SACR

NOV 2026

# THE FUTURE OF SIEM, PART II:

## The Five Emerging SIEM Architectures

SEAN SOSNOWSKI

UCHE OBIORA




Cribl

# We explore the newest frontiers of cybersecurity.

Whether you're looking at emerging vendors, evolving threats, or shifting architectures, our timely, opinionated insights help modern security leaders make smarter, faster decisions.

## About Us

Software Analyst Cyber Research (SACR) is a modern research and advisory firm built for today's cybersecurity leaders. We deliver in-depth, timely analysis across SOC operations, Identity, Network, Cloud, Application Security, Data, and AI Security; equipping CISOs, security teams, founders, investors, and practitioners with the insight they need to navigate high-stakes decisions.

With an engaged community of over **80,000 readers and followers**, SACR connects with a global network of cybersecurity decision-makers and innovators. Our access to leaders across categories and industries gives us a direct line to the conversations shaping the market. By pairing these insights with rigorous technical analysis and continuous market tracking, we produce research that is both data-driven and grounded in the realities of modern security operations.

Whether you're seeking clarity on emerging technologies, evaluating vendors, or tracking market shifts, SACR delivers trusted, independent research designed to help you see clearly and decide with confidence.

## The Authors

**Sean Sosnowski** serves as the Research Director for Security Operations and Cloud Security at SACR, where he leads research on SOC strategy and operations, detection engineering, and the evolving role of automation and agentic AI in security workflows. Drawing on a decade of intelligence experience in the U.S. Marine Corps he has authored several analytic reports on emerging technologies and threats regarding sensitive national security and military operations.

**Uche Obiora** is a cybersecurity analyst and AI security researcher at Software Analyst Cyber Research (SACR), focusing on SIEM modernization, threat detection, and AI-driven security operations. He brings hands-on experience in SIEM administration, vulnerability assessment, and incident response across financial services and nonprofit organizations. A CompTIA CySA+ certified analyst, Uche combines practical security experience with published research on quantum computing threats to help security leaders understand emerging risks and make informed decisions.

# TABLE OF CONTENTS


Executive Summary ..... 3

The MSAM Framework ..... 6

Converged SIEM/XDR Model ..... 8

Observability-Connected Model ..... 12

Pipeline-First Model ..... 14

Data Platform Model ..... 17

Federated Model ..... 20

Practical Recommendations ..... 23

Conclusion ..... 26

Cribl ..... 27

# Executive Summary

Most security information and event management (SIEM) products now list the same capabilities: search, detection-as-code, AI-assisted investigation, case management, and long-term retention. Across the 19 vendors assessed in Part II, the differences that change an evaluation sit underneath that checklist. They come down to whether the storage tier decides which logs can trigger detections, whether third-party telemetry gets the same detections and response actions as first-party data, what a federated vendor copies to make queries fast, how much detection content, case management, and response ships with a customer-controlled data platform, and how the platform meters, scopes, and grounds the agents that query it.

*Part I* made the case that the next SIEM decision is an architecture decision and introduced SACR's Modern SIEM Architecture Models (MSAM) framework, which groups the market by where the primary control point sits. Part II applies MSAM at the vendor level, compares how vendors in each model handle storage tiers, detection placement, third-party telemetry, and agent access, and sets out proof-of-concept (POC) tests that separate them.

