Pre-requisites

Table 1: Tools and licenses

Tool	License
Microsoft Purview	Microsoft 365 E5 Compliance
Microsoft Defender	Microsoft Defender for Office 365 (Plan 2)
Office 365	Office 365 E5, Exchange Online Protection

5 Implementation

Organizations can deploy Microsoft Purview Information Protection's sensitivity labels automatically with Microsoft Defender for Cloud Apps. Depending on how the label is configured, these labels can apply encryption for further security when applied to files as a file policy governance action. Investigating files can also be done through the Defender for Cloud Apps site by filtering for the applied sensitivity label. Organizations can have more control and visibility over your sensitive data in the cloud by using labels. It only takes enabling one checkbox to integrate Microsoft Purview Information Protection with Defender for Cloud Apps. Organizations may utilize the full potential of both services and secure files in your cloud by integrating Microsoft Purview Information Protection with Defender for Cloud Apps.(Chris Fox MSFT et al., 2024)

5.1 DLP Planning

1. DLP policies can be adopted across all departments any organization once they are implemented. Businesses should appoint stakeholders who can:
    - Outline the laws, regulations, and industry standards the organization must comply with.
    - Determine which categories of sensitive assets need protection.
    - Evaluate which business processes they are utilized in, and which risky activity needs to be restricted.
    - Based on the risk and sensitivity, prioritize which data should be protected primarily
    - Detail the process for reviewing and remediating DLP policy match events.
2. Sensitivity labels and policies to protect the data of organizations are defined: These labels can apply protection actions including content markings (headers, watermarks, footers,), encryption, and other access controls in addition to indicating the sensitivity of the content.
3. Data from all the Microsoft 365 apps and services are labelled and protected: PowerPoint, Teams meetings, Outlook, Microsoft 365 Word, Excel, and containers such as SharePoint and OneDrive sites and Microsoft 365 groups are all compatible with sensitivity labels. Implement a combination of labelling techniques, including mandatory labelling, default labelling, automatic labelling, and manual labelling.
4. Sensitive data in the cloud and on-premises are discovered, labelled, and protected by implementing information protection scanner along with sensitivity labels.

5.2 DLP Policy Configuration

1. Define information to be monitored: There are pre-defined policies in Purview, custom policies can also be created to suit the requirements of organizations.
2. Administrative scoping for policy: Only users, groups, distribution groups, and accounts that they are assigned to can have policies created and managed by administrators attached to an administrative unit. Therefore, an administrator can scope policies to administrative units or apply them to all users and groups.

3. Set locations to be monitored: One or more locations like SharePoint sites, OneDrive accounts, Exchange email to be monitored by DLP are selected with options for distribution lists, groups, etc.
4. Select the conditions that need to be matched for a policy to be applied to an item: Custom conditions can be defined, or predefined ones can be modified accordingly. For example, a document containing confidential data of the organization is shared externally.
5. Specify an action to be executed when the policy requirements are satisfied: Actions are determined based on the activity's location. Examples:
    - Exchange/OneDrive/SharePoint: Block external users from accessing the content. Notify the user via email that they are attempting a prohibited action which is against the DLP policy and display a policy tip.
    - Office 365 Apps: Display a popup to the user alerting them to the risky actions they engage in, then block or block but permit override.

5.3 DLP Policy Deployment

1. Execute the policy in simulation mode excluding policy tips, then use incident and DLP reports to evaluate the impact. Policies can be fine-tuned based on the results. DLP policies won't affect the organization's productivity when they are in simulation mode.
2. Run it with policy tips and notifications so that users can be trained about compliance policies. A link to the organization's policy website with additional information regarding the policy is helpful when it comes to the Policy Tip.
3. Initiate full policy adoption to ensure that the rules are followed, and data is secured. To ensure that the outcomes are what you intended, keep an eye on the DLP reports as well as any incident reports or notifications.
4. User Training: Policies can be configured to automatically send email notifications and display policy tips to administrators and end users when a DLP policy is triggered. Policy tips are helpful in raising awareness about risky activities and educate users to avoid them.
5. Review DLP specifications and update the strategy: Organization's legal requirements, standards, and regulations will evolve over time, just as the business objectives for DLP will. Ensure the DLP implementation continues to suit business needs, and that the organization maintains compliance by including regular assessments of all these areas.

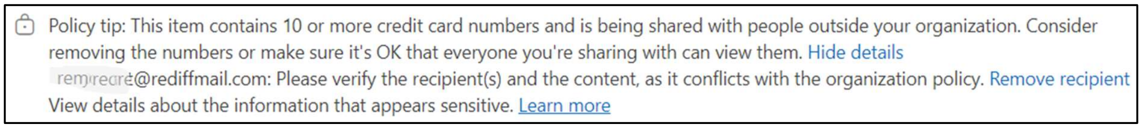
Alerts triggered can be monitored on the ‘Alerts’ tab of the Data Loss Prevention solution in the Microsoft Purview portal. DLP incidents/alerts can also be triaged in the Microsoft Defender portal if connected.

