panther

Reducing Costs by Moving to a Cloud-Native SIEM

Leveraging cloud services, detection-as-code, and security data lakes to improve security posture and reduce the operational burden of SIEM.

runpanther.io

panther

_01 Introduction

Security, compliance, and system availability are the cornerstones of modern business. With organizations increasingly adopting cloud -based technologies, the amount of information needed to maintain secure and efficient operations becomes untenable.

Traditionally, IT and security teams use event log data to help detect, trace, and respond to security and availability issues. However, with distributed workforces and networks, the number of devices, users, applications, and systems generating log data fundamentally alters the way organizations need to collect and store data to achieve visibility across their environment.

Additionally, with more organizational stakeholders needing more information to make educated decisions, siloed data is more problematic than ever, leading to reduced visibility and productivity. Security, operations, engineering, and even compliance teams frequently need access to the same data sets and information. Segregating this information across multiple tools often means duplicating data, storage costs, and activities, which leads to skewed and inaccurate data.

As traditional security boundaries disappear, organizations need solutions to meet this new agile infrastructure where it lives - in the cloud. Instead of purchasing multiple "modernizing" tools, organizations need to modernize their mindset around data collection, storage, and analysis to reduce costs, increase productivity, and enhance security.

panther

_02 Too Much Information, Too Little Space, Never Enough Money

In on-premises environments where organizations could control the number of devices, ingesting data into a single location made sense. With a limited attack surface and a finite set of digital assets, organizations could reasonably estimate their required storage needs. With the move to the cloud, however, infinite resources and distributed workforces increase overall data volume changing data collection processes and requiring different tools to manage security.

Expanded Attack Surface

The term "expanded attack surface" has become so ubiquitous that, at this point, it might seem like nothing more than a series of buzzwords. But with remote workforces bringing their own devices and connecting to networks from anywhere, organizations are struggling to manage all the new entry points and backdoors that attackers can exploit.
Most professionals recognize that the COVID-19 pandemic accelerated digital transformation, and Netskope's February 2021 "Cloud and Threat Report" highlights the impact by noting the following tangible changes:

- 20% increase in cloud application adoption
- 83% of enterprise users use personal applications, leading to shadow IT
- 690 distinct cloud apps are used per month by organizations with 500-2,000 employees

From a security standpoint, these numbers can feel overwhelming. IT departments need to secure too many applications across too many users. Meanwhile, threat actors doubled down on their ransomware and malware attacks, with the Federal Bureau of Investigation (FBI) reporting more than 240,000 phishing attack complaints in 2020, more than doubling the volume of attacks in 2019.

While IT professionals focus on securing endpoints more than ever, an article from Expert Insights helps quantify how difficult it is for IT teams to identify and manage endpoints, noting that:

- 70% of IT professionals estimate an average of 750 endpoints; and
- 30% of IT professionals have no idea how many endpoints they need to manage

As the attack surface expands beyond traditional perimeter boundaries, organizations need solutions that offer flexibility and visibility at scale.

Cost of Tools

As organizations need more visibility into security, they gravitate towards tried and tested solutions. Problematically, these tools were born out of on-premises environments, and they often fail to scale and meet cloud-related security needs at a cost-efficient price point.

Importantly, the security problems faced today are not for lack of spending. Ponemon's 2021 "Second Annual Study on the Economics of Security Operations Centers: What is the True Cost for Effective Results?" found that in 2020 organizations intended their average annual spends to be:

- $183,150 for security information and event management systems (SIEMs)
- $345,150 for Security Orchestration, Automation, and Response (SOAR)
- $285,150 for Managed Detection and Response (MDR)
- $333,150 for Extended Detection and Response (XDR)

Additionally, respondents noted that they choose to invest in SOARs primarily for "Alert or Event Analysis" and "Alert or Case Enrichment." Further, a cost analysis from Blumira breaks down the costs associated with SIEM to show where the bulk of the money goes:

