Cribl

Get Real-Time Access to Archived Data with Cribl Search

HOW-TO GUIDE

CONTENTS

OVERVIEW

Use Cribl Search to eliminate re-hydrating ‘frozen’ data

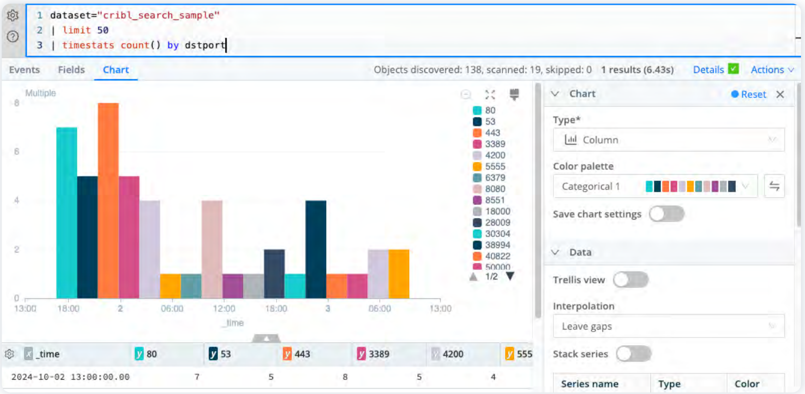
Cribl Search value in bullets

- Cribl Search flips the traditional search process on its head, delivering **a federated, search-in-place processing engine** that can access data where it is, where it was generated, or where it's been stored. In any storage medium and in any format using a single, intuitive query interface. With Search, you are able to explore data anywhere.

- **Search-in-place technology** empowers users to explore and analyze telemetry data that was previously unreachable—directly at its source, across any cloud, even from API endpoints. The power of Cribl Search lies in its strategic approach: locate and then forward only actionable data to your systems of analysis.

- **Separating data retention systems from analysis systems** saves significant costs. Build a data storage strategy that minimizes costs by using your storage of choice, including Amazon S3, Amazon Security Lake, Azure Blob, Google Cloud Storage, and compatible systems. Then collect and forward only what's needed.

Cribl Search solves your archive challenge



```
dataset="cribl_search_sample"
| limit 50
| timestats count() by dstport
```

Objects discovered: 138, scanned: 19, skipped: 0 1 results (6.43s)

Chart
Type*: Column
Color palette: Categorical 1
Save chart settings

Data
Trellis view
Interpolation: Leave gaps
Stack series

Series name	Type	Color

_time	80	53	443	3389	4200	555
2024-10-02 13:00:00.00	7	5	8	5	4	

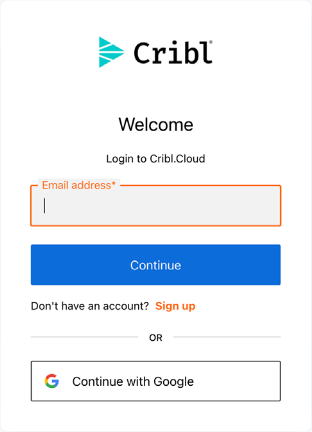
Archived data

As data ages out of a system of analysis it is often relegated to the lowest cost storage, most times this means some form of “frozen” data. Getting data out of this type of archive is often a slow and painful process that typically means the data is not readily available and may take hours or days to retrieve, convert, thaw, and format correctly to be used. That’s not a good option when investigating an event in real time and the stored data is invaluable for investigating the event, identifying trend analysis, decision-making, and strategic planning.

Cribl Search accesses data archives in real time

1 | Create a Cribl.Cloud account

If you don’t already have one, you can create one in no time:

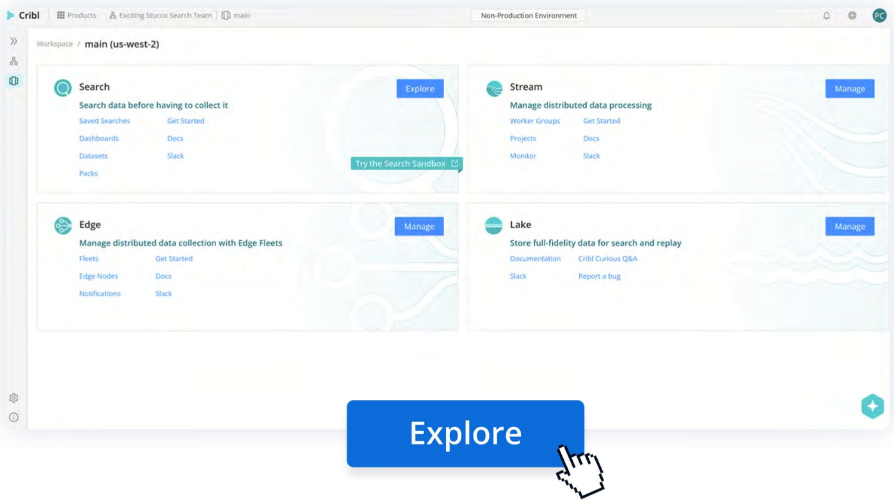


- Get started at **Cribl.Cloud:** [FREE cloud account]
- Only takes a minute
- No credit card required!

2 | Launch your Search instance

Gain immediate access to all Cribl resources. Just click and go:

- From the main screen select **Explore Search**
- Don’t forget to check out the other Cribl apps, too
- All apps are integrated on the back end

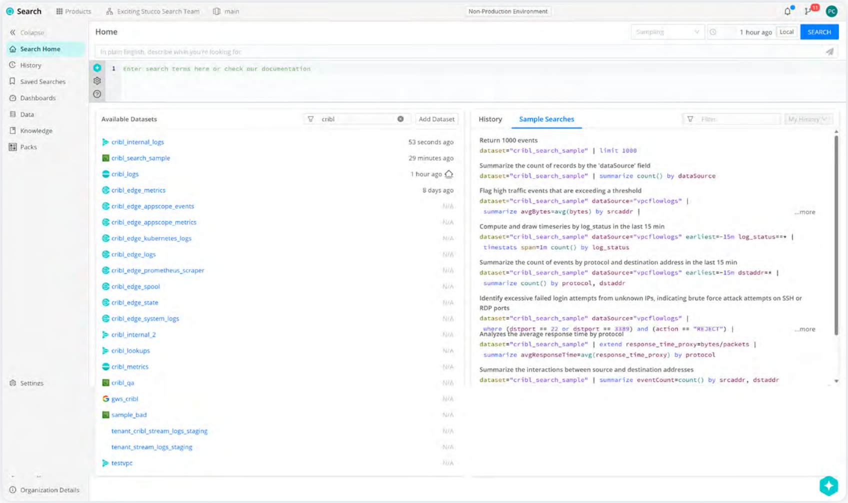


Cribl Search accesses data archives in real time

3 | Try a quick search to get the feel

Take a quick test drive, make a couple of clicks and your searching:

- Looks complicated—It isn't!
- Just click on one of the sample searches (right side of page)
- Now click on one of the results to see what's provided



```
1  Enter search terms here or check our documentation
```

Available Datasets	
cribl_internal_logs	53 seconds ago
cribl_search_sample	29 minutes ago
cribl_logs	1 hour ago
cribl_edge_metrics	8 days ago
cribl_edge_appscope_events	N/A
cribl_edge_appscope_metrics	N/A
cribl_edge_kubernetes_logs	N/A
cribl_edge_logs	N/A
cribl_edge_prometheus_scraper	N/A
cribl_edge_spool	N/A
cribl_edge_state	N/A
cribl_edge_system_logs	N/A
cribl_internal_2	N/A
cribl_lookups	N/A
cribl_metrics	N/A
cribl_qa	N/A
gws_cribl	N/A
sample_bad	N/A
tenant_cribl_stream_logs_staging	N/A
tenant_stream_logs_staging	N/A
testvpc	N/A

Return 1000 events

```
dataset="cribl_search_sample" | limit 1000
```

Summarize the count of records by the 'dataSource' field

```
dataset="cribl_search_sample" | summarize count() by dataSource
```

Flag high traffic events that are exceeding a threshold

```
dataset="cribl_search_sample" dataSource="vpcflowlogs" |
 summarize avgBytes=avg(bytes) by srcaddr |
```

...more

Compute and draw timeseries by log_status in the last 15 min

```
dataset="cribl_search_sample" dataSource="vpcflowlogs" earliest=-15m log_status==* |
 timestats span=1m count() by log_status
```

Summarize the count of events by protocol and destination address in the last 15 min

```
dataset="cribl_search_sample" dataSource="vpcflowlogs" earliest=-15m dstaddr=* |
 summarize count() by protocol, dstaddr
```

Identify excessive failed login attempts from unknown IPs, indicating brute force attack attempts on SSH or RDP ports

```
dataset="cribl_search_sample" dataSource="vpcflowlogs" |
 where (dstport == 22 or dstport == 3389) and (action == "REJECT") |
```

...more

Analyzes the average response time by protocol

```
dataset="cribl_search_sample" | extend response_time_proxy=bytes/packets |
 summarize avgResponseTime=avg(response_time_proxy) by protocol
```

Summarize the interactions between source and destination addresses

```
dataset="cribl_search_sample" | summarize eventCount=count() by srcaddr, dstaddr
```

4 | Now try searching your own archive (it's still free!)