6 Evaluation

This research aims to classify different types of sensitive information and protect confidential data to prevent data loss. Sample data from DLPTest[2] containing credit card numbers, Social

[2] https://dlptest.com/

Security numbers (SSNs), and other personally identifiable information (PII) was shared over an email and to observe how effectively Defender for Cloud Apps could detect and manage the data. A policy tip was displayed when the file was attached, the message was blocked, a detailed email with user, receiver, timestamp, location and file details was sent to the admin and an alert was generated in the Defender portal.

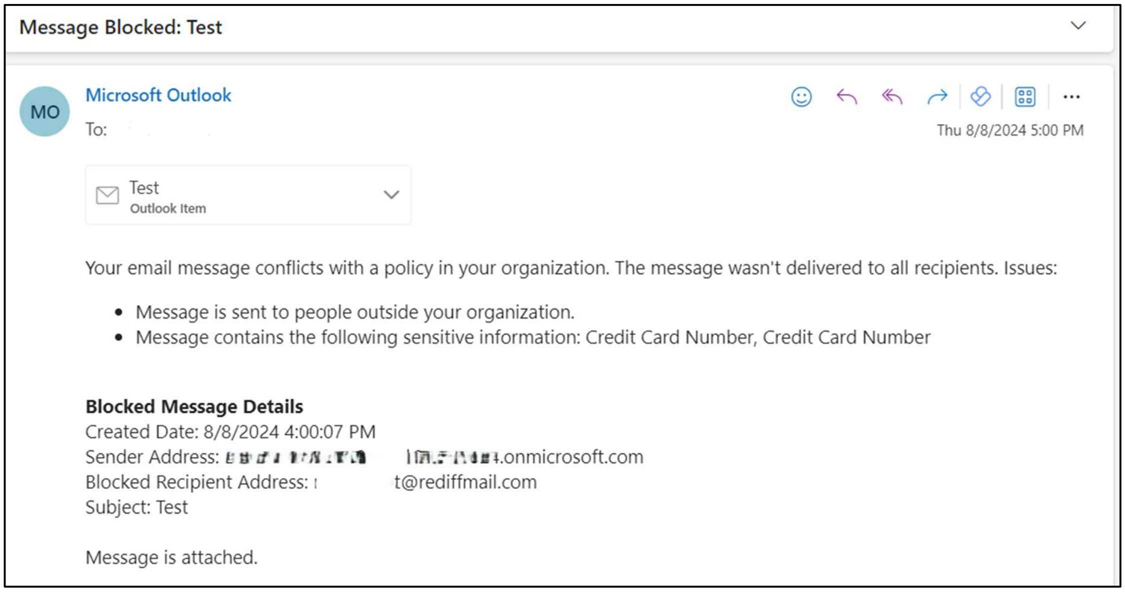


Policy tip: This item contains 10 or more credit card numbers and is being shared with people outside your organization. Consider removing the numbers or make sure it's OK that everyone you're sharing with can view them. Hide details

[redacted]@rediffmail.com: Please verify the recipient(s) and the content, as it conflicts with the organization policy. Remove recipient

View details about the information that appears sensitive. Learn more

Figure 9: Policy tip



Message Blocked: Test

Microsoft Outlook

To:

Thu 8/8/2024 5:00 PM

Your email message conflicts with a policy in your organization. The message wasn't delivered to all recipients. Issues:

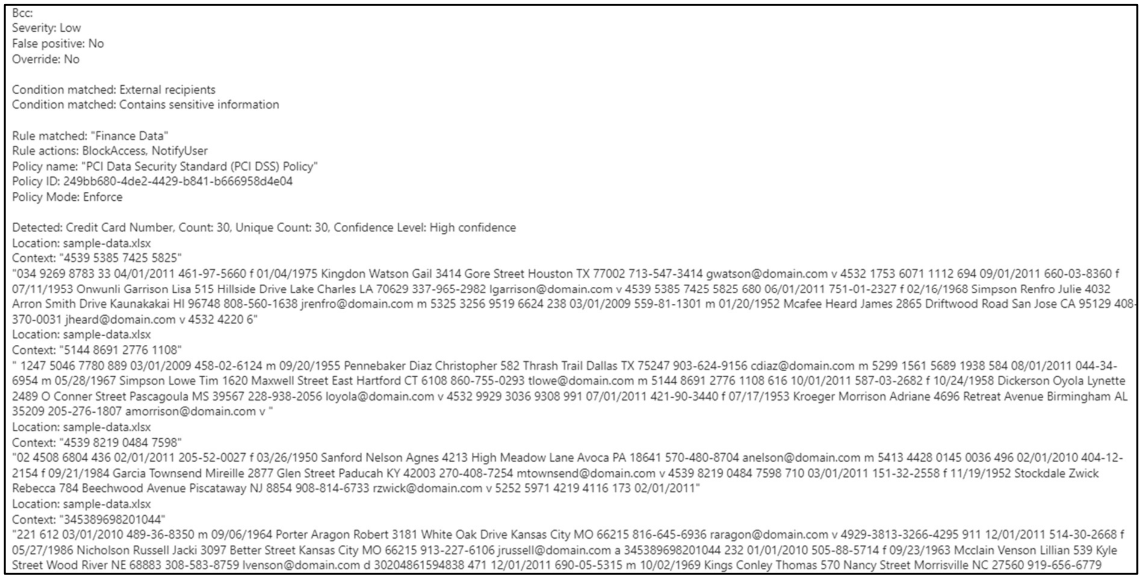
- Message is sent to people outside your organization.
- Message contains the following sensitive information: Credit Card Number, Credit Card Number

