Cribl®

Solution Guide

GUIDE

Cribl Solution Guide

April 20, 2026 | Download the PDF

Document Purpose

This guide provides a holistic overview of how Cribl solutions can help your organization unlock the value of data, ensuring it's not just abundant but actionable and impactful. Read on for an overview of the solution's features, key capabilities, benefits, and deployment options.

Executive summary

Most teams are drowning in data but starved for insight. You're trying to keep up with exploding data volumes, rising cyberthreats, shrinking budgets, and limited headcount. That means IT and Security teams end up making tradeoffs about what data they can afford to collect, store, and analyze.

You shouldn't have to choose. You should be able to use all your data, at any scale, and deliver it in any format to any tool that needs it.

Cribl's suite of products is built for exactly that. Cribl gives you a vendor-agnostic, flexible data engine so you can route, shape, store, and search data on your terms. With Cribl, teams get the control, flexibility, and efficiency they need to make data work for them… and for their customers.

Solution overview

Introduction

The Cribl suite is built around a single idea: give you control over your data landscape.

You decide what's best for your organization. You control where data goes. You choose the format, on the fly, without replacing agents or locking into a single tool.

- **Cribl Stream** optimizes data in motion, so you can route, enrich, and transform data efficiently.
- **Cribl Edge** pushes those capabilities out to the edge, closer to where data is created, to cut latency and network cost.
- **Cribl Search** lets you search data in place, across live and stored data, without first moving or rehydrating it.
- **Cribl Lake** makes it easy to store, manage, and access data in open formats so every team and tool that needs it can use it.

Together, these products give you an end-to-end data engine with real choice and control, while preserving your existing investments.

Solution architecture

Cribl's architecture is designed to meet you where your data lives—on endpoints, in transit, in existing tools, and in object storage.

- Cribl Stream and Cribl Edge collect, process, and deliver observability, security, and telemetry data in real time to any destination.
- Use built-in and ad hoc collection to pull data from your data lakes into your analytics tools when you actually need it.
- Cribl Search runs search-in-place queries against data at the edge, in flight, in Cribl Lake, or in your existing systems, so you widen your analysis without duplicating data.
- Cribl Lake acts as a central hub for open-format storage. You can share and re-route data downstream through Stream and Edge, in any format, to any tool, at any time.
- Search unifies the query experience no matter where data is stored, so you don't have to move or rehydrate data just to analyze it.

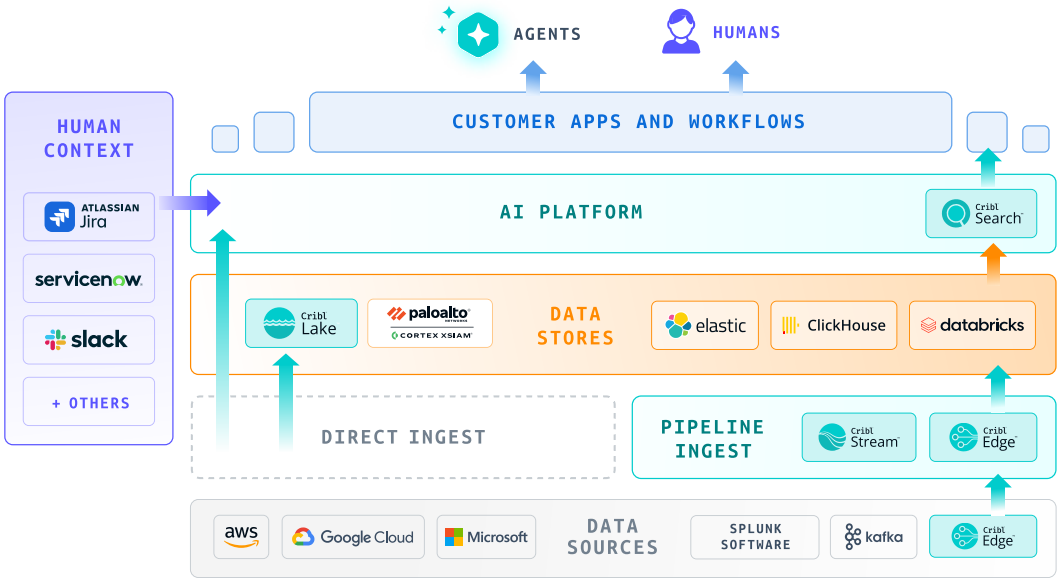

Fig. 01: High-level solution architecture.