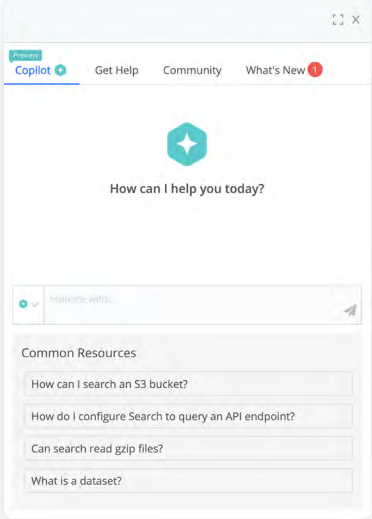
Go from test drive to searching your own data in minutes:

- Create a dataset to search in place the data from dozens of reachable resources, then populate your dataset.
- Or, create your own Cribl Lake dataset to store your data, then populate it with a few clicks
- Then, just click on the dataset to search it, or for more filtering and control follow along to build your search
- Finally, if fields/tables are not enough, visualize your results to your liking

NEXT STEPS

Want to know more about what you did… and could do?

- Learn more about the Search UI
- Need immediate answers? Get AI guidance by clicking on the ✦



How can I help you today?

Common Resources

How can I search an S3 bucket?

How do I configure Search to query an API endpoint?

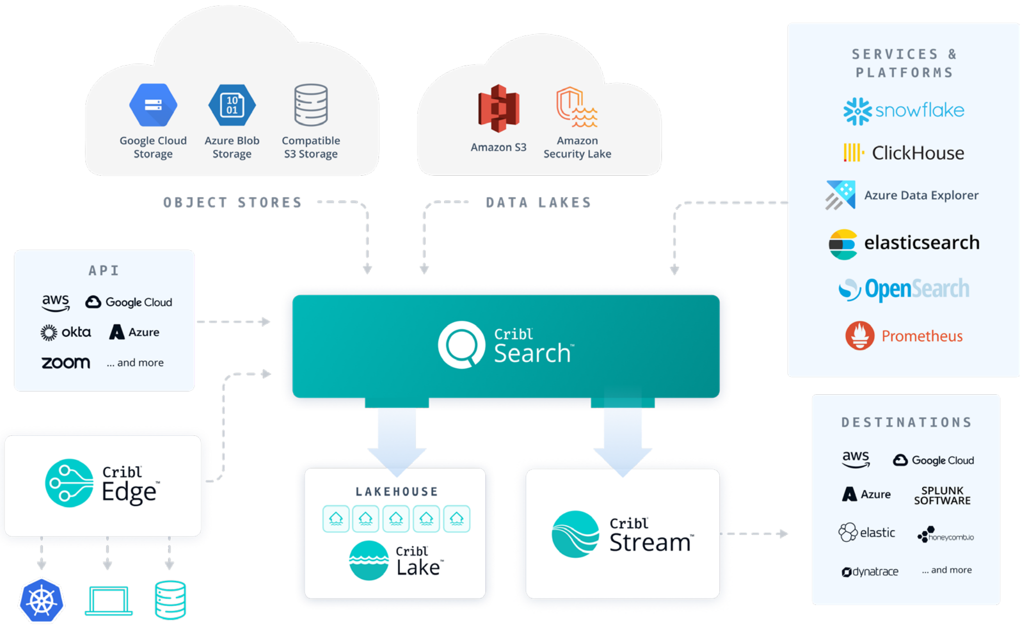
Can search read gzip files?

What is a dataset?

So, where can Cribl Search reach?