Blocked Message Details

Created Date: 8/8/2024 4:00:07 PM
Sender Address: [redacted].onmicrosoft.com
Blocked Recipient Address: t@rediffmail.com
Subject: Test

Message is attached.

Figure 10: Blocked notification



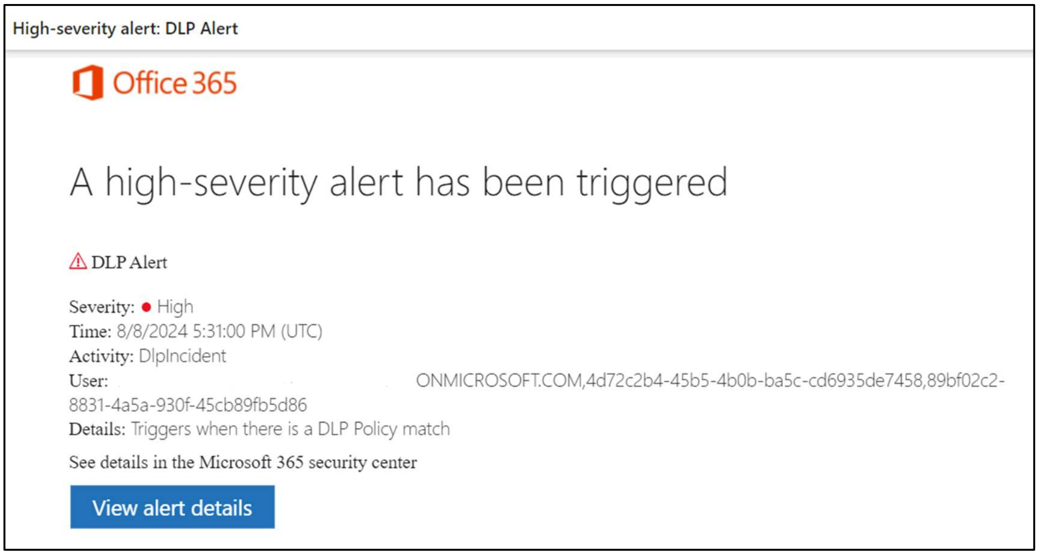
```
Bcc:
Severity: Low
False positive: No
Override: No

Condition matched: External recipients
Condition matched: Contains sensitive information

Rule matched: "Finance Data"
Rule actions: BlockAccess, NotifyUser
Policy name: "PCI Data Security Standard (PCI DSS) Policy"
Policy ID: 249bb680-4de2-4429-b841-b666958d4e04
Policy Mode: Enforce

Detected: Credit Card Number, Count: 30, Unique Count: 30, Confidence Level: High confidence
Location: sample-data.xlsx
Context: "4539 5385 7425 5825"
"034 9269 8783 33 04/01/2011 461-97-5660 f 01/04/1975 Kingdon Watson Gail 3414 Gore Street Houston TX 77002 713-547-3414 gwatson@domain.com v 4532 1753 6071 1112 694 09/01/2011 660-03-8360 f
07/11/1953 Onwunli Garrison Lisa 515 Hillside Drive Lake Charles LA 70629 337-965-2982 lgarrison@domain.com v 4539 5385 7425 5825 680 06/01/2011 751-01-2327 f 02/16/1968 Simpson Renfro Julie 4032
Arron Smith Drive Kaunakakai HI 96748 808-560-1638 jrenfro@domain.com m 5325 3256 9519 6624 238 03/01/2009 559-81-1301 m 01/20/1952 Mcafee Heard James 2865 Driftwood Road San Jose CA 95129 408-
370-0031 jheard@domain.com v 4532 4220 6"
Location: sample-data.xlsx
Context: "5144 8691 2776 1108"
" 1247 5046 7780 889 03/01/2009 458-02-6124 m 09/20/1955 Pennebaker Diaz Christopher 582 Thrash Trail Dallas TX 75247 903-624-9156 cdiaz@domain.com m 5299 1561 5689 1938 584 08/01/2011 044-34-
6954 m 05/28/1967 Simpson Lowe Tim 1620 Maxwell Street East Hartford CT 6108 860-755-0293 tlowe@domain.com m 5144 8691 2776 1108 616 10/01/2011 587-03-2682 f 10/24/1958 Dickerson Oyola Lynette
2489 O Conner Street Pascagoula MS 39567 228-938-2056 loyola@domain.com v 4532 9929 3036 9308 991 07/01/2011 421-90-3440 f 07/17/1953 Kroeger Morrison Adriane 4696 Retreat Avenue Birmingham AL
35209 205-276-1807 amorrison@domain.com v "
Location: sample-data.xlsx
Context: "4539 8219 0484 7598"
"02 4508 6804 436 02/01/2011 205-52-0027 f 03/26/1950 Sanford Nelson Agnes 4213 High Meadow Lane Avoca PA 18641 570-480-8704 anelson@domain.com m 5413 4428 0145 0036 496 02/01/2010 404-12-
2154 f 09/21/1984 Garcia Townsend Mireille 2877 Glen Street Paducah KY 42003 270-408-7254 mtownsend@domain.com v 4539 8219 0484 7598 710 03/01/2011 151-32-2558 f 11/19/1952 Stockdale Zwick
Rebecca 784 Beechwood Avenue Piscataway NJ 8854 908-814-6733 rzwick@domain.com v 5252 5971 4219 4116 173 02/01/2011"
Location: sample-data.xlsx
Context: "345389698201044"
"221 612 03/01/2010 489-36-8350 m 09/06/1964 Porter Aragon Robert 3181 White Oak Drive Kansas City MO 66215 816-645-6936 raragon@domain.com v 4929-3813-3266-4295 911 12/01/2011 514-30-2668 f
05/27/1986 Nicholson Russell Jacki 3097 Better Street Kansas City MO 66215 913-227-6106 jrussell@domain.com a 345389698201044 232 01/01/2010 505-88-5714 f 09/23/1963 Mcclain Venson Lillian 539 Kyle
Street Wood River NE 68883 308-583-8759 lvenson@domain.com d 30204861594838 471 12/01/2011 690-05-5315 m 10/02/1969 Kings Conley Thomas 570 Nancy Street Morrisville NC 27560 919-656-6779
```