Solution benefits

- **Flexibility to eliminate tool lock-in:** Choose the best vendor tool(s) without requiring new agents or collectors.
- **Simplify observability data engineering:** Easiest way to discover, explore, collect, enrich, and route data from anywhere to anywhere.
- **Complete control over your data:** Route data where it has the most value. Control access to sensitive data. From core to the edge.
- **Observe more:** Build and automate tiered storage. Search data in place, only forward useful data tailored to your needs. Access data previously inaccessible.
- **Stretch your investment:** Reduce management costs and complexity of exploring, collecting, processing, and accessing data at scale.

Product overview

Cribl Stream

Functionality

Cribl Stream is an observability pipeline that lets you route, shape, restructure, and enrich data from any source to any destination without adding new agents. You gain control over your data and simplify IT and Security workflows so you can instrument more, analyze more, and pay less.

Customer needs addressed

Stream gives you control over data in motion. You can:

- Intelligently route, reduce, and enrich data.
- Make noisy, high-volume streams manageable and insightful.
- Cut waste while keeping what matters for security, operations, and business analytics.

You end up with data that's not just big—it's useful.

Cribl Edge

Functionality

Cribl Edge is a highly scalable, edge-based data collection system for logs, metrics, and application data. It lets administrators reliably collect and process this data in real time from Windows and Linux machines, apps, and microservices, and send it to any supported destination.

Customer needs addressed

Edge brings processing closer to where data is born. That means:

- Lower latency and smarter use of network resources.
- Data can be reduced, transformed, and routed at the edge before it travels.
- You can scale collection as your environment grows, with a centralized management view for insight and efficiency.

You get edge autonomy with central control.

Functionality

Cribl Search is a vendor-agnostic analytics tool that lets teams run search-in-place queries across data wherever it is — at the edge, in flight, in an observability lake, or inside existing systems.

Customer needs addressed

Search reshapes how you explore data:

- Query historical and real-time data without moving or copying it first.
- Run fast investigations on data in S3, data lakes, TSDBs, log stores, and more.
- Make decisions with a current, complete view of your data, not yesterday's subset.

This "analyze before you ingest" model reduces risk and cost.

Functionality

Cribl Lake is a central hub for storing, managing, and accessing large volumes of data from many sources. It's designed to integrate tightly with Cribl Stream and Edge so you can ingest and route data downstream with minimal friction.

With Lake and Search together, you can:

- Query data at rest in Cribl Lake.
- Search across other data lakes, object stores, search APIs, and analytics platforms.
- Use federated search to query multiple stores at once, enabling quick investigations and rapid time to value.

Customer needs addressed

Lake solves for long-term, cost-effective data retention and analysis:

- Keep data in open formats without constantly moving or rehydrating it.
- Support historical analysis, compliance, and security investigations with one consistent storage strategy.
- Provide a unified platform for ingestion, processing, and long-term storage so insights are never stranded.

You get durable, accessible data that keeps its value over time.

How it all works together

Cribl Stream and Edge give you powerful processing and transformation options inside a common architecture:

- **Sources** collect data and hand it off to Destinations through **QuickConnect** or **Routes**, which send data through Pipelines for processing.
- From there, you can easily land any data in **Cribl Lake** for low-cost storage and future analysis.
- **Cribl Search** lets you transform and analyze log events in place, across Cribl Lake and any other supported destination.

Fig. 02: Cribl Stream basic solution architecture.