- **Hardware**
  - SIEM Hardware Small $25,000
  - SIEM Hardware Medium $60,000
  - SIEM Hardware Large $100,000
- **Infrastructure**
  - Servers $8,000

- Storage $1,500
- Switches $3,000

- **Software**
  - Event volume – 5G $8,000
  - Event volume – 20G $24,000
  - Event Volume – 100G $40,000
  - Event Volume – Other $100,000

- **Annual Support**
  - 20% of cost of software + hardware

While these costs might have been manageable, albeit high, for on-premises infrastructures, cloud migration makes these costs difficult to estimate accurately. Overestimating costs means paying for resources the organization does not need. Underestimating costs leaves the organization unable to optimize its use of the tool.

Difficult to Customize

For organizations that can afford the technology, customization becomes the next hurdle. With no "one size fits all" approach to security, organizations need to configure their tools to meet their needs.

Deployment Issues

Deploying most SIEM's, and especially legacy SIEM or SOAR is a cumbersome and time-consuming process. According to a TechTarget article, this process can take 90 days or more. Organizations need to define rules, ensure appropriate configurations, and test the deployment to ensure it works as intended. Despite "out of the box" integrations and rulesets, organizations will frequently need to also build out customized use cases. All of this takes time, and in security, time is always of the essence. Simultaneously, organizations still need to manage security, meaning it is likely using its original toolset while setting up the new SIEM/SOAR. Ultimately, all of these costs add up, even if they are not always visible when the tool is purchased.

To deploy a SIEM or SOAR, the organization needs at least a:

- SIEM administrator
- Security analyst
- SIEM expert to fine-tune

The "Second Annual Study on the Economics of Security Operations Centers: What is the True Cost for Effective Results?", linked above, also supports the existence of these "hidden costs." It notes that the security engineering work used to integrate disparate security data, build-out rules and content, and automate processes costs an organization, an average of $2.76m per year.

In short, organizations can only use new security tools if they have the right people to make the tools work, and staff costs money.

Data Collection Issues

Since data storage costs need to be valued into the purchase, organizations need to make difficult decisions about the security information they choose to collect, store, and analyze. Organizations find themselves making data decisions based on cost rather than need.

Limiting the data collected also limits the data available for investigations. This then limits the security team's ability to use disparate data to build higher fidelity alerts and makes investigating threat actor behaviors more difficult and time consuming.

While best practices for collecting and analyzing data exist, threat actors also understand these concepts. Since the most malicious actors also understand how detection is performed, they also understand how to better hide their actions. Fundamentally, data collection becomes a series of "best guesses" before an incident occurs, guesses that threat actors know in advance.

panther

_03 Getting Your Feet Wet: Build a Security Data Lake

As organizations modernize their business processes, they can also modernize their approach to security. By embracing the same big data principles that enable business intelligence, security teams can make better-informed decisions and mitigate risks more rapidly.

What is a security data lake?

Modern data lake and data warehouse architectures use cloud-native storage to centralize enterprise data for analytics. By using cheap and limitless blob storage, cloud data platforms support cost-effective and flexible analytics for business intelligence and data science use cases.

A security data lake is where the cloud-based data platform architecture is applied to security use cases. Generally, the security data lake is not an isolated datastore but an extension of the business's central data platform. This means that security and business datasets can be analyzed together by the security team, including with support from the company's existing data analytics specialists. In addition to consolidating storage and analytics, a security data lake supports using the enterprise BI reporting tool of choice for security metrics and self-service dashboards.

_03

Why create a security data lake?

Security teams have traditionally relied on SIEM for security data collection and analytics separately from the enterprise's central data platform. Modern cloud data platforms support semi-structured data and fast search, making them an alternative to traditional SIEM backends with greater scalability, a better cost model and more powerful analytics. By moving to a security data lake, security teams can remove limitations imposed by the legacy architecture of traditional SIEM solutions and collect all their data cost-effectively to the company's cloud data platform. This enables a unified source of security information with all the data needed for threat detection and response in a single location. Through better visibility and context, detections can be significantly more accurate, false positives can be eliminated and breach investigations accelerated.