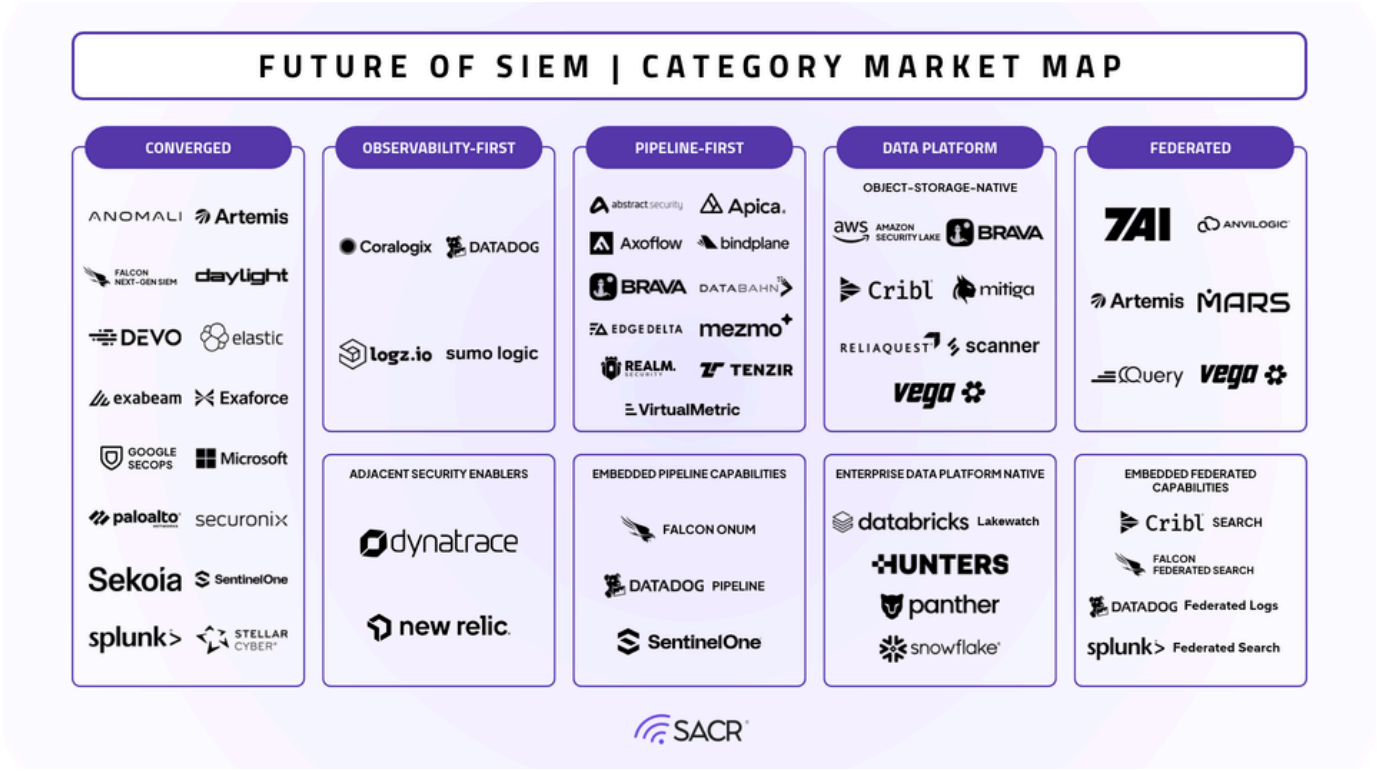
## Key Insights

- SACR finds that the storage tier now decides detection coverage: in tiered platforms such as Microsoft Sentinel, logs held only in the lower-cost tier do not trigger analytics rules.
- Third-party telemetry parity is the most important test for converged platforms, because each vendor's detections and response actions are deepest on its own first-party telemetry.
- Our research indicates that pipelines are moving into storage and detection coverage, with vendors such as Axoflow keeping full-fidelity copies and testing in-stream detection in early access. A pipeline that stores and searches data belongs in the SIEM evaluation.

- Customer custody of raw logs leaves portability open, because indexes, normalized tables, and detection logic can stay tied to the vendor's indexing and normalization pipeline.
- SACR finds that federated vendors often copy more data than the label suggests, and detection against federated sources remains uneven across vendors.
- Agents turn query cost and schema quality into architecture constraints: one agent investigation fans out into many metered queries, and where field names are undocumented, agents guess.

## SACR Thesis

The right SIEM architecture puts control at the layer the SOC can operate well, keeps detection coverage intact as data moves between tiers and systems, and lets the organization change direction without rebuilding its detections, cases, and playbooks.




*Future of SIEM Category Market Map*

## Why Now?

Vendors are expanding outward from their original control point. Converged platforms are adding data lake tiers, pipelines, and federated search, as Microsoft Sentinel has with *its Data Lake tier* and *data federation (Preview)*. Pipeline vendors are adding storage and detection, and federated vendors are adding cases, response, and their own storage. A model label now shows where a vendor started, and the vendor-level comparison shows what the platform does today.