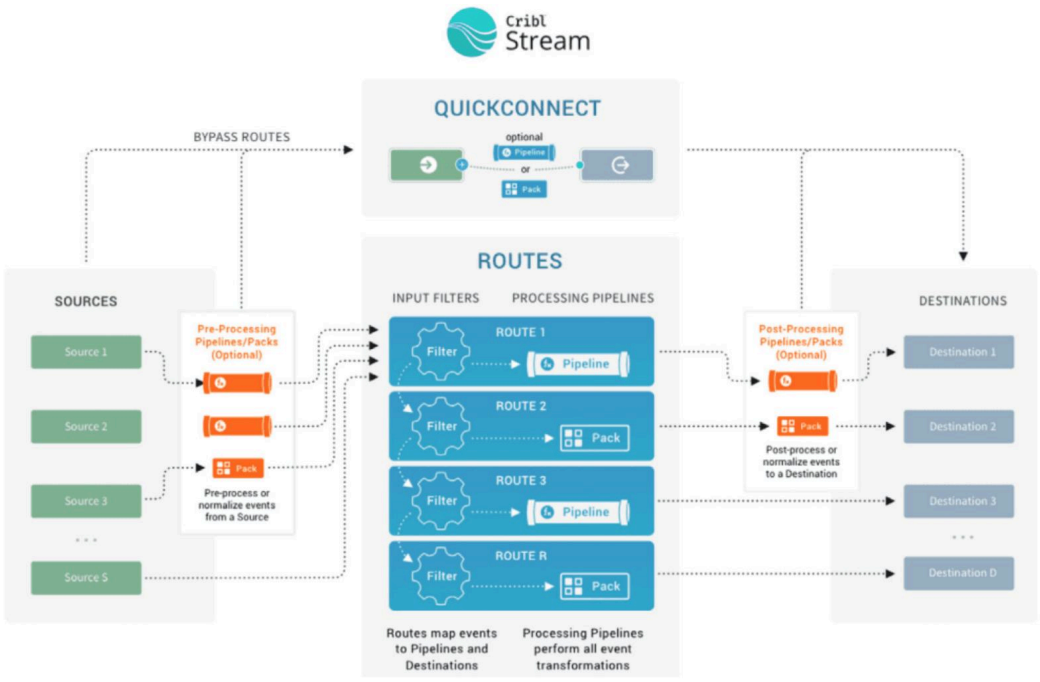

Fig. 03: Closer look at a single Stream Worker (instance).

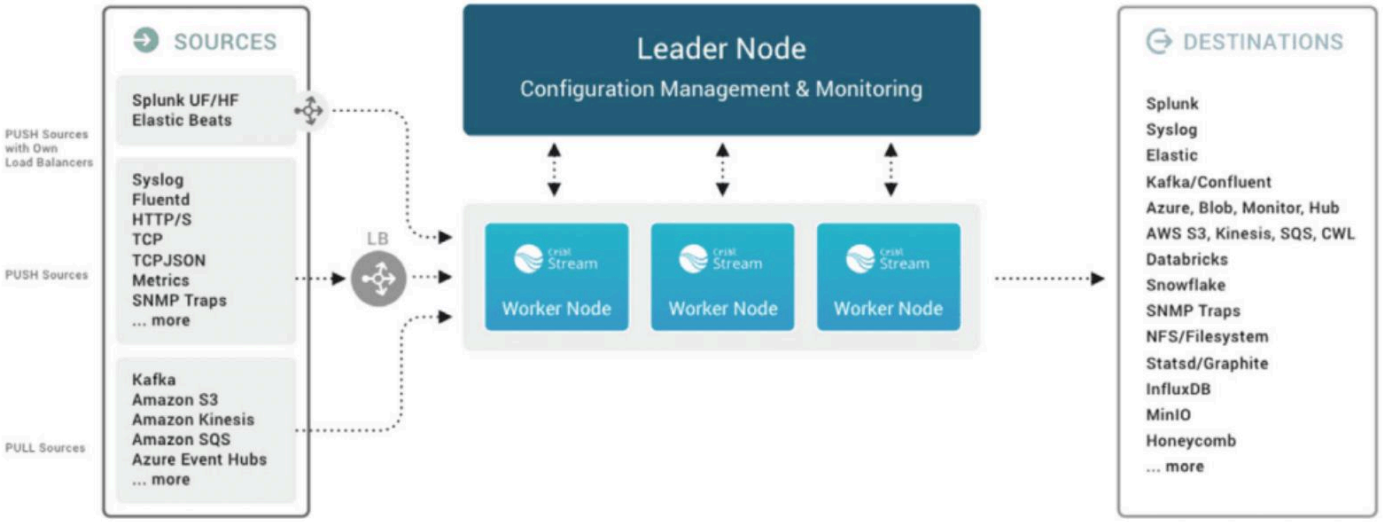

Fig. 04: Cribl Stream scales up to meet enterprise needs in a distributed deployment.