Figure 11: Alert details



High-severity alert: DLP Alert

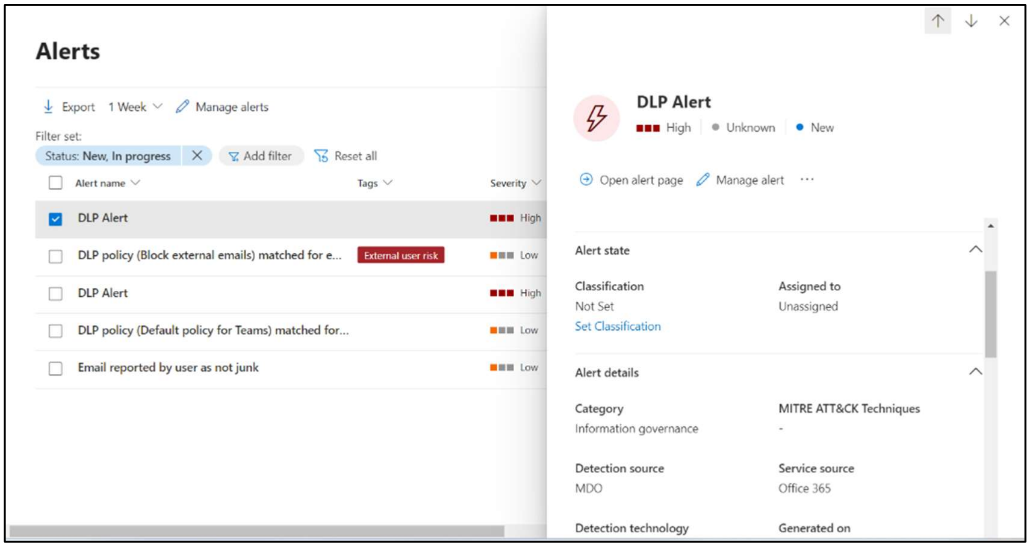
A high-severity alert has been triggered

⚠ DLP Alert

Severity: ● High
Time: 8/8/2024 5:31:00 PM (UTC)
Activity: DlpIncident
User: ONMICROSOFT.COM,4d72c2b4-45b5-4b0b-ba5c-cd6935de7458,89bf02c2-8831-4a5a-930f-45cb89fb5d86
Details: Triggers when there is a DLP Policy match

See details in the Microsoft 365 security center

Figure 12: Defender alert notification



Alerts

	Alert name	Tags	Severity
☑	DLP Alert		■■■ High
□	DLP policy (Block external emails) matched for e...	External user risk	■■■ Low
□	DLP Alert		■■■ High
□	DLP policy (Default policy for Teams) matched for...		■■■ Low
□	Email reported by user as not junk		■■■ Low