## Actionable Summary

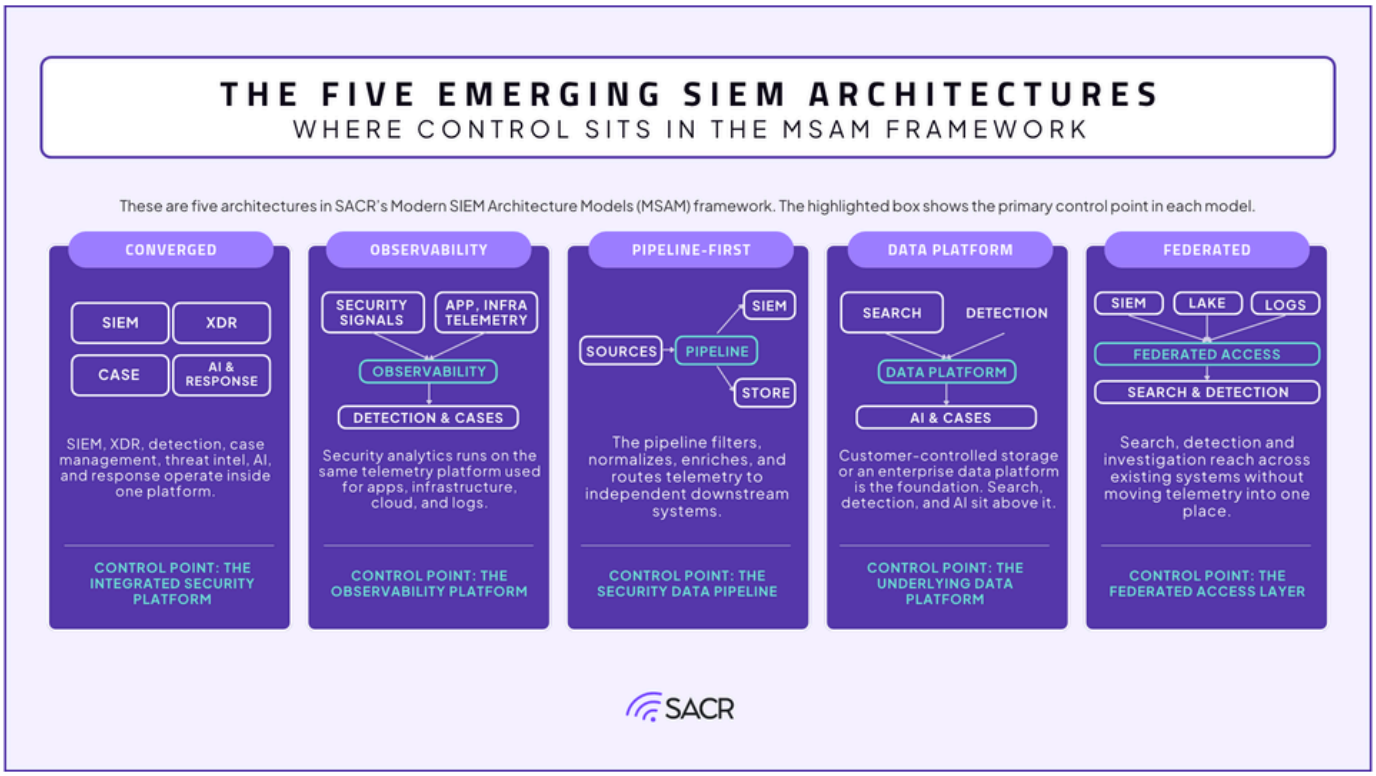
- Before changing retention or storage tiers, map which sources each tier excludes from real-time detection and replay an attack against the lowest-cost tier.
- Where the endpoint detection and response (EDR) and identity estate is mixed, weight third-party detection and response parity above feature breadth, and test it on the two non-native sources the security operations center (SOC) relies on most.
- Treat a vendor's data copies, reduction rules, and indexing and normalization pipeline as part of the SIEM architecture: get a written inventory of what each vendor holds, and confirm what remains usable from customer storage without it.
- Price the dataset a breach investigation needs, not the steady-state dataset, and list which sources and retention windows were cut because of the license metric.

## Scope and Methodology

Vendors are in scope when they own or operate a security data storage and analytics layer, and they are classified by core operating model rather than full feature set. AI SOC platforms are included where they have extended into that layer, as 7AI and Exaforce have. Profiles and criteria draw on vendor briefings, documentation, practitioner interviews, SACR CISO Advisory Board discussions, public practitioner accounts, and SACR analysis. Practitioners SACR interviewed are anonymized by role and organization type. Vendor-supplied figures are marked as vendor-reported, and the profiles are not rankings or procurement recommendations.

# The MSAM Framework

MSAM places each vendor by its primary control point, the layer where the system makes its most consequential decisions about security data. The five models are Converged SIEM/extended detection and response (XDR), Observability-Connected, Pipeline-First, Data Platform, and Federated, each defined in full in *Part I*.



