>SOLUTION BRIEF_

Replay from Low-cost Storage

THE CHALLENGE

- Manual replay is cumbersome and time-consuming. Pulling data from different sources, owned by different teams, delays incident response.
- Significant time, expertise, and resources are required to ensure data is in the correct format to be replayed for the destination.
- Storing large volumes of data long-term for potential future replay is extremely costly.

THE SOLUTION

Cribl Stream enables you to store full-fidelity data in low-cost storage — such as Cribl Lake and/or object stores — while making it simple to pull data from any source and send it downstream for analysis, testing, or troubleshooting.

THE BENEFITS

- Easily collect data from multiple sources
- Seamless workflows to transform and route data, reducing unnecessary overhead
- Quickly and easily replay data on demand, avoid the expense of keeping all data live in analytics tools
- Cost-effectively store full-fidelity data in Cribl Lake or object stores for long-term retention

Organizations require always-accessible data to ensure they can quickly replay it to the appropriate destinations when the need arises. This allows for fast and flexible responses to critical challenges such as security incidents or audit reporting.

The challenge.

Every minute spent wrestling with data access and replay logistics is time lost in incident response. Rising data volumes make storing data costly, forcing tough retention tradeoffs and leaving organizations unprepared for security incidents or compliance needs.

Manual replay often involves pulling raw data from multiple sources owned by different teams, writing scripts to parse and resend it, and ensuring proper formatting for target systems.

The challenges don't stop there—storing data for future replay can be costly, especially as data volumes grow, forcing tough decisions about what to retain. License costs and tool limitations often prevent organizations from ingesting all their data or retaining it long enough to meet evolving needs. This poses significant risks, particularly in security scenarios where incidents may surface after retention periods end or involve data that was never ingested.

The Cribl solution.

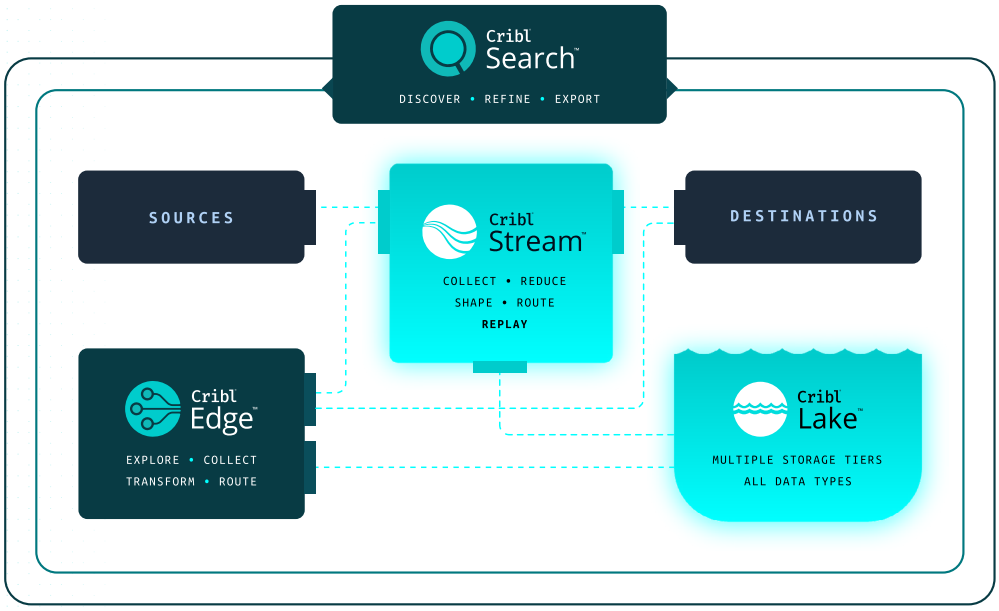
Cribl Stream is a powerful observability pipeline designed to precisely parse, process, and route your data. It ensures you get the data you need, in the formats you want, delivered to the right destinations.

Cribl Stream's Replay feature transform data management by simplifying data retrieval, allowing you to easily pull and replay data from Cribl Lake or object storage to various destinations. Store raw data like logs, metrics, and traces in cost-effective storage with customizable partitioning for easier searches. When needed, quickly re-ingest relevant data into analytics systems for investigations, audits, or compliance checks. Replay ensures your data is always accessible, enabling fast, flexible responses to challenges like security incidents or compliance needs.

Cribl Stream + Cribl Lake

Cribl Lake is a simple, quick-to-deploy, easy-to-use, data lake just for telemetry. Data is stored in open formats in Cribl Lake, allowing for seamless replay to analytics tools without the need for complex scripts or manual processes.

