Cribl Stream™

Cribl Stream is an observability and data streaming platform for real-time processing of logs, metrics, traces, and O11y data that enables the ITops/SRE/SecOps/O11y teams to collect the data they want, shape the data in the formats they need, route the data wherever they want it to go, and replay data on-demand; thereby enabling customers to observe more and spend less, to have choice and flexibility, and to provide control over their data.

CHOICE AND FLEXIBILITY

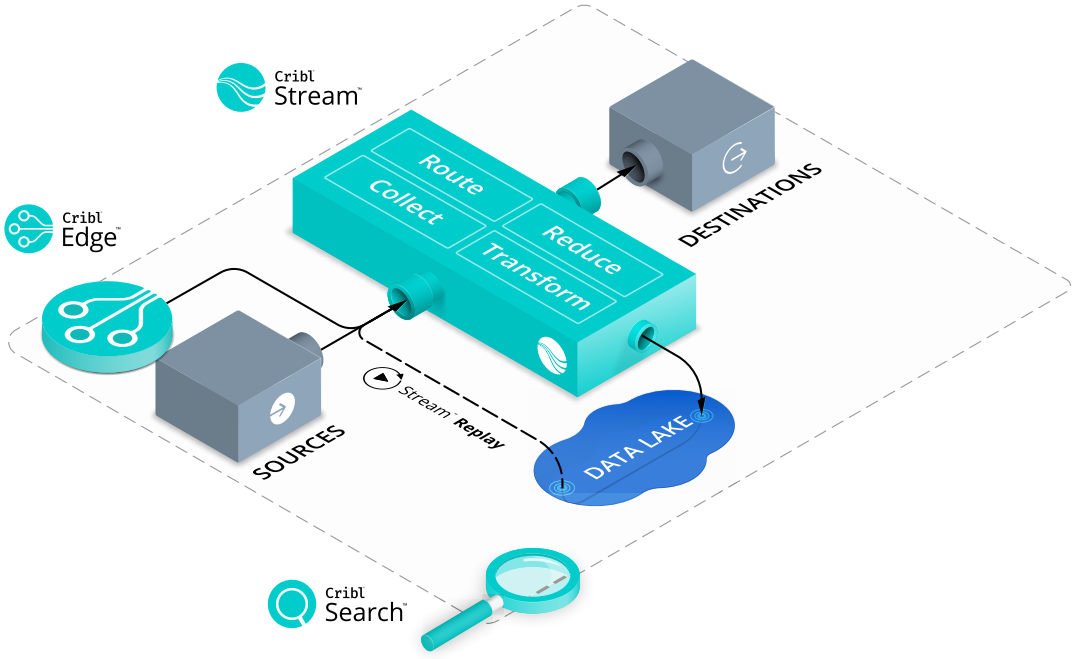
- Stream events, metrics, and traces from any source to any destination, in any format.
- Add new analytics tools and other destinations without adding agents or collectors.

CONTROL YOUR DATA

- Route data to the right teams and the right tools.
- Deploy enterprise-grade security.
- Enforce data policies, standards, and formats across toolsets.

OBSERVE MORE AND SPEND LESS

- Streamline data onboarding and collection to expose data you didn't even know existed.
- Optimize infrastructure, data ingestion, tool performance, and people hours.



COLLECT THE RIGHT DATA
Get data from any source to any tool in the right format

REDUCE YOUR DATA
Eliminate uninteresting data to control costs

SHAPE YOUR DATA
Take data as it comes, shape it into what you need

ROUTE YOUR DATA
Put data where it has the most value

REPLAY YOUR DATA
Keep high-fidelity data on low-cost S3 to replay on-demand

Product Features

ARCHITECTURE

- *Single binary distribution with zero dependencies*
- *Shared-nothing, super scalable distributed architecture*
- *Scales from laptop to 100s of nodes and 10000s of cores*
- *Highly parallelizable, highly performant platform built for extensibility*
- *Sub-millisecond latency*
- *Tested to upwards of 20PB/day*
- *Deployment Options include:*
  - *SW including linux binaries, docker containers and helm charts for easy deployment in any K8s environment*
  - *Cloud provides Saas experience through Cribl.Cloud, entirely Cribl managed, no infrastructure overhead and scales as needed*
  - *Hybrid-Leader/Control plane in the cloud and workers performing local processing Integrations*
- *Over 80 source/destination integrations available out of the box*
- *Native protocol support for leading sources and destinations of logs, metrics, and traces*
- *Out-of-the-box TLS support for all integrations that support it*
- *Out of the box support for SAML 2.0 supported IAM and Assume roles*
- *Live data capture for integration for troubleshooting and inspection*
- *Rich logging, metrics and real-time status for each integration*
- *Baked-in connectivity tests & results for each integration*
- *Support for arbitrary REST endpoint data collection*
- *Support for arbitrary Script based data collection*
- *Support for sending and collecting from all major Cloud PaaS storage services*