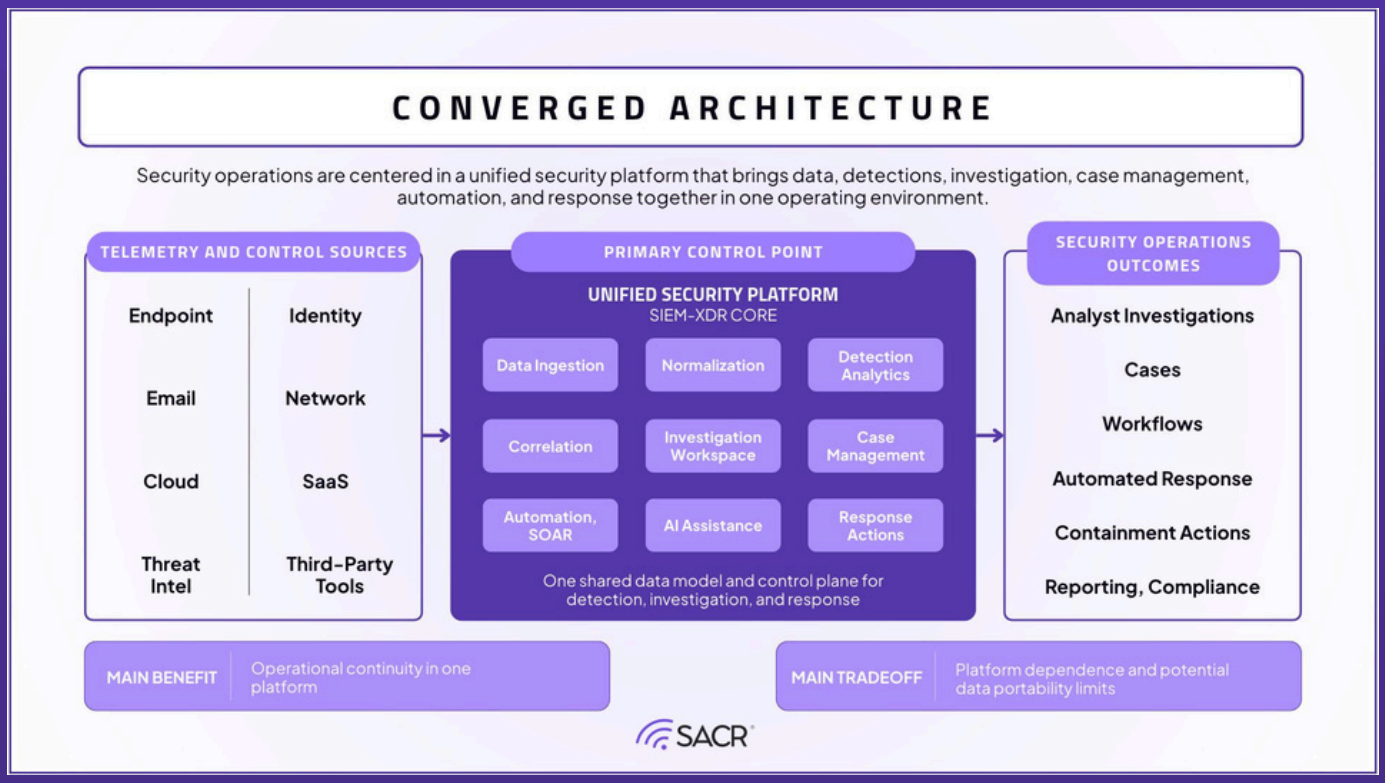

*The five emerging SIEM architectures*

In each model, the primary control point is a different layer: the integrated security platform, the observability platform, the security data pipeline, the underlying data platform, or the federated access layer. Vendor placements appear on the market map, where each vendor sits at the control point that best explains how customers operate it day to day. A vendor whose core architecture is built on two control points appears in both models. Vega queries existing SIEMs in place and also indexes data in customer object storage, so it sits in Federated and Data Platform. Brava Security routes and reduces telemetry ahead of the SIEM and keeps the rest in its own searchable store, so it sits in Pipeline-First and Data Platform. Cribl Stream remains a Pipeline-First product, and Cribl Detect places Cribl in Data Platform and Federated.

AI agents run inside all five models and do not form a sixth. An agent inherits the data access, schemas, retention, response permissions, and metering of the architecture it runs in.

Three of those inheritances show up first in practitioner deployments. The first is cost. Agents fan one question out into many queries, so on platforms that bill by data scanned, compute, or agent runs, cost per investigation becomes a variable line that most SIEM business cases do not model. Superhuman's security team pointed its first autonomous agent at its SIEM's query API and found *the bill was steep*, and Microsoft Sentinel bills *data lake queries per GB of data scanned*. The second is schema. Where field names vary by source and are undocumented, agents guess, and a query that returns results is not the same as a query that is right. A 1Password detection engineer *saw an agent write queries that would not work in the team's SIEM* until it had vendor documentation for each log source. The third is identity. Each agent needs its own scoped, read-only access and a trail that ties every finding to the exact query, time range, and results behind it, separate from human users and service accounts. In a vendor-sponsored survey of 300 CISOs, *53% said they run agents on shared or long-lived keys*.