For example, a security data lake can help teams:

- Scale resources dynamically, enabling affordable long-term data retention, faster searches, and more robust investigations
- Reduce data duplication to streamline security activities
- Automate activity correlation across security event logs
- Enrich time-series data with context, like incorporating geolocation information and device
- Manage dynamic as well as static information, like IP addresses
- Reduce the cybersecurity skills needed for robust search
- Make historical data more accessible for enhanced threat hunting and forensics

In addition, security data lakes overcome the primary problems associated with the collection, aggregation, correlation, and analysis of high-scale security data. For instance:

- Scalability and flexibility mean never having to take “best guesses” with which data to collect and which data to ignore.
- Cloud storage makes costs manageable and predictable.
- Customized reporting with business intelligence tools provides broader visibility and enables easier communication across business-level stakeholders.
- Ease of use reduces the cybersecurity skills gap’s impact on staffing and budgets.
- Data deduplication reduces silos across security, operations, development, and compliance teams.

panther

_04 If You Build It, They Will Come: 5 Steps to Building Your Security Data Lake and Analytics

Although building a security data lake is easier and more efficient than traditional solutions, organizations still want to create a plan. However, while planning most security technology deployments focuses on what organizations can't have, planning a security data lake focuses on what they can do.

Step 1: Know Your Why, Not Your What

Building a data lake focuses on how your organization wants to use data instead of what data your organization wants to use.

When building your use case, think about your current and future plans, including:

- Business roadmap: will your build sustainably align with your business goals?
- Risk roadmap: will your build be flexible enough to manage emerging risks?
- Compliance roadmap: what reports does your compliance/audit team need today, and what might they need based on your regulatory landscape?

Panther's data-first approach leverages serverless infrastructure, decoupling storage and compute. Security teams can store everything they need to build out their security data lake, creating a scalable and maintainable approach with clean data.

Step 2: Choose a home for your data

After establishing your use cases, decide how and where you want to store your security data. Security teams should work with their counterparts on the data analytics teams to consider storage options. In some cases, the data analytics teams may already have a storage location. In other cases, the security team may be building from the ground up.

While choosing where to store your data, consider the following capabilities:

- Performance, manageability, and scale
- Value-added native integrations for your data
- Support for structured and unstructured data
- Ease of connecting data sources
- Security controls for protecting sensitive information
- Business intelligence connectors for visualizations and reporting

By defining your use cases, especially those that enable cross-functionality within the organization, you can make meaningful decisions around where to store your security data.

Panther's native integration with Snowflake enables organizations to easily build a security data lake that fits their security, operations, and business needs.

Step 3: Clean, Structured Data

The biggest cost associated with security tools is unmanageable data. Nearly every technology that your security team needs to secure has its own data format. To optimize your security data lake, you need a solution to normalize and standardize all of these different formats. If your security team still needs to understand the intricacies of every data format to intelligently search and analyze the data, you're losing efficiency and time.

When trying to find the right solution, you should consider how to:

- Parse and normalize data before storing it in the data lake
- De-duplicate data
- Use common formats rather than tools-specific schema
- Automatically structure data into tables, columns, and rows

Panther converts raw logs into structured, formatted JSON data, normalizing common fields like:

- IP addresses
- domain names
- hashes
- Usernames
- Email addresses

Panther structures raw logs as it ingests them, making them available for programmatic analysis in real-time to help reduce key security metrics like mean time to detect (MTTD) and mean time to investigate (MTTI).

Step 4: Design Alerts, Notifications, and Queries

By building a security data lake instead of buying a traditional SIEM solution, security teams can customize alerts, notifications, and queries. Security teams can tailor how they create alerts and notifications because they can use any data and any context available.