DLP Alert
■■■ High | ● Unknown | ● New

Alert state

Classification: Not Set
Set Classification
Assigned to: Unassigned

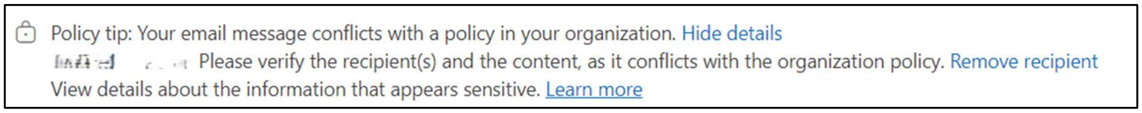
Alert details

Category: Information governance
MITRE ATT&CK Techniques: -
Detection source: MDO
Service source: Office 365
Detection technology
Generated on

Figure 13: Microsoft Defender Alerts Dashboard

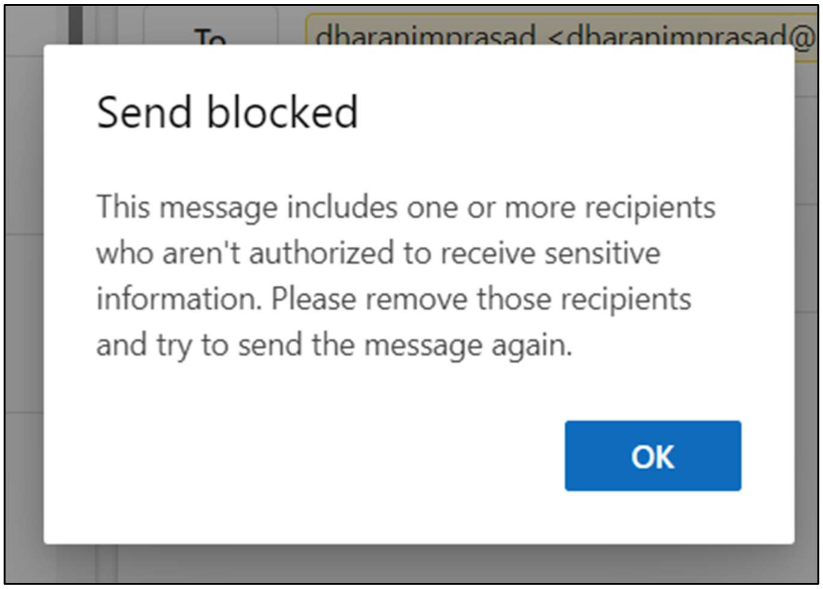
6.1 Use case 1 - Block external emails

A policy was created to restrict users from sending emails to their personal addresses to prevent sharing of internal data.



Policy tip: Your email message conflicts with a policy in your organization. Hide details
[illegible] Please verify the recipient(s) and the content, as it conflicts with the organization policy. Remove recipient
View details about the information that appears sensitive. Learn more

Figure 14: Policy tip



To dharanimprasad <dharanimprasad@

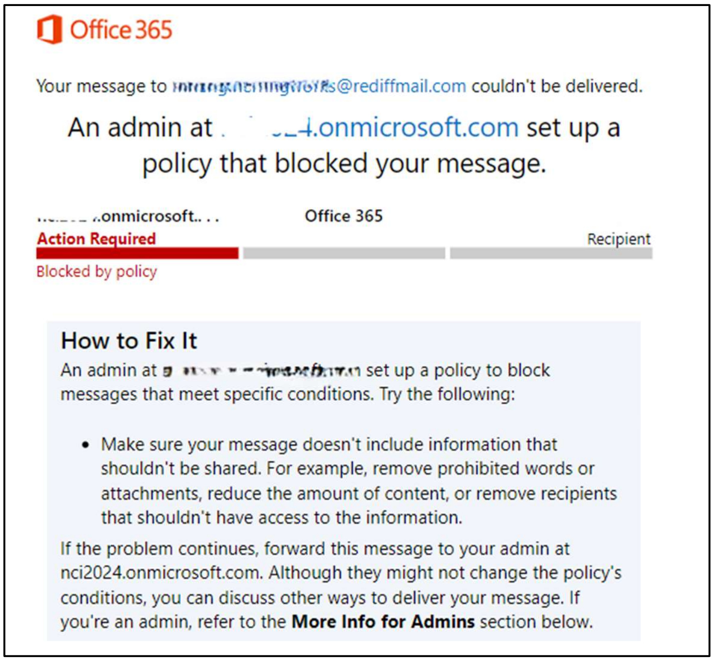
Send blocked

This message includes one or more recipients who aren't authorized to receive sensitive information. Please remove those recipients and try to send the message again.