# Converged SIEM/XDR Model




*Converged architecture*

The Converged model runs SIEM, XDR, threat intelligence, case management, automation, and AI investigation on one platform and one entity model. It cuts integration work and raises switching cost. Converged vendors split mostly on how far detection and response extend beyond their own first-party telemetry. A SOC whose investigations already run on that telemetry gets the most from the platform, and one with a mixed EDR or identity estate should weight third-party parity above feature breadth and map which sources each storage tier excludes from real-time detection.

## Architecture analysis

The first difference between converged vendors is third-party telemetry parity. Every converged vendor has its own first-party telemetry, such as EDR, identity, network, productivity, or threat intelligence, and its out-of-the-box detections, entity resolution, and response actions are deepest there. A connector that ingests a third-party source does not guarantee the same correlation or response options. Palo Alto Networks makes the split explicit: its *standard data sources are limited to ingestion, parsing, and normalization*, while content packs add playbooks, correlation rules, and data model rules on top.

The second is the storage tier, where retention and detection coverage are now linked decisions. Logs kept only in Microsoft Sentinel's Data Lake tier *do not generate alerts from analytics rules or custom detections*. Federated sources sit outside both tiers: Sentinel's *data federation (Preview)* queries Microsoft Fabric, Azure Data Lake Storage (ADLS) Gen2, and Azure Databricks tables in place for hunting and investigation, but Microsoft does not document federated tables as a source for analytics rules or custom detections. Google SecOps *includes 12 months of hot retention* in its packages, and Sekoia's *ExaLog engine* keeps the full retention window searchable without rehydration in the regions where it has rolled out. Tier placement decides which sources are eligible for detection and how long a historical query takes during an incident.

The third is economics. Converged pricing is built around the vendor's own telemetry. Microsoft offers eligible Microsoft 365 E5 customers *a data grant of up to 5 MB per user per day*, and CrowdStrike lets Falcon Insight XDR customers *ingest up to 10 GB per day of third-party data at no additional cost*. The bill, and the sources teams drop first, come from the firewall, SaaS, OT, and cloud data the vendor does not produce. Under per-GB pricing, noisy firewall and network logs are often the first sources left out, because they cost the most to ingest and their detection value is hardest to show, and *68% of organizations in a Red Canary survey* (vendor-sponsored) said they discard data because of SIEM storage costs. Those coverage gaps never appear as detection decisions.

## How the vendors differ

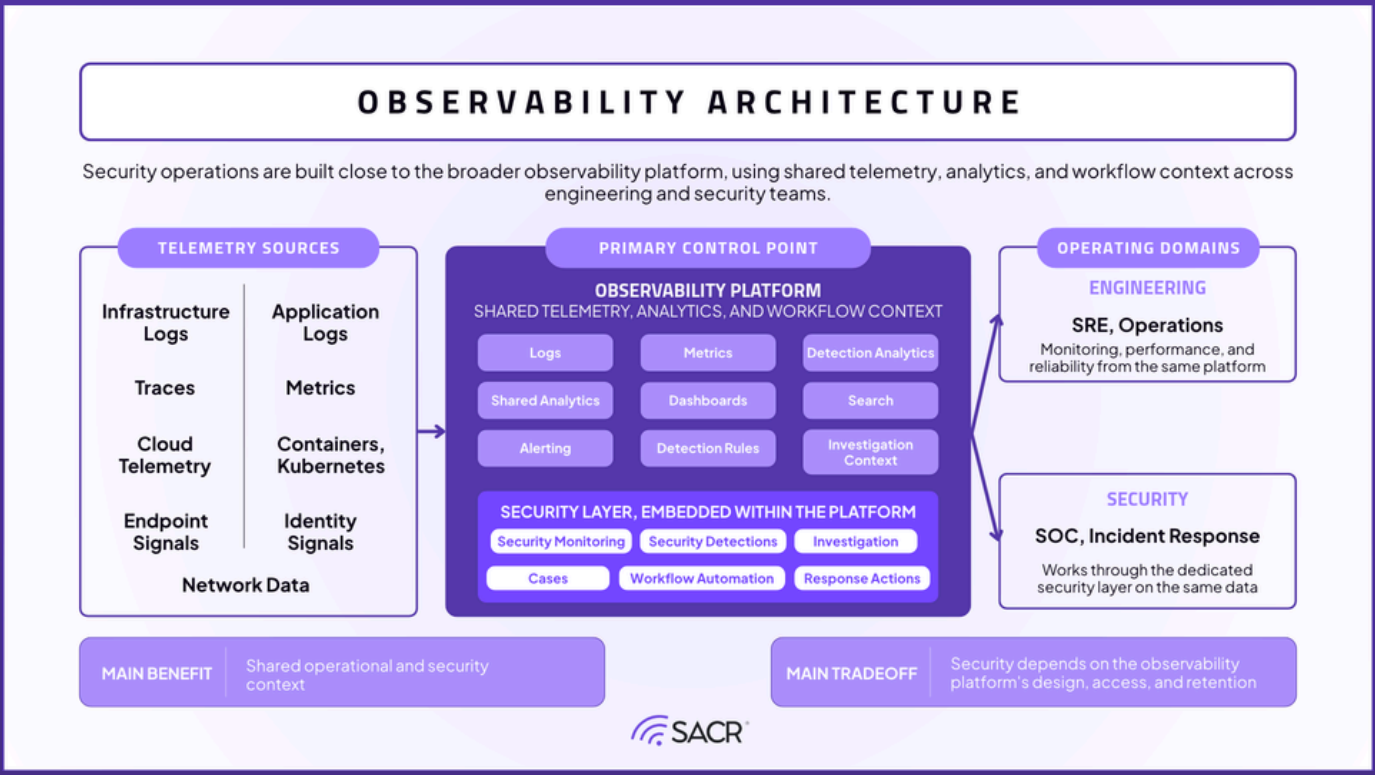
Microsoft Sentinel, CrowdStrike, and Palo Alto Networks build out from their own identity, endpoint, and firewall telemetry, while SentinelOne *normalizes native and third-party telemetry* into the Open Cybersecurity Schema Framework (OCSF). Anomali starts from threat intelligence applied at ingest, Stellar Cyber from its own network sensors and *correlation of alerts from third-party EDR tools*, and Exaforce and Artemis from a model of the environment's entities and relationships. Google SecOps and Sekoia pair in-house detection content and threat intelligence with long hot retention. Daylight *delivers the model as a managed service,* which places much of the control point with the provider. The same question applies to any converged platform run by a managed security service provider (MSSP) or managed detection and response (MDR) provider. Microsoft recommends that MSSPs *host analytics rules and hunting queries in their own tenant,* so that content stays with the provider, and some MDR services act only on the platform's native detections rather than customer-written rules.

