SOLUTION BRIEF

# StreamSets for Cybersecurity

## Overview

In an increasingly interconnected world, threats to our digital welfare are everywhere. Cybersecurity attacks can menace individuals, organizations, and infrastructure. The multibillion-dollar global cybersecurity industry relies on technologies, regulations, and best practices to help keep data, devices, applications, networks, and systems—and the people and organizations depending on them—safe from attack.

## Challenges

Reminders of potential cyber threats are everywhere, whether it's news of a distributed denial of service (DDoS) attack on a major website, stolen credit card numbers or other sensitive personally identifiable information (PII) from a national retailer, hacking for political or financial reasons, or warnings about possible infrastructure breaches. At a corporate level, every organization must be on alert to potential cybersecurity threats—and have ways to prevent, detect, and mitigate those threats.

Minimizing the effects of cybersecurity incidents requires the ability to understand and deal with:

- An expanding list of threat vectors.
- High-velocity data sources such as network and web logs.
- Real-time alerting for instant remediation.

Standing in the way of organizations' implementation of effective cybersecurity systems and processes are certain realities, such as:

- The rise in cybercrime creates ingestion scenarios that can challenge legacy systems. Most single-source threat systems struggle to keep pace with the rising demand for data.
- Effective cybersecurity mediation involves detecting and often acting in real time, but many existing solutions do not handle real-time data efficiently.
- Defending against expanding threats in real time requires increasingly expensive resources that often don't scale well.
- Traditional Security Information and Event Management (SIEM) systems were designed to handle only specific data sources, so they're not well equipped to deal with today's diverse array of sources.
- Addressing vulnerabilities as they happen has required specialized systems and skills, too often leaving forensic analysis as the default approach to cybersecurity breaches.

Organizations are moving more of their business and operating assets online, making it imperative that they be able to monitor and control their data movement to and from the full range of systems and data sources. Unstructured and semi-structured data—including everything from website traffic logs to images and social media data—are particularly difficult to manage, leaving serious gaps in visibility.

Streaming data presents an opportunity to combat cyber threats by incorporating data sources such as logs from security systems, network servers, and Windows endpoints to better detect threats and attacks. But because many organizations lack the ability to reliably manage streaming data, it is more likely to add to their cybersecurity vulnerabilities than to help alleviate them.

### ABOUT CYBERSECURITY

Cybersecurity encompasses all the technologies, techniques, and practices employed to protect information, systems, and businesses from digital threats, which could include attacks, hacks, malware, ransomware, phishing, damage, data destruction, or unauthorized access.

### INDUSTRY EXAMPLES

**Financial Services**

Online banking and trading sites use cybersecurity solutions to reduce improper access to client funds. By monitoring customer behavior and website diagnostics along with network activity, financial services organizations can gain a clear picture of their threats.

**Telecom**

When bad actors commandeer IP addresses and run up large charges for data and telephone carriers, it can cause big problems for network providers. Actively monitoring network activity and usage and can help companies get ahead of cybersecurity events.

**Retail**

With more retail business being hosted online, retailers are looking for ways to keep their digital assets safe and keep customer data protected. GDPR imposes strict, new guidelines on how retailers need to manage fraud events. More than ever, retailers need effective cybersecurity solutions.

## INDUSTRY EXAMPLES

### Healthcare

In healthcare, many systems that house customer data are a closed loop and do not access the internet. But the addition of new data sources, medical sensors, and device data opens up new areas of concern around access and security.

### Technology

DDoS and other attacks can take down applications and services and ruin brand reputation. To ensure uptime and reliability, technology and SaaS solutions need to defend their environments from cyber criminals. Technology companies, at the same time, also must protect their own internal customer data.

### Insurance

Insurance companies house critically important and sensitive data about customers. Ensuring that medical, personal, and private information is kept from outside access is especially critical in these regulated industries.

## ABOUT STREAMSETS

StreamSets transforms how enterprises flow big and fast data from myriad sources into data centers and cloud analytics platforms. Its DataOps platform helps companies build and operate continuous dataflow topologies, combining award-winning open source data movement software with a cloud-native Control Hub. Enterprises use StreamSets to enable cloud analytics, data lakes, Apache Kafka, IoT, and cybersecurity.

Founded by Girish Pancha, former chief product officer of Informatica, and Arvind Prabhakar, a former engineering leader at Cloudera, StreamSets is backed by top-tier Silicon Valley venture capital firms, including Battery Ventures, New Enterprise Associates (NEA), and Accel Partners.

For more information, visit streamsets.com

## Solution

StreamSets enables real-time threat detection by giving its users the tools to ingest data from the full range of sources and to apply analytics in real time. As a result, organizations can detect advanced persistent threats and also improve their cybersecurity forensics.

**Deal with expanding threat vectors.** StreamSets users create data pipelines that can feed expanded analysis environments, also known as data lakes. StreamSets' real-time data delivery lets organizations move beyond forensic reporting to predictive remediation of cyber threats. And the StreamSets solution scales with the addition of more data sources, without schema re-design or the risk of data drift.

**Manage the full range of modern data sources.** StreamSets participates in a rich ecosystem of partners and tools to help address new data formats, providing pre-built connectors and destinations for common data sources and platforms. Using StreamSets, organizations can deliver faster value to analytics projects, irrespective of the data source—while expanding the capabilities of their current security systems with additional data sources.

**Perform real-time alerting and remediation.** StreamSets leverages the Apache Spark unified analytics engine for real-time detection of cyber threats via streaming data, using the popular Python/R modeling technology. Organizations benefit from StreamSets' collection of pre-built sources and destinations, which addresses all types of generated data and applies common machine learning and reporting processes. Enterprises can monitor and gauge performance and meet cybersecurity service-level agreements (SLAs) using StreamSets Control Hub.

## StreamSets Benefits

StreamSets enables organizations to:

- Understand streaming data to set up data pipelines along with thresholds for alerts to particular threats, such as DDoS attacks.
- Integrate with Apache Spark for machine learning and data science.
- Quickly add new data sources as requirements change.
- Manage entire topologies and individual pipelines.
- Track changes in threat vectors over time.
- Implement key performance indicators (KPIs) for data availability and accuracy.
- Create data SLAs to detect and remediate violations, then monitor data usage to meet data availability SLAs.
- Achieve affordable management, performance monitoring, and data encryption of real-time streaming data.

## Closing

Cybersecurity threats are arriving faster, and from more diverse sources, than ever before. The StreamSets platform helps deliver threat data in real time so that organizations can stay one step ahead of cyber threats.

Find out more about how StreamSets can help protect your data, including streaming data, from cybersecurity threats. Contact a StreamSets representative today.