>WHITE PAPER_

Migrating log analytics to the cloud.

>WHITE PAPER_

Migrating log analytics to the cloud.

Everything is moving to the cloud. At least that's what the narrative in the industry is. A large shift from on-premises to cloud log management and analysis comes with plenty of upside, but challenges can quickly follow.

It isn't a silver bullet for all use cases. Teams need to be cognizant that the data is readily ingested and available in an economically viable way. Plus, cloud service egress charges can limit choice and control.

Cribl knows IT departments often manage change carefully during transitions to keep risks low. So why move to the cloud without putting strong protective measures in place? A company wouldn't let a third party connect directly into their network without some kind of firewall, so why would they move their data to the cloud without similar considerations?

> A company wouldn't let a third party connect directly into their network without some kind of firewall, so why would they move their data to the cloud without similar considerations?

The challenges.

Costs.

Cloud-based log systems take infrastructure responsibility and the associated costs off the table, but the potential cost of sending high-volume, low-value data to them can escalate costs quickly. These can include ingestion costs, increased network bandwidth requirements, and possible network-data egress fees. That's all before even mentioning security and compliance mandates that can require years of retention. Add these up, and it's critical that enterprises get the most value possible out of every byte they send to log system(s).

Logging systems have organically become both analysis systems and data retention systems, to the detriment of both purposes. Many organizations say most of the data in their logging systems is never analyzed or acted on, and is only there to satisfy their retention requirements.

For example, in an AWS instance that's sending 5TB/day to a log analytics system, an organization would be paying $.09 per GB making it $450/day x 290 = $132,000 in egress charges per year, according to the AWS Pricing Calculator. That's just the cost of moving the data, not including compute or storage. Of course, most enterprises will reduce that cost through discounts or other avenues, but it remains a cost of moving log analytics to a cloud environment.

Flexibility limitations.

When sending data to a cloud environment, control of that data is sacrificed. Redirecting Elastic Beats Agents or Splunk Universal Forwarders, for example, to feed into a cloud environment is simple. What is complicated is doing anything with that data — like enrichment, shaping, or routing to other tools. For example, if the same data that's going to the cloud log provider also needs to be sent to a data lake, there needs to be a way to facilitate that delivery.

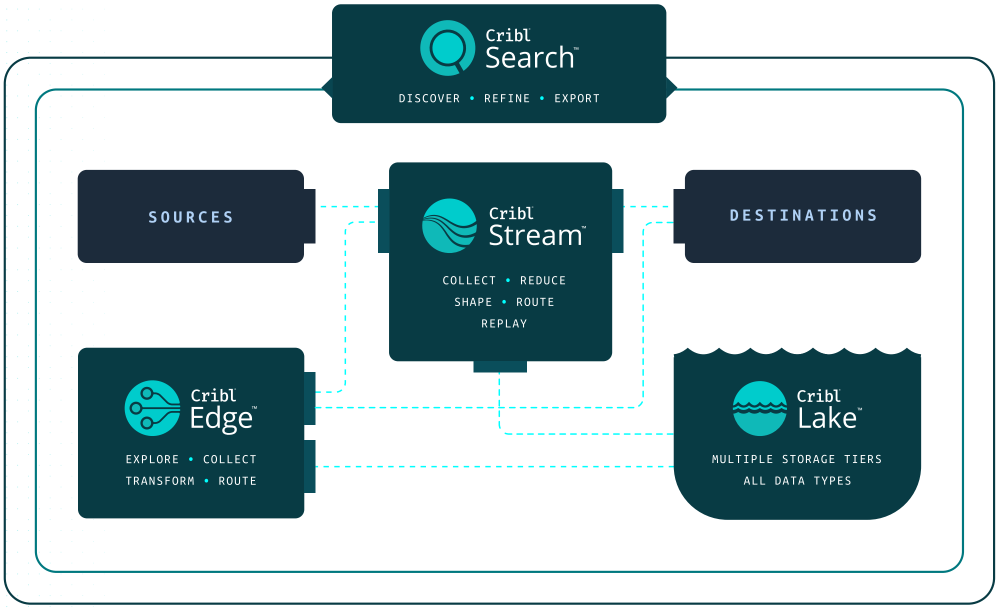
Vendor lock-in.

In theory, enterprises no longer need to maintain deep expertise in day-to-day running of the migrated system once moving to a cloud environment. But, systems can become a black box that controls the data if they depend on the vendor for both data operations and stewardship of the data without the ability to adapt or change as the data evolves and needs change. This can lead to incremental costs of legacy architecture that teams can't get out of and can snowball out of control.