Converged vendors also own the connector lifecycle, so detections can need rework without any migration. Microsoft ended incident support for connectors built on the HTTP Data Collector API on September 14, 2026, and the replacement path *can introduce new table names and schemas* that require analytics rules, hunting queries, workbooks, and playbooks to be updated. Cloudflare *told its Enterprise customers to migrate* to its new Sentinel connector by the same date. Export is the other side of the lock-in, because raw native telemetry often leaves the platform only through a separately licensed path. Elastic's *comparison of SIEM data export,* a competitor analysis, reports that CrowdStrike Falcon Data Replicator (FDR) is a licensed add-on, Cortex XSIAM export requires paid add-ons and can take up to two hours, and the Defender streaming API delivers to Azure destinations only.

## What to test in a POC

- Move one source to the lowest-cost storage tier and replay an attack that depends on it. Record which detections stop firing and how long the historical query takes, then run the same replay on a single-tier platform for comparison.
- Onboard the two non-native sources the SOC relies on most, such as an identity provider and a secure web gateway. List which out-of-the-box detections, entity mappings, and response actions are missing compared with first-party telemetry, including any that need add-on content beyond the standard connectors.
- For a managed variant such as Daylight, or any platform an MSSP or MDR provider operates, confirm which detections the provider acts on (native only, or custom rules too), where its rules and playbooks live, and what transfers if the service ends, including detections, case history, and investigation context.
- Price and test the raw-telemetry export path for native sources: the SKU, latency, buffer retention, field completeness against the native console, and supported destinations.
- Ask for the connector deprecation history for the top 10 sources over the last 24 months, which changes altered tables or schemas, and what tooling found the dependent detections.
- Have the SOC onboard one custom source and write its parser without opening a vendor support ticket or waiting on a feature flag, and list every configuration setting only the vendor can change.
- Stop one batch SaaS source and one high-volume source, such as a domain controller, and record how long each takes to raise a source-health alert. Then confirm the batch source does not alert during its normal delivery gap.

# Observability-Connected Model




*Observability-Connected Model*

The Observability-Connected model runs security detection and investigation on the platform that already holds application logs, metrics, traces, service ownership, and deployment history. It fits cloud-native organizations where the same incident shows up in security signals, application errors, and infrastructure changes. It fits less well where the main exposure sits outside the observability platform, such as on-premises networks, operational technology (OT), or legacy infrastructure the platform does not monitor. Buyers should judge it on whether service ownership, deployment history, and traces change the outcome of real investigations.