With the freedom to decide what data to use and how to use it, security teams no longer need to rely on generic alerts and queries. They can gain greater visibility and find new ways of detecting risks that threat actors may not already be privy to. This capability helps security teams align tactics, techniques, and procedures for their detection and response program, giving them the ability to improve security metrics like MTTD, MTTI, and mean time to recover (MTTR).

When considering a security data lake solution, organizations should consider how they are able to:

- Customize alerts
- Leverage out-of-the-box, pre-written detections
- Correlate activity across disparate data sets

Panther offers a variety of out-of-the-box detection content to identify malicious activity like Command and Control (C2) beacons, unusual logins, and password spraying. Additionally, with Panther, security teams can run queries and rules on a schedule to create detections that work over large time spans while aggregating data from multiple sources.

 _04

Step 5: Leverage Business Intelligence Reporting Tools

Data analytics teams have a wealth of business intelligence reporting tools available to them. With a security data lake, security teams can use these tools, and leverage their counterparts on the data team, to enhance their reporting across multiple internal stakeholders.

Instead of being tethered to built-in reporting capabilities or choosing a security solution based on the reports you want, you can communicate effectively with various stakeholders by building out the reports you need with a dedicated BI tool.

By leveraging business intelligence tools, your security data lake can be used to improve cross-functional communications. For instance, you can:

- Report security posture to senior leadership and Board of Directors
- Document security for risk, compliance, and audit teams
- Provide development teams with security visibility into their software development life cycle management processes
- Work with operations teams to identify root causes

With Panther's Snowflake integration, security teams have access to the same industry-leading business intelligence solutions used by data teams, including Tableau, Mode, Looker, and more. By treating security as a data problem, organizations can demystify security using tried and true best practices that enable stronger security and better governance.

panther

_05 Detection-as-Code: A Modern Approach to Security

By leveraging modern tooling, organizations can modernize their entire approach to security detection, investigation, and response. Panther's platform reduces the operational and financial burdens associated with traditional SIEM tool.

Panther believes that strong security starts with good data. With Panther, security teams can perform log collection and analysis, cloud security, and security analytics in a single platform, leveraging the scalability of the cloud and the flexibility of developer tooling. Panther supports collecting, correlating, and analyzing data from various sources, including:

- Cloud services providers like AWS, Google Cloud Platform (GCP), and Azure
- Security tools like Osquery and OSSEC
- Endpoint monitoring tools like Crowdstrike
- Software-as-a-Service (SaaS) applications like G Suite, Box, and Slack
- Single-sign-on (SSO) providers like Okta and OneLogin
- Network monitoring tools like Cisco Umbrella, Juniper, and Cloudflare

As Panther ingests logs, the platform parses and normalizes the data in real-time to provide security teams rapid visibility into potential threats. Leveraging this real-time stream processing, security teams can build their own customized detections using Python. By treating Detection-as-Code, security teams can write, test, and deploy alerts customized for their organization's environment, reducing the need for specialized skills or a proprietary language.

In addition, teams can rapidly onboard AWS accounts to manage their cloud security using prebuilt CloudFormation templates. With Panther's flexible, out-of-the-box policies, organizations can begin monitoring their cloud configurations, reducing risk through daily scans and real-time alerts.

Bringing everything together with a single, unified security data lake, Panther enables security teams to power threat detection and investigation at scale. By decoupling storage and compute, security teams can scale up their infrastructure just-in-time to complete investigations and forensics against large data sets.

Leveraging the power and speed of the cloud no longer needs to be constrained to business operations. With Panther, organizations can apply those same dynamic qualities to their security processes, gaining the visibility they need while reducing silos. By unifying data and bringing together internal stakeholders, security teams using Panther communicate more effectively and accelerate and enhance their security and compliance postures.

panther

Thank you.

runpanther.io