Cribl Lake provides a cost-effective solution for retaining large volumes of data over extended periods while keeping it readily accessible for investigations, audits, and testing. While Cribl Stream excels at data routing and replay orchestration, combining it with Cribl Lake creates a seamless telemetry lifecycle. When you pair Cribl Stream's smart parsing and processing capabilities with Cribl Lake's purpose-built telemetry storage, you can finally eliminate the traditional friction points of data accessibility, processing overhead, and cross-team dependencies.



Facets of Replay with Cribl

Route full-fidelity data to cost-effective storage.

Send your data to the most effective destinations, including cost-efficient storage options like Cribl Lake and object stores for long-term retention. Replay this data as needed, allowing you to analyze only what's essential now while storing the rest cost-effectively for future use.

Replay data to analytics systems as needed.

Cribl Stream revolutionized data management by introducing batch log reprocessing with replay. Beyond processing real-time streaming data, Stream enables you to collect data from a wide variety of sources, including object storage and REST APIs. While most data may be analyzed in real time, the addition of batch processing and Replay broadens both the range of data sources you can analyze and the timing of your analysis. Batch collection and replay put you in full control of your data.

Easily collect and replay from multiple sources.

With Stream, you can process and replay logs and metrics data from all REST endpoints. Cribl offers several different ways to discover and retrieve REST data, with both known-structure and schema-agnostic retrieval options.

> "If a big event occurs, we know we've got a longer tail of data that we can go back and investigate further That peace of mind Cribl gives us has been really helpful to have."
>
> Director of Enterprise Security at a Leading Energy Utility Company

Stream receives and replays data from many APIs and other data sources. These include Kinesis Firehose via the Kinesis HTTP endpoint, and raw HTTP data. Stream can also replay batch data collected from the Office 365 Service Communications API, for service incidents on Microsoft cloud services. Similarly, Stream can replay batch data from the Office 365 Management Activity API, for actions and events on Azure Active Directory, Exchange, SharePoint, and other Microsoft servers.

Schedule batch collection to automate replay.

Simplify operations by scheduling batch data collections and automating replay to seamlessly integrate with your workflows, reducing the need for manual intervention. Set recurring schedules for the distributed collection of data from multiple sources, and for replay to an analytics tool. Stream allows you to configure collections based on resource filters and constraints. You can also limit concurrent running instances of ad-hoc and scheduled jobs.

Better security breach investigations.

Many security breaches are discovered long after they start, sometimes several years later. Because of costs, most organizations do not retain the necessary data needed to investigate these breaches in their analytics systems, for more than a few months. Stream allows you to park full-fidelity data in a low-cost storage location such as Cribl Lake for as long as seven years.

When security breaches are discovered, Stream can efficiently collect data from Cribl Lake and/or object storage. Stream then replays security data to SIEM or UEBA systems, to quickly diagnose and resolve existing breaches and potential threats. Stream enables an affordable way to retain more data for longer periods of time — while still making that data easily accessible for necessary investigations, whenever they happen.

Meet stringent compliance requirements.

Cribl's solution also aids in meeting stringent compliance requirements. By enabling cost-effective long-term data retention and easy accessibility, organizations can meet data retention mandates for regulations like GDPR, CCPA, and industry-specific requirements. With Stream and Lake, teams can also quickly produce audit trails and reports for regulatory inquiries, and demonstrate due diligence in data management and security practices.

Summary.

As data volumes continue to grow exponentially and security threats evolve, Cribl's solution ensures you're ready for the future:

- Scalable architecture to handle increasing data loads
- Flexible data routing to accommodate new analytics tools and storage options
- Continuous updates to address emerging security threats and compliance requirements
- API-first approach for easy integration with future technologies

Combining flexibility, automation, and cost efficiency, Cribl's Replay functionality ensures your data is always ready for whatever your teams need—without breaking your budget.

ABOUT CRIBL

Cribl, the Data Engine for IT and Security, empowers organizations to transform their data strategy. Customers use Cribl's vendor-agnostic solutions to analyze, collect, process, and route all IT and security data from any source or in any destination, delivering the choice, control, and flexibility required to adapt to their ever-changing needs. Cribl's product suite, which is used by Fortune 1000 companies globally, is purpose-built for IT and Security, including Cribl Stream, the industry's leading observability pipeline, Cribl Edge, an intelligent vendor-neutral agent, Cribl Search, the industry's first search-in-place solution, and Cribl Lake, a turnkey data lake. Founded in 2018, Cribl is a remote-first workforce with an office in San Francisco, CA.

Learn more: www.cribl.io | Try now: Cribl sandboxes | Join us: Slack community | Follow us: LinkedIn and Twitter

SB-0018-EN-3-0125