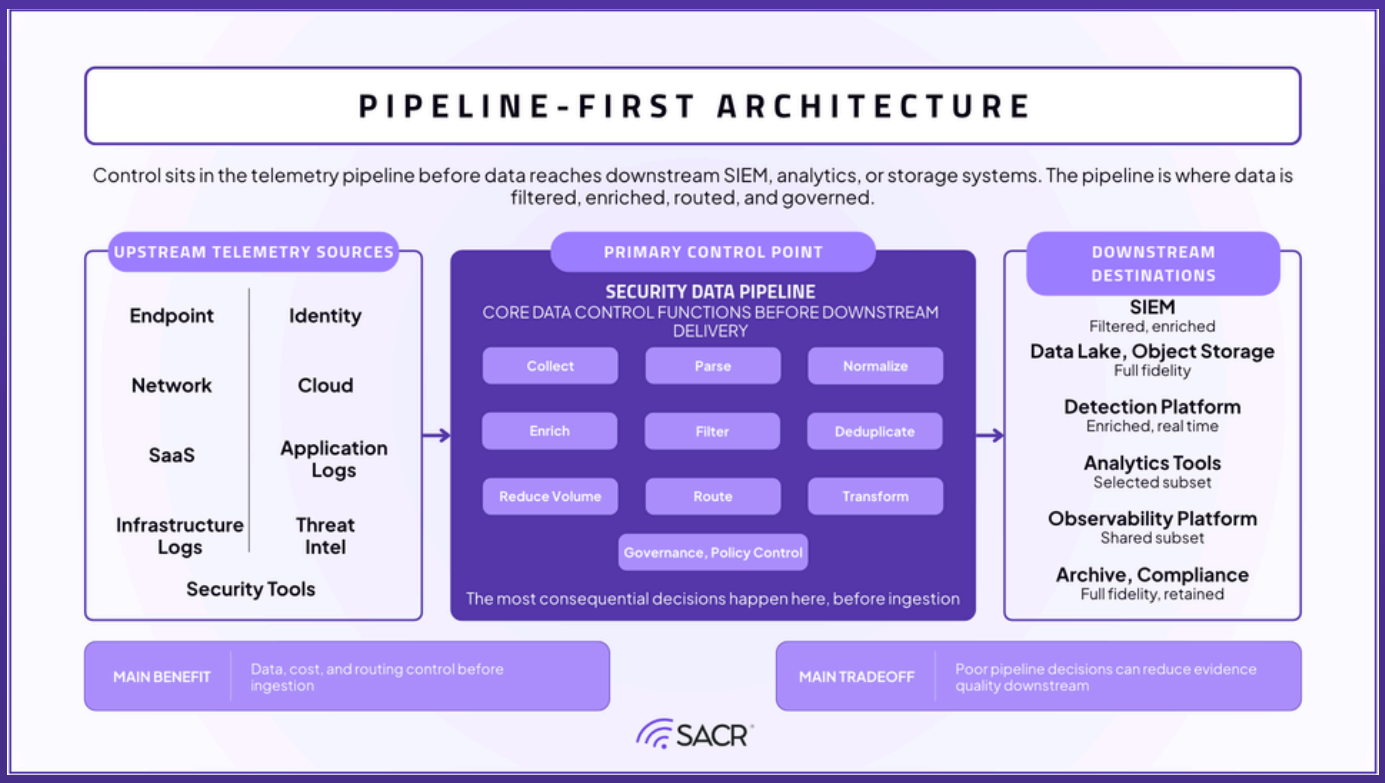
## Architecture analysis

Datadog is the only vendor placed in this model in Part II. Alongside ingestion and tiered storage, it can *query logs held in external data stores* from the same interface, a capability in preview as of June 2026.
With one vendor in the model, the evaluation is about configuration: whether security-grade lookback, tamper resistance, and role-based access control (RBAC) can be set on a platform engineering owns, with change control that both teams sign off on for security-critical sources.

## What to test in a POC

- Investigate a cloud incident from the first signal to the service owner, the most recent deployment, and the affected traces in one workflow, and compare the time against the current SIEM.
- Apply RBAC so developers can query application logs but not identity or security logs, then confirm the restriction also holds in saved queries, dashboards, and AI-assisted searches.
- Price ingestion, indexing, retention, and federated queries at full security log volume with a 12-month lookback.

# Pipeline-First Model




*Pipeline-First Model*

The Pipeline-First model puts the control point before ingestion: the pipeline decides which events and fields are kept, how they are normalized and enriched, and which destinations receive each copy. Part I and SACR's 2025 market guide on security data pipelines cover the economics. Vendors here differ in how they decide what a detection needs. A pipeline that keeps a full-fidelity copy lowers the cost of a bad reduction rule and is harder to replace, so reduction rules need an owner and the same version control, review, and CI/CD pipeline as detection content.

## Architecture analysis

Rules written by a data engineer, attack-simulation mapping, and AI recommendations all cut volume, and each can drop a field that a future detection or investigation needs. The method decides what kind of mistake the pipeline makes and how visible it is.