MANAGEMENT

- *Full control of all your observability data from a central control plane*
- *High-availability architecture offering seamless Leader(Control plane) failover near real-time fault-tolerance on-premises and 99.9% availability on Cribl.Cloud*
- *AuthZ support enhances security by giving control over who has permissions and privileges to access Cribl products, capabilities, and resources.*
- *Stream Projects allows Cribl users to* **securely** *access data throughself-service model, freeing up Cribl admin time to work on business-critical tasks.*
- *Enterprise grade authentication support (LDAP, SSO etc)*
- *Policy-based RBAC for fine-grained permissioning*
- *Intuitive, rich user interface for distributed system management*
- *Single, centralized management via cloud or self-hosted software for 100s of groups/nodes*
- *Dependable configuration version control with ability to revert changes*
- *Built-in, real-time configuration change validation*
- *Centralized support for certificate and key management*
- *Built-in data generators for pipeline and destination testing*
- *Fully automated and distributed upgrades of all Stream workers*
- *Leverage external Key Management Services for managing secrets/tokens across all nodes*
- *Built-in synchronization with external code repositories for CI/CD integrations and disaster recovery*

MONITORING

- *Notification system alerts operators when data flows have stopped*
- *Notification system alerts operators when there is too much variation in data volume from sources/collectors to identify upstream issues*
- *Built-in Monitoring covering all aspects of a distributed deployment*
- *Built-in centralized log search across 100s of groups/nodes*
- *Rich, visually dense, dashboards built for admins/operators*
- *Contextual monitoring for all sources and destinations*
- *Ability to forward full-fidelity internal logs & metrics to external solutions*
- *Dataflow visualizations provide Birds Eye View of all sources, routes, pipelines and destinations*

WORKING WITH DATA

- *Interactive, user-friendly, efficient UI for working with streaming data*
- *Visual authoring, validation and troubleshooting of data pipelines*
- *Data Preview with instant feedback for visual inspection of events as they're being transformed*
- *Live capture on multiple points as events travel from source to destination*
- *Built-in documentation and contextual tooltips help on every screen*
- *Over 30 out of the box Functions that support arbitrary data transformations, securing and enrichment.*
- *Over 40 built-in C. function methods for finer processing capabilities*
- *… plus all the power of JavaScript for almost-arbitrary data transformations*
- *IDE-like experience with auto-complete and typeahead assist*
- *Automatic byte-stream to events conversion/breaking using intelligent rules with optional user overrides*
- *Automatic timestamp format recognition with optional user overrides*
- *Timezone recognition and/or correction*
- *Built-in JavaScript expression editor with live result preview*
- *Built-in Regex editor with live match & capturing group preview*
- *Built-in Regex Library for most common regex, extensible*
- *Out-of-the-box parsing support for many well known data sources*
- *User-defined data parsers for K=V, CSV, ELFF, CLF, JSON and delimiter based values*
- *Regex-based field extractions and native Grok pattern support*
- *Event schema validation support using JSON Schema standard*
- *Support for Global Variables - re-usable and composable JS expressions that can be referenced by any Function*
- *Granular analysis of pipelines for better monitoring and quicker troubleshooting in preview mode using Pipeline Profiling*
- *Real-time data enrichment via lookup tables. Exact, Regex and CIDR support out of the box*
- *Support for geoip enrichment using Maxmind binary databases*
- *Packs support for building, packaging and sharing routes, pipelines, data samples, functions internally or with members of the community*
- *Access to a growing community of Stream Packs with pre-built pipelines, custom functions and other out-of-the-box features for speeding up and improving the value of Stream*
- *Global search makes finding anything in Stream easy and fast*
- *Gather live data samples to aid in development of pipelines or to share with teammates working on similar projects.*

TECHNICAL REQUIREMENTS

System

- *+4 physical cores*
- *+8GB RAM*
- *5GB free disk space (more if PQ enabled)*
- *Also available as a SaaS solution through Cribl.Cloud*

Sizing Guidance

- *1 physical core for each 400GB/day of IN+OUT throughput*

  E.g., 4 TB IN -> full 4TB to Destination A, plus 2 TB to Destination B = 10TB total = 25 physical cores.