Figure 15: Blocked message

6.2 Use case 2 - Block sensitive documents

Policy created to restrict sharing documents containing confidential data related to the organization by checking for sensitive words in the name and content of the document.



Your message to [redacted]@rediffmail.com couldn't be delivered.

An admin at [redacted]4.onmicrosoft.com set up a policy that blocked your message.

[redacted].onmicrosoft...	Office 365	
Action Required		Recipient
Blocked by policy		

How to Fix It

An admin at [redacted] set up a policy to block messages that meet specific conditions. Try the following:

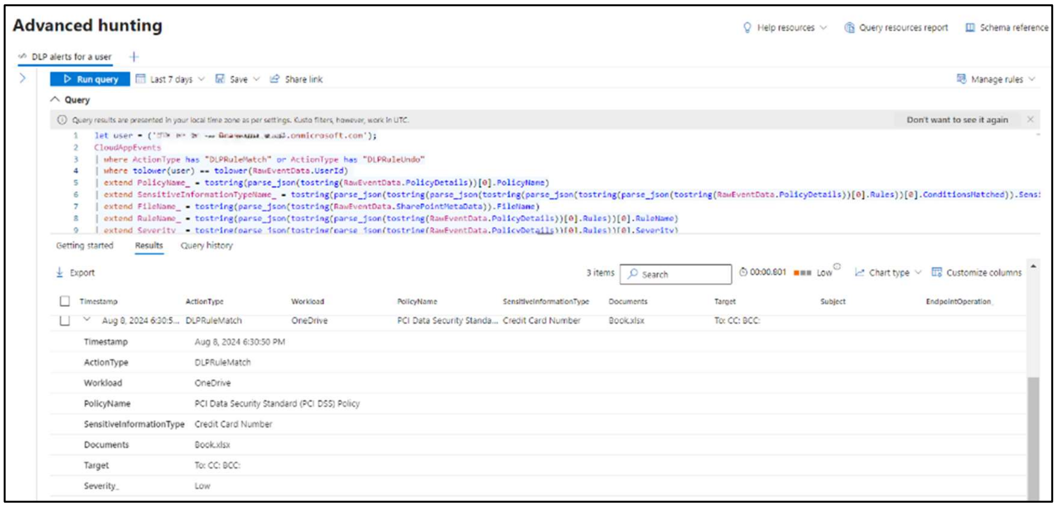
- Make sure your message doesn't include information that shouldn't be shared. For example, remove prohibited words or attachments, reduce the amount of content, or remove recipients that shouldn't have access to the information.

If the problem continues, forward this message to your admin at nci2024.onmicrosoft.com. Although they might not change the policy's conditions, you can discuss other ways to deliver your message. If you're an admin, refer to the **More Info for Admins** section below.

Figure 16: Blocked notification

6.3 Use case 3 - Visualisation of raw data

KQL Queries were built based on existing IOCs in the advanced hunting section of Defender portal to visualize the attack chain and query the raw compliance and security data signals produced by Microsoft 365 to proactively identify known and potential vulnerabilities inside the organization. In addition to improving the investigative workflow, advanced hunting reveals more about the different kinds of signals you receive throughout the environment.



```
let user = ('[redacted].onmicrosoft.com');
CloudAppEvents
| where ActionType has "DLPRuleMatch" or ActionType has "DLPRuleUndo"
| where tolower(user) == tolower(RawEventData.UserId)
| extend PolicyName_ = tostring(parse_json(tostring(RawEventData.PolicyDetails))[0].PolicyName)
| extend SensitiveInformationTypeName_ = tostring(parse_json(tostring(parse_json(tostring(parse_json(tostring(RawEventData.PolicyDetails))[0].Rules))[0].ConditionsMatched)).Sens:
| extend FileName_ = tostring(parse_json(tostring(RawEventData.SharePointMetaData)).FileName)
| extend RuleName_ = tostring(parse_json(tostring(parse_json(tostring(RawEventData.PolicyDetails))[0].Rules))[0].RuleName)
| extend Severity_ = tostring(parse_json(tostring(parse_json(tostring(RawEventData.PolicyDetails))[0].Rules))[0].Severity)
```

Timestamp	ActionType	Workload	PolicyName	SensitiveInformationType	Documents	Target	Subject	EndpointOperation_
Aug 8, 2024 6:30:5...	DLPRuleMatch	OneDrive	PCI Data Security Standa...	Credit Card Number	Book.xlsx	To: CC: BCC:		

Timestamp	Aug 8, 2024 6:30:50 PM
ActionType	DLPRuleMatch
Workload	OneDrive
PolicyName	PCI Data Security Standard (PCI DSS) Policy
SensitiveInformationType	Credit Card Number
Documents	Book.xlsx
Target	To: CC: BCC:
Severity_	Low

Figure 17: Advanced threat hunting query and results

6.4 Discussion

Defender for Cloud Apps proved to be very accurate in locating sensitive information based on the classification by Purview. Most sensitive data categories were successfully detected by the pre-configured and custom DLP policies. Its ability to modify DLP policies in a flexible way made it possible to improve the detection process and was very effective in improving the detection of complex patterns and decreasing false positives. After testing, it was discovered that Microsoft Defender for Cloud Apps' alerting functions were responsive. As soon as sensitive data was found, alerts were created, ensuring fast action. The DLP rules' remediation actions like limiting access, encrypting files, or notifying administrators were carried out as planned. These measures offered a solid framework for reducing risk connected to data loss.