The solution: A Data Engine for IT and Security.

Cribl's next-generation data engine is built specifically for IT and Security data and provides a unified data management solution for exploring, collecting, processing, and accessing observability data at scale. With complete control and flexibility to access, explore, discover, collect, and process data at scale, enterprises can experience the difference between an alright analysis experience to a great analysis experience.

- **Stream –** A data collection, reduction, enrichment, and routing system for IT and security data. Retain a full-fidelity data copy in a lower-cost cloud storage and replay/analyze that data when needed.
- **Edge –** An intelligent, scalable, edge-based data collection system for logs, metrics, and application data. Enable administrators to easily provision datasets to their security and operation teams.
- **Search –** Find and access data regardless of where it is landed and in any format. Perform federated search-in-place queries on any data, in any form without having to move it or index it first.



Cribl Stream™

Retain a full-fidelity data copy in a lower-cost cloud storage, but still be able to replay and analyze it at will.

Cribl Edge™

Enable administrators to easily provision datasets to their security and operation teams.

Cribl Search™

Perform federated search-in-place queries on any data, in any form.

Managing costs.

Revisiting the cost example addressed earlier — with Cribl, take raw events, reduce them, and then compress them up to a 10:1 compression rate. By optimizing the data and allowing analysis and data retention systems to be used as they were intended, teams are able to maintain the integrity of their data within budget.

For example, cloud services, such as Amazon S3 and Microsoft Sentinel, create opportunities to retain data for cents on the dollar compared to traditional storage or datastore technologies. When a security breach occurs, teams can quickly investigate by either replaying their full-fidelity copy of that data or search it in place.

Creating flexibility.

Select and employ different software tools without requiring the installation of new agents or software components. Log data holds a lot of power, but it is rarely pretty — it's largely unstructured or semistructured, and is not very valuable without context or cleaning.

Plan ahead to mitigate challenges like new costs, flexibility limitations, and vendor lock-in.

For example, a switch reporting in its log that a port is flapping doesn't really help anyone — not unless they can also understand what server is connected to that port, what application is running on that server, and what business process that application supports.

If that contextual data is available alongside the log entry, it helps an analyst decide how much priority to put on the error. Moreover, that same set of "switch port flapping" log entries uses a lot of space to identify a simple problem: The single log entry is not relevant on its own, but a collection of, say, 50 of them over a short period indicates a real problem – at the cost of ingesting and storage 50 entries.

Reducing that collection to a single item that summarizes the problem is far more efficient. Reshaping and enriching data in the pipeline increases the value of the data being sent to the analytics system, while potentially reducing the cost of sending it.

Fighting vendor lock-in.

The freedom to select the best tools for the job is a lot more difficult to come by then one would think. But, with Cribl suite of products, organizations get to decide where they want to route data, process that data in flight, or search it in the most optimized manner. Export large volumes of data with minimal impact to existing logging solutions. Then, fork open-format data off to low-cost storage to abide by retention requirements, freeing up ingest volume for additional data sources.

While an organization might have bought Splunk, a couple of teams may want to use Elasticsearch or Grafana as their analysis tool. Without a control point in front of the log analytics systems, supporting multiple tools gets complex and difficult. But, with Cribl they're able to have different teams use different tools that all feed into a unified system of analysis.

Putting it all together.

The migration of log analytics from on-prem to the cloud needs to be handled with care. In Spider-Man they say, “With great power comes great responsibility.” The sentiment remains true when transitioning to the cloud. By being cognizant of new costs, limitations on flexibility, and vendor lock-in, organizations are setting themselves up for success.

Implementing a unified data management solution in front of a cloud log analytics environment allows the use of the analysis tools more effectively. Shape, enrich, and transform data to provide analysts with the context they need to derive insights and get more value out of their data. The Cribl suite of products provide out-of-the-box capabilities and gives users the superpowers they need to thrive in a cloud environment.

Organizations get the right data, where they want, in the formats they need. Instrument everything, analyze more data, and pay less. Query the data wherever it lives, and finally go from petabytes to insights.

Take this Fortune 50 Professional Services Firm who were able to scale back on data costs while accelerating their migration to the cloud. The enterprise cut their 5-6TB / day Windows Event Log volume by 20%.

See Cribl in action by exploring Cribl Sandboxes or creating a free Cribl.Cloud account.

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, and Cribl Search, the industry’s first search-in-place solution. 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

WP-0012-EN-1-0224