



How to Deploy Complete Security Analytics

CLOSE THE "DARK DATA" BLACK HOLE
WITHOUT OPENING A MONEY PIT





What's inside:

- 3 The untapped value of log data
- 4 Why SIEM vendors fail to deliver security
- 6 Holistic security analytics delivers complete peace of mind
- 8 What cloud-built data warehouses deliver beyond security analytics
- 9 Security you can bet the network on
- 10 About Snowflake

The untapped value of log data

When it comes to cyber-threats, security professionals are in a tough position. They are charged with maintaining a secure IT environment that protects against malicious threats, prevents network breaches and guards against data theft. At the same time, attackers have become increasingly sophisticated as security events continue to escalate in size and number. According to IDG, the average organization experienced 130 breaches in 2017, which represents a 27% increase over the previous year and almost twice as many as five years ago. All the while, bad actors can dwell unnoticed inside a network for over 200 days at a time.

One of the best cybersecurity defenses is to analyze log data because it programmatically records every network system change. However, IT teams struggle to monitor, use and analyze log data in an efficient and cost-effective manner due to the constraints and prohibitive costs of existing security information and event management (SIEM) tools. As a result, log data is a largely untapped resource, and organizations continue to find themselves vulnerable to security threats.

Modern cloud data warehousing opens up new opportunities for robust security protection by storing and analyzing all log data in order to reveal deeper insights. In this eBook, you'll learn about the limitations of SIEM tools and how they deliver substandard security analytics, putting organizations at risk. On the flipside, the benefits of adopting a built-for-the-cloud data warehouse demonstrate how organizations can enable holistic security analytics for entire business teams at significantly lower costs.



Why SIEM vendors fail to deliver security

Log data is used to record every change within a network. In theory, organizations that capture log data have all the information necessary to detect and stop cyberthreats. However, companies today are hampered by their SIEM tools. Using these tools requires employees with special skill sets. They are also designed and priced primarily to deliver live data. Any use of long-term log data for ad hoc or historical analysis is difficult due to cost, storage restrictions, data structure and accessibility.

The result: a massive black hole of “dark data” (i.e., the information organizations collect and store but fail to use for activities such as analytics). When it comes to cybersecurity, dark data is the enemy that opens the gates and exposes organizations to numerous attacks. As long as organizations continue to use SIEM tools to manage and store their log data, they will face the following challenges:

HIGH STORAGE COSTS

Log data is a vast resource that continues to grow in size and scope with the spread of IoT and web-enabled technologies. While organizations increasingly recognize how valuable log data is and would like to analyze it extensively, the costs associated with storage for SIEM tools is prohibitive.

In fact, the leading enterprise SIEM tools are some of the most expensive resources available for data analytics. SIEM providers charge exorbitant amounts per gigabyte for streaming and storage, which makes pricing for tens or hundreds of terabytes per day unaffordable. Consequently, companies of all sizes end up storing only a small fraction of their log and security data. A recent 451 Research Information Security Survey revealed that only 21.6% of organizations pass more than 81% of their log data through their SIEM system, while 38% pass less than 30%.

As a result, organizations are forced to make a difficult choice: pay more than they can afford to monitor all of their log data, or ignore a large portion of that data and any threats that may exist within it.

DIFFICULT TO COMBINE DATA

In addition to expensive storage, the log data is kept separate from the other tools and systems within an organization. This situation creates a silo of log data that is challenging to decipher on its own, especially when IT employees lack insight into whether a particular activity deserves investigation with other data or is a normal occurrence. For example, is a sudden uptick in website activity attributable to a distributed denial-of-service (DDoS) attack or simply the result of a successful marketing strategy?

Protecting development codebases is one area in which combining log data with other system data could be particularly useful. If data from an issue tracking system such as Jira can be compared to log data, then it's easy to monitor and flag any unusual commit activity that doesn't match expectations, thus stopping nefarious activity in its tracks. Unfortunately, SIEM tools do not allow these data sources to interact.

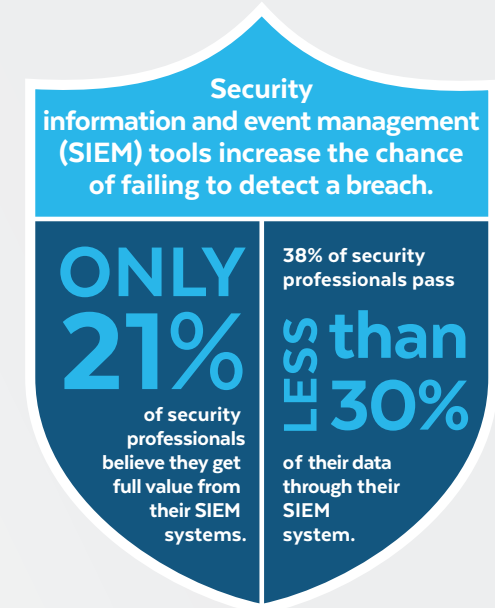
Some organizations attempt to work around this situation by using complicated ETL and ELT tasks or programmatic jobs that require additional maintenance. However, without the ability to seamlessly combine log data with additional data sources, most organizations find it nearly impossible to conduct analysis, detect important patterns and answer pressing business or security questions.

NO DATA DEMOCRATIZATION

Data is everyone's business. While database administrators have traditionally handled all enterprise analytics activity, today's organizations are full of business users, data scientists and analysts who want to query that data as well. With data democratization as today's norm, many teams inside an enterprise need to run analytics against all types of data, including log data.

However, SIEM tools are not built for unlimited data querying by concurrent users with high performance. SIEM tools tightly couple storage with compute. This makes it cost-prohibitive for organizations to scale up, down and out (concurrency) for an unlimited number of users and jobs accessing a single copy of the data with their own, instantly elastic compute resources.