The DLPTest[3] sample data was helpful, but it did not capture the variability of real-world data. Thus, additional validation of the results using a wider range of data sets is necessary. Sensitive data that was complicated or hidden proved harder to detect, showing that fine tuning of DLP rules and machine learning models could be helpful. More data sources would have improved results of the research. A more thorough assessment of the DLP system would have been possible with the inclusion of a larger range of sensitive data types and formats. The efficiency and technical implementation of DLP in Defender for Cloud Apps were the primary focus of this research. A better grasp of the solution might be gained by expanding the evaluation to include user experience, simplicity of policy management, and integration with other security tools. The DLP policies were not tested in a real-world setting with actual users as part of the project. Practical challenges may be encountered while implementing the solution in a production environment.

[3] https://dlptest.com/

7 Conclusion and Future Work

The strong features of the platform in protecting sensitive data in cloud environments have been highlighted by this research on implementing DLP in Defender for Cloud Apps (CASB). Enabling DLP in Defender for Cloud Apps was proven to be effective in identifying, classifying, and protecting data from unauthorized access. It offers an integrated solution for organizations looking to improve their cloud security posture because of its real-time monitoring, customized rules, and seamless integration with other security tools. The project has highlighted some limitations, in handling false positives and detecting complex data patterns. The solution did well when processing classified and widely recognized data types, but more work is required to ensure it can handle all kinds of data in real-world situations. Despite these challenges, the solution performed well and provided useful details on how DLP operates in a cloud security system.

Future work should concentrate on further improving and modifying DLP policies to address the issue of false positives. This might involve developing context-aware filtering and creating more detailed rules to increase accuracy without reducing security. To improve the detection of complex and new data patterns, research is needed to investigate the integration of advanced AI and machine learning technologies into Defender for Cloud Apps. These technologies have the potential to detect threats that may go unnoticed by conventional DLP techniques. Integrating a SIEM solution like Azure Sentinel with Defender for Cloud Apps can help to centrally monitor and correlate DLP alerts with other security incidents. Benefits include automated incident response and threat intelligence, continuous compliance audits, correlated DLP and SIEM alerts for improved threat detection, and the utilization of threat insights and user behaviour analytics for proactive security measures.

References

Ahmad, S. et al. (2022) ‘RSM analysis based cloud access security broker: a systematic literature review’, Cluster Computing, 25(5), pp. 3733–3763. Available at: https://doi.org/10.1007/s10586-022-03598-z.

Ahmad, S., Mehfuz, S. and Beg, J. (2020) ‘Securely Work from Home with CASB Policies under COVID-19 Pandemic: A Short Review’, in 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART). IEEE, pp. 109–114. Available at: https://doi.org/10.1109/SMART50582.2020.9337121.

Ahmad, S., Mehfuz, S. and Beg, J. (2021) ‘Enhancing Security of Cloud Platform with Cloud Access Security Broker’, in Information and Communication Technology for Competitive Strategies (ICTCS 2020), pp. 325–335. Available at: https://doi.org/10.1007/978-981-16-0882-7_27.

Arjun Ramakrishnan (2023) Tech Guide: Microsoft Defender for Cloud Apps - DLP Policies, Cybersecurity Dev Blog. Available at: https://www.arjunrcybersec.dev/mipcasbdlp1/ (Accessed: 7 August 2024).

Batami Gold et al. (2024) Tutorial: Discover and protect sensitive information in your organization, Microsoft Learn. Available at: https://learn.microsoft.com/en-us/defender-cloud-apps/tutorial-dlp (Accessed: 5 August 2024).

Bhansali, A. (2023) 'Cloud Security and Privacy', International Journal for Research in Applied Science and Engineering Technology, 11(8), pp. 1539–1542. Available at: https://doi.org/10.22214/ijraset.2023.55416.

Bhattacharya, D. et al. (2021) 'Dynamic Cloud Access Security Broker Using Artificial Intelligence', in, pp. 335–342. Available at: https://doi.org/10.1007/978-981-15-7106-0_33.

Brouwer, M. and Groenewegen, A. (2021) Cloud Access Security Brokers (CASBs): Characterization of the CASB market and its alignment with corporate expectations Cloud Access Security Brokers (CASBs). Available at: https://doi.org/10.13140/RG.2.2.33067.36648.