It reaches almost anywhere you have data you need to investigate. Cribl Search allows you to search, explore, and analyze telemetry data—logs, instrumentation data, application data, metrics, etc.—in place without first moving it to specialized storage. And if its already stored we can look there too. Search's accessible resources include the following types of data storage providers:

- **Data lakes**, including Amazon Security Lake, Amazon S3, and S3 compatible stores.
- **Object stores**, including the likes of Amazon S3, Azure Blob Storage, Google Cloud Storage, and more.
- **Analytics services and platforms**, like Azure Data Explorer, Elasticsearch, Opensearch, Prometheus, and more.
- **API endpoints**, including Azure, AWS, Google Workspace, Okta, Zoom, and even a Generic HTTP API option allowing you to query any HTTP API.
- **Cribl family**: Cribl Edge, Cribl Lake, and Lakehouse



Cribl Search is a key component of a modern data strategy

Rethink your archival strategy

- Use low-cost storage without giving up visibility and instant retrieval—no waiting for data to “thaw.”
- Optimize systems of analysis processing—ingest only actionable data.

Investigate data where it lives

- Data lakes, object stores, API endpoints, host systems, and more
- Search puts your data at your service.
- Optimize query performance while minimizing system of analysis costs.

Route results to any destination

- Search data where it is, retrieve only what you need, store data where you want.
- Route results to optimal processing, never compromise between cost and value.

Ultra-fast access to Cribl Lake data

- Effortlessly store and search massive volumes of ever-changing telemetry data.
- Lakehouse enables real-time, high-performance searching of the most recent data.

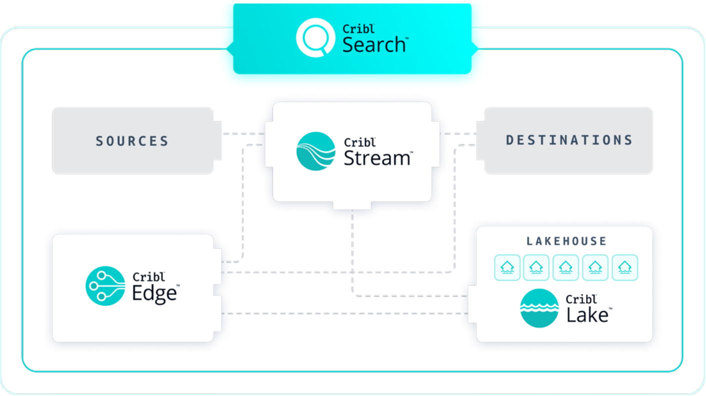
Search any and all telemetry easily

- Never have to settle for sampled data due to license costs.
- Full-fidelity searching of any data type, storage type, application, or system.

Learn more about Cribl Search

- Existing Cribl.Cloud user? Getting Started with Search
- New to Cribl.Cloud? Go to try cloud and get started
- Watch the Search Video
- Learn more about Cribl Search
- Have questions? Get support from the Cribl Community
- Learn more about Cribl with free training at Cribl University

Conclusion



Cribl isn't just Search. The Cribl portfolio enhances your data management journey with tools that give you visibility and control across all your telemetry data. Cribl provides choice, control, and flexibility to ensure you manage your data efficiently and cost-effectively. Modernizing collection, routing, storing, and analysis improves visibility, while modernizing integration and transformation enhances orchestration between different platforms and providers. Additionally, upgrading storage and automation streamlines and optimizes data handling, and modernizing discovery and analysis offers better tactical value by making it easier to find necessary data. This iterative modernization process gradually increases your organization's data maturity.

Cribl Search success story

92% **Reduction in MTTD**

(24 hrs to 1.5 hrs)

"Incident Response team needed to recover a specific dataset, existing vendor said 24hrs to rehydrate all 26 TB of data. Cribl Search was able to query data in storage, retrieving only the specific dataset (50k events) and then forwarding only useful data to their existing systems of analysis for deeper investigation."

– CUSTOMER, MULTINATIONAL NEWS & MEDIA CORPORATION

Cribl®

Cribl, the Data Engine for IT and Security, empowers organizations to transform their data strategy. Powered by a data processing engine purpose-built for IT and Security, Cribl's product suite is a vendor-agnostic data management solution capable of collecting data from any source, processing billions of events per second, automatically routing data for optimized storage, and analyzing any data, at any time, in any location. With Cribl, IT and Security teams have the choice, control, and flexibility required to adapt to their ever-changing data needs. Cribl's offerings — Stream, Edge, Search, and Lake— are available either as discrete products or as a holistic solution.

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

EB-0012-EN-1-0525