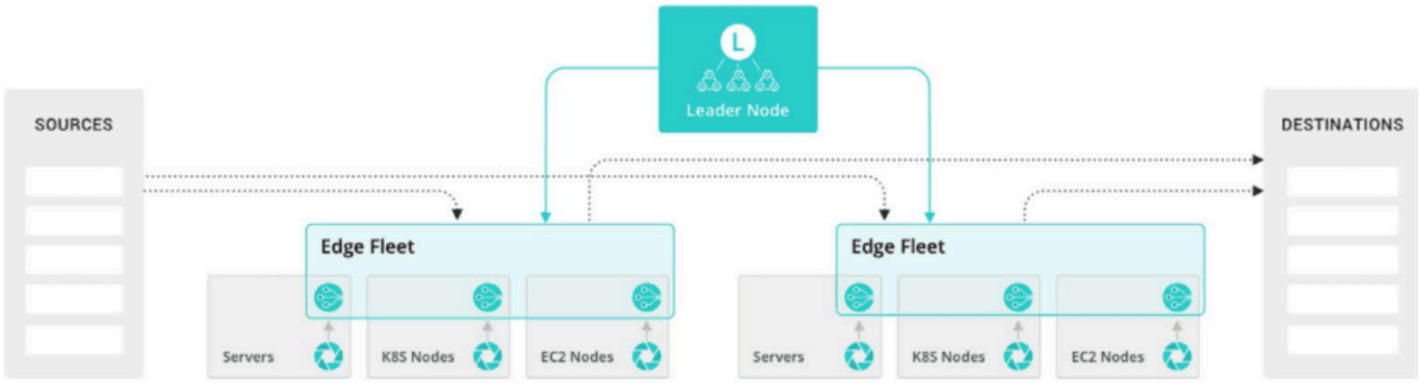

Fig. 05: Cribl Edge basic solution architecture. Edge processing, management, and receivers.

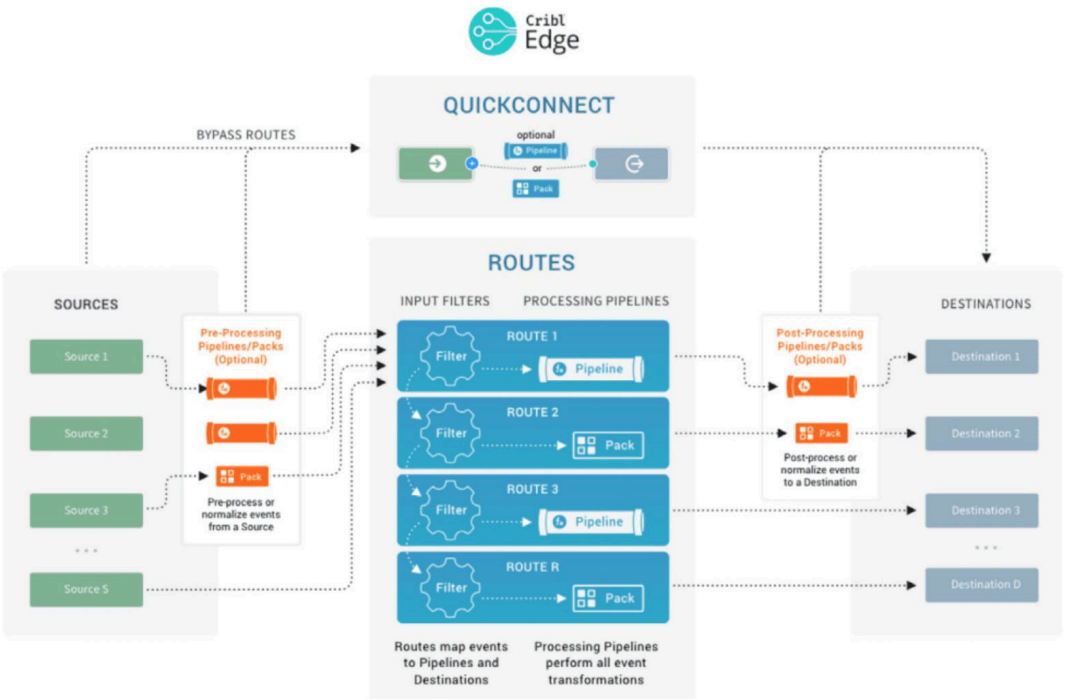

Fig. 06: Double-click into a single Edge node.

Key capabilities

Explore data first, then collect

Most tools force you to collect everything before you know what's useful. Cribl flips that.