SIEM limitations also lead to poor performance when querying historical data for long-term analysis. Some providers don't even store log and machine-generated data beyond 30, 60 or 90 days. Is it any wonder that bad actors can lurk in networks for over six months without being detected?



Holistic security analytics delivers complete peace of mind

To plug the security hole that SIEM tools create, organizations need look no further than a modern cloud data warehouse. Anyone with a security question can immediately find answers when all log data is stored alongside other enterprise data in a data warehouse built for the cloud that's designed with affordability, scalability and flexibility in mind.

This modern solution provides organizations with a holistic and collaborative approach to security data that easily loads, integrates and analyzes all of your dark data for effective security insights to detect and thwart attacks. Here's how:

AFFORDABLE STORAGE AND ANALYSIS OF LOG DATA

Data cannot be analyzed if it's too expensive to store. It's that simple. With a cloud-built data warehouse, the flexible and elastic architecture that truly separates compute and storage makes it possible to store unlimited amounts of log and security data in a single location at affordable, long-term storage rates. A simple cost comparison demonstrates that organizations pay just 5-10 percent of the cost of SIEM services by opting for cloud data warehouse storage, and they receive a wealth of additional capabilities and benefits.

A cloud-built data warehouse unifies log data from SIEM silos and integrates and analyzes it alongside other data and contextual information. Related to the software development example mentioned above, bringing together data from Jira and log files in one place allows security teams to quickly identify commits that weren't anticipated. By flagging these discrepancies, security teams can easily and automatically detect potential threats.

It never makes sense to throw away log data because of cost. Losing long-term historical data puts organizations at risk of repeating the same security mistakes and makes it challenging to spot bad actors lurking in the network. With a cloud-built data warehouse, it's finally affordable to store unlimited amounts of log data indefinitely, use it to spot trends and anomalies and develop more effective security measures.





ANALYTICAL TOOLS OF CHOICE

Rather than putting limitations on the tools used to load, query and analyze security and log data, a cloud-built data warehouse provides the opportunity to leverage business intelligence (BI), ETL and ELT, standard SQL and open source tools that best match a user's needs. Regardless of the data type or its source, full support exists for structured and semi-structured data, which allows diverse data to be integrated and analyzed with ease.

In addition, the continuous data loading capabilities of a cloud-built data warehouse help ensure that users always access the most up-to-date data possible in order to deliver timely and holistic security analytics. With combined log and enterprise data, it's easier for security teams to distinguish real threats from harmless day-to-day activities.

COMPLETE ACCESSIBILITY AND UNLIMITED CONCURRENCY

With unlimited amounts of security data stored affordably and alongside all other enterprise data, the final piece of the puzzle is allowing anyone within the organization to access and use the data. With a cloud-built data warehouse that separates compute from storage, this modern architecture ensures unlimited concurrency against a single copy of the data without impacting performance.

In addition, users across the organization can access the same shared datasets at any time to enable analysis without disruption. That means security teams can access log data at the same time as any other team through individual and elastic compute clusters for each user's query or any other job against the data. With automatic scaling and the addition of compute clusters as concurrency increases, virtual warehouses provide a cost-effective way to empower teams to conduct analysis.



What cloud-built data warehouses deliver beyond security analytics

THE BENEFITS OF ADOPTING A CLOUD-BUILT DATA WAREHOUSE GO BEYOND ENABLING HOLISTIC SECURITY ANALYTICS TO INCLUDE:



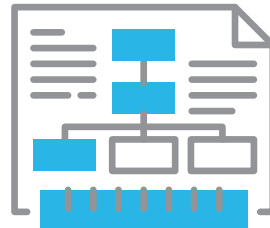
Scalability

- Independent, unlimited and instant scalability of compute and storage
- Multiple independent compute clusters that share data without contention between workloads
- Auto-scaling during concurrency surges



Affordability

- Pay-by-the-second usage
- Automatically adjust resources for new use cases or data surges
- Reduce capacity planning exercises to simple assessment of need



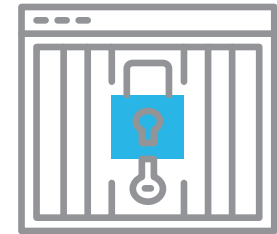
Semi-structured Data

- Full support for JSON, Avro, Parquet and XML data. Load data in its native format into a Variant type column and transform it, if needed, with dot notation extensions to standard SQL. No need to transform or define a schema before load.



Modern Data Sharing

- Secure, governed and live enterprise data sharing between data providers and data consumers
- Share data across an organization or an enterprise and between enterprises
- Cross-region data replication for data providers with worldwide data consumers and multi-region data resiliency



Security, Management and Data Protection

- Out-of-the-box, always-on secure data environment, including a virtual private cloud (VPC)
- Near-zero administration with built-in performance tuning
- Zero-copy cloning to eliminate any impact to performance
- No need for conventional backups with zero-copy clones and rollback features such as time travel



Security you can bet the network on

Your security analytics is only as strong as the data that powers it. While traditional SIEM tools fall short with their prohibitive costs, isolated data silos and lack of historical data, a cloud-built data warehouse sheds light on the dark data that SIEM tools leave behind. Adopting a security analytics warehouse provides the up-to-date and holistic view of log and security data to empower the most secure network.



About Snowflake

Snowflake is the only data warehouse built for the cloud. Snowflake delivers the performance, concurrency and simplicity needed to store and analyze all data available to an organization in one location. Snowflake's technology combines the power of data warehousing, the flexibility of big data platforms, the elasticity of the cloud and secure data sharing at a fraction of the cost of traditional solutions. Snowflake: Your data, no limits. Find out more a [snowflake.net](https://www.snowflake.net)