Pipeline vendors are also moving into storage and detection. A pipeline that only reduces volume for the SIEM is straightforward to replace. One that also stores the full-fidelity copy, runs search, and maps telemetry to detection coverage is harder to remove and should be evaluated as part of the SIEM architecture.

Savings also arrive later than most business cases assume. Early effort goes into migrating sources onto the pipeline rather than reducing them, and governance policies that require data to stay centralized can block re-tiering. Reduction pays off only after data owners agree on what can be dropped or moved to a lower-cost tier. The pipeline also becomes production infrastructure, with worker capacity, pack management, and configuration that need the same operational discipline as the SIEM.

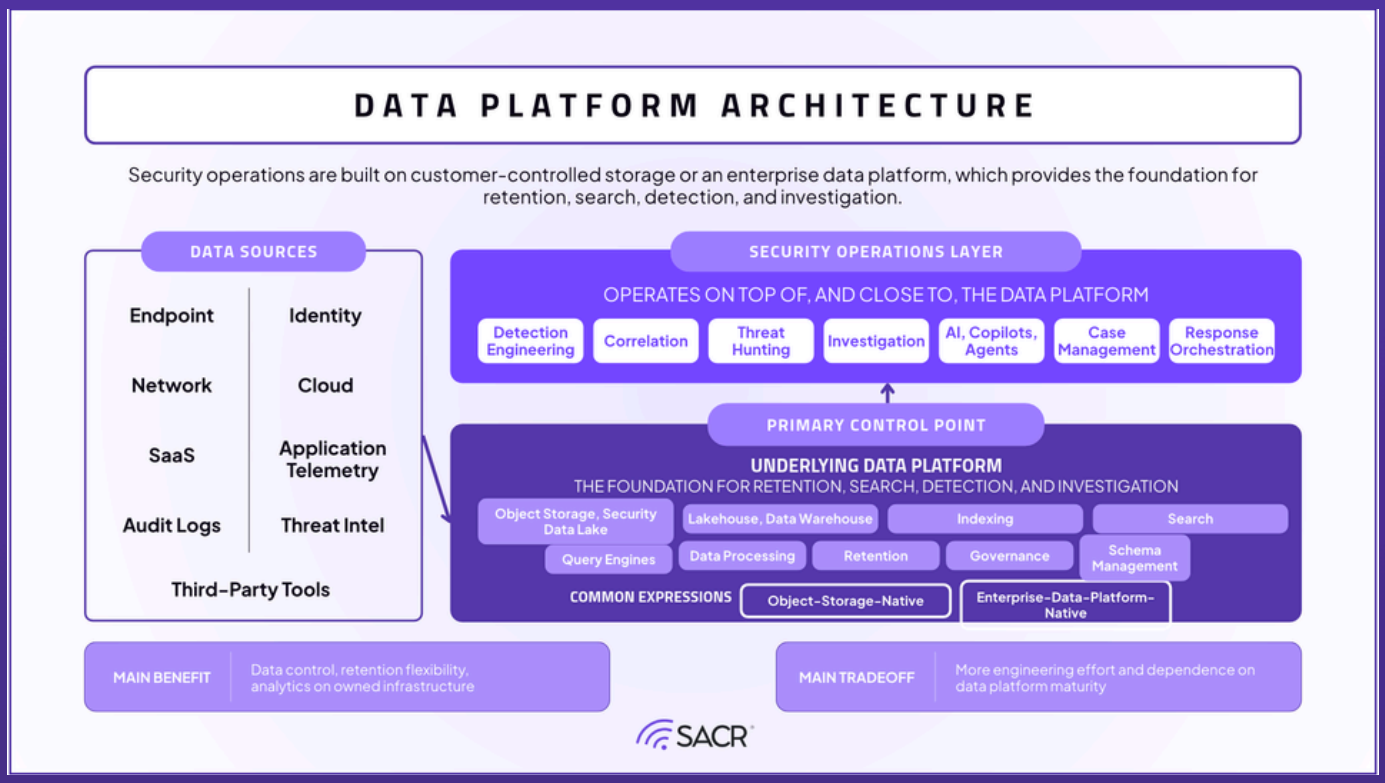
## How the vendors differ

Brava Security *runs the attack each detection rule was written for* and sends the SIEM the events those rules read, moving the rest to low-cost storage it can search through Aquarium and bring back, which also places it in Data Platform. Axoflow *classifies each message against a classification database* and maps it to the schemas each destination needs, keeps *full-fidelity logs in AxoLake*, and has in-stream Sigma detection in early access that flags where a rule lacks the data it needs. Cribl Stream remains a general-purpose pipeline many