With a unique **"Explore First, Collect Second"** approach, you can:

- Inspect logs, metrics, and application data at their egress points using Edge and Stream.
- Dive into data before you commit to collecting it.
- Collect at massive scale from any source, and even schedule batch collection from multiple APIs.
- Use ad hoc collection to recall data from your lake or storage, and replay it into your analytics tools.

You only collect what earns its keep.

Cost-effective routing and processing wherever you need it

Cribl's architecture pushes processing to the right place in your environment:

With a unique "Explore First, Collect Second" approach, you can:

- **Cribl Edge** performs data reduction, transformation, and routing at the edge, cutting bandwidth and storage costs.
- You can send data from Edge to **Cribl Stream** for deeper processing, or forward it straight to your tool of choice.

You design the flow that best fits your infrastructure and budget.

More data value through long-term retention and analysis

Cribl Lake plays a critical role in the lifecycle of your data:

- It gives you a scalable, secure space to keep large volumes of data for the long haul—perfect for historical analysis, regulatory compliance, and security investigations.
- Lake works hand-in-hand with **Stream** and **Edge** so data moves smoothly from creation to long-term storage.
- You can run time-series analysis, retrospective audits, and advanced ML models directly on retained data.

Lake extends the life and value of your data while keeping storage and retrieval simple.

A “Search Then Forward” approach

Cribl Search introduces a **“Search Then Forward”** pattern:

- Run federated, search-in-place queries across multiple data types and stores—edge nodes, pipelines, data lakes, TSDBs, log stores, and more.
- Validate exactly which data you need before you collect or move it.
- Eliminate guesswork by querying data where it lives, then selectively forwarding only what’s useful.

This reduces risk, avoids unnecessary ingestion, and keeps you in control of costs.

More choice, less complexity

Across the suite, Cribl gives you:

- **Data choice and flexibility** so you can route data where it has the most value and easily support multi-tool or consolidation projects.
- **End-to-end control** over collection, routing, processing, storage, and analysis, including the ability to sanitize or restrict sensitive data in flight.
- **Lower risk and better compliance** through consistent controls and search-in-place capabilities.
- **Better ROI** on licenses, hardware, storage, and people by making every byte and every tool work harder.

Solution deployment

Before you deploy: key considerations

You'll choose a deployment model based on how much data you handle and how you plan to process it. To narrow it down, think about:

- **Amount of data ingest**
  How much data do you plan to ingest per time unit—MB, GB, or TB per day?
- **Amount of data at the endpoint**
  How much of that data will you collect and process at endpoints (edge) per time unit?
- **Amount of data processing**
  How heavy is the processing on incoming data—lots of transformations, regex extraction, parsing, masking, or obfuscation?
- **Routing and cloning**
  Is data mostly going to a single destination, or will you be cloning and routing it to many?
- **Connectivity constraints**
  Do you need to support air-gapped, on-prem servers with no internet access?

Once you've answered these, it's much easier to pick the right deployment pattern.

Deploying the Cribl suite

This section covers how to set up Stream, Edge, Search, and Lake in a way that fits your environment—from simple to large distributed deployments.

Common deployment types:

- **Single-instance deployment:** Use this when incoming data volume is low and processing is light. For details, see the Single-Instance/Basic Deployment and the Cribl Stream Getting Started Guide.
- **Distributed deployment:** Use this to support higher loads or more complex processing. See Distributed Deployment for setup, Sizing and Scaling for capacity planning, and Bootstrap Workers from Leader to automate Worker rollout.
- **Cribl.Cloud deployment:** Use Cribl.Cloud to quickly launch a Cribl-hosted deployment of Stream, Edge, and Search. Cribl manages the infrastructure for you. See Cribl.Cloud for details and comparison with self-hosted options.

In all cases, you can evolve from simple to hybrid models over time without re-architecting everything.

Setting up Cribl Lake for long-term data management

Cribl Lake is designed for large-scale, long-term IT and security data retention. To integrate Lake into your environment:

1 ## Direct data to Cribl Lake

Configure Cribl Stream or Cribl Edge to route the right data streams into Cribl Lake, in open formats. This is key to capturing the data you'll want later.

2 ## Activate collection from Cribl systems

Cribl Lake can automatically collect operational data from your Cribl deployments to help you monitor health and performance. To enable this in Cribl.Cloud:

- Go to your Cribl.Cloud Organization homepage.
- In the Cribl Lake section, turn on internal data storage.