Bs, V., Arvindhan, M. and Kalimuthu, S. (2023) 'The Crucial Function that Clouds Access Security Brokers Play in Ensuring the Safety of Cloud Computing', in 2023 16th International Conference on Security of Information and Networks (SIN). IEEE, pp. 1–5. Available at: https://doi.org/10.1109/SIN60469.2023.10475014.

Carol Bailey et al. (2024) Protect your sensitive data with Microsoft Purview. Available at: https://learn.microsoft.com/en-us/purview/information-protection (Accessed: 6 August 2024).

Chimpiri, T.R. (2024) Enhancing Cloud Security with Oracle Cloud Security Applications, EJBSOS European Journal of Business Startups and Open Society |. Available at: http://innovatus.es/index.php/ejbsos.

Choudhary, A. et al. (2022) 'Evolution and comparative analysis of different Cloud Access Security Brokers in current era', in 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP). IEEE, pp. 36–43. Available at: https://doi.org/10.1109/ICFIRTP56122.2022.10059416.

Chris Fox MSFT et al. (2024) Plan for data loss prevention (DLP), Microsoft Learn. Available at: https://learn.microsoft.com/en-us/purview/dlp-overview-plan-for-dlp (Accessed: 8 August 2024).

Chris Fox MSFT, Katy Koenen and Robert Mazzoli (2024) Create and Deploy data loss prevention policies.

Collier, B. (2023) Considerations for Selecting and Implementing Cloud Security Solutions Using Cloud Access Security Brokers. Marymount University.

Deloitte (2021) Data Protection. Available at: https://www.marketsandmarkets.com/Market-Reports/data-protection-market-214254944.html.

IT Governance (2024) Global Data Breaches and Cyber Attacks in 2024, IT Governance.

Javed, H. et al. (2024) 'Blockchain-Based Logging to Defeat Malicious Insiders: The Case of Remote Health Monitoring Systems', IEEE Access, 12, pp. 12062–12079. Available at: https://doi.org/10.1109/ACCESS.2023.3346432.

Jayaraman Soundarya (2024) 160+ Fascinating Cloud Computing Statistics for 2024. Available at: https://www.g2.com/articles/cloud-computing-statistics (Accessed: 3 August 2024).

Kaur, S. and Gupta, R. (2019) ‘Enhancing Features of Cloud Computing Using Cloud Access Security Brokers to Avoid Data Breaches’, European Journal of Engineering and Technology Research, 4(10), pp. 185–189. Available at: https://doi.org/10.24018/ejeng.2019.4.10.1518.

Khan, K. (no date) Cloud Access Security Brokers (CASBs): Enhancing Cloud Security Posture.

Kharb, L. and Chahal, D. (2023) ‘Cloud Access Security Brokers: Strengthening Cloud Security’, International Journal of Research Publication and Reviews, 4(8), pp. 642–644. Available at: https://doi.org/10.55248/gengpi.4.823.50412.

Kodurupati, P. (2022) ‘Loss of Data Control Scenarios and Best Data Loss Prevention (DLP) Practices’, PragmaEdge LLC, 1(3), pp. 1–3. Available at: https://doi.org/10.47363/JAICC/2022(1)272.

Kumar, H. and Gupta, H. (2023) ‘Cloud Security: An Innovative Technique for the Enhancement of Cloud Security’, in 2023 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). IEEE, pp. 411–416. Available at: https://doi.org/10.1109/ICAC3N60023.2023.10541567.

Obregon, L. (2021) A Technical Approach at Securing SaaS using Cloud Access Security Brokers.

Ramesha, K. and Sahni, V. (2023) Adaptive Cloud Access Security Broker.

Reddy, A.R. et al. (2023) ‘Detecting and Preventing Unauthorized User Access to Cloud Services by CASBs’, in 2023 Second International Conference on Electronics and Renewable Systems (ICEARS). IEEE, pp. 868–873. Available at: https://doi.org/10.1109/ICEARS56392.2023.10085406.

Selvam, P.A. (2022) ‘Secure Cloud Services by Integrating CASB Based Approach’, INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, 04(04). Available at: https://doi.org/10.55041/ijsrem15210.

Shahzad, A. and Machado de Sousa, E. (2021) Data Loss Prevention from a Malicious Insider : Cloud Service Providers’ Perspective. Available at: https://aisel.aisnet.org/pacis2021.

Tayouri, D. et al. (2022) Cybersecurity in Agile Cloud Computing–Cybersecurity Guidelines for Cloud Access, IEEE Access.

Wason, R., Aghili, S. and Zavarsky, P. (2020) An integrated CASB implementation model to enhance enterprise cloud security. Concordia University of Edmonton.

Yiliyaer, S. and Kim, Y. (2022) ‘Secure Access Service Edge: A Zero Trust Based Framework For Accessing Data Securely’, in 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC). IEEE, pp. 0586–0591. Available at: https://doi.org/10.1109/CCWC54503.2022.9720872.