3 ## Organize data with datasets

Cribl Lake uses datasets to structure stored data for easier access and management.

To create a custom dataset:

- In the Cribl Lake UI, add a new dataset.
- Give it a clear name and description.
- Save it so it's ready for use in routing and search.

4 ## Replay data via Cribl Lake Collector

When you need to reprocess stored data, configure the Cribl Lake Collector in Cribl Stream. This lets you pull historical data back into your pipelines and tools.

5 ## Leverage Cribl.Cloud

Cribl Lake is built for the cloud and is a core part of Cribl.Cloud, giving you scalable, secure storage without the operational burden of traditional setups.

Adding Lake to your deployment ensures valuable data is both retained and easy to act on later.

Recommended solution deployment option: Cribl.Cloud

The fastest way to get going is Cribl.Cloud.

Cribl.Cloud spins up all Cribl products—Stream, Edge, Search, and Lake—in just a few minutes. In this model:

- The Leader and Worker Nodes run in Cribl.Cloud.
- Cribl manages the underlying infrastructure for you.
- You can later extend into any hybrid deployment, with any mix of cloud and on-prem Workers and Edge nodes.

To get started, use the Cribl Stream Cloud deployment docs and sign up through the Cribl.Cloud portal.

Summary

Running an effective data engine is getting harder every year. Data volumes skyrocket. Licenses and infrastructure cost more. Retention windows get longer. At the same time, security requirements tighten and business needs change constantly.

Tool sprawl and proprietary formats make it hard to move data between systems or share across teams. You end up with islands of data, partial visibility, and expensive blind spots.

Cribl's data engine changes that.

By adding a flexible processing pipeline, powerful search-in-place, and open storage to your stack, you regain choice and control over all your data—without discarding the tools you already use.

Cribl's end-to-end observability solution brings together:

- Data collection with Cribl Edge and Stream.
- Data routing and processing in Stream and at the edge.
- Data storage and retention with Cribl Lake.
- Data search and analysis with Cribl Search.

You get better visibility and control while reducing cost and complexity. You can deploy in the cloud, on-prem, or in any hybrid pattern.

Specifically, the Cribl solution:

- **Uses Edge and Stream to "Explore First, Collect Second"**
  Explore logs, metrics, and application data at egress points. Decide what to collect and how to process it before it hits your expensive tools.

- **Enables reduction, transformation, and routing at the edge**
  Use Cribl Edge to cut noise and cost before data goes anywhere. Send it to Cribl Stream for deeper processing or directly to your target systems.

- **Simplifies full lifecycle data management with Cribl Lake**
  Keep data safe, open, and accessible for long-term retention. Lake works tightly with Stream and Edge so you can explore, collect, retain, and analyze with one consistent approach.

- **Eliminates risk and guesswork with search-in-place**
  Use Cribl Search to run federated queries across multiple data types and stores—Cribl Lake, major object stores, time-series databases, and log platforms — before you decide what to store and where.

The fastest way to get started with Cribl is in the cloud! Cribl.Cloud platform can spin up Cribl Stream, Edge, Search, and Lake in just a few minutes. To get started, check out our Launch Guide and sign up on the Cribl.Cloud portal.

Additional Resources:

- Cribl Documentation
- Getting Started Guide
- Distributed Quick Start
- Launch Guide
- Cribl Community
- Cribl Support
- Cribl University
- Cribl Curious
- Self-Guided Trials
- Cribl Github Repos
- Docker Hub

Cribl®

The **AI Platform** for Telemetry

Learn more at cribl.io | Join our Slack community

Try Cribl Sandboxes | Follow us on LinkedIn and X

Cribl, the AI Platform for Telemetry, empowers enterprises to manage and analyze telemetry for both humans and agents. Trusted by organizations worldwide, including half of the Fortune 100, Cribl bridges the gap between AI ambition and infrastructure reality. No lock-in. No data loss. No compromises. Cribl's vendor-agnostic platform ensures data remains portable and interoperable. By cost-effectively handling increasing data volume and variety without delay, Cribl gives enterprises the choice, control, and flexibility to build what's next. Founded in 2018, Cribl is a remote-first workforce with an office in San Francisco, CA.

SGDE-0001-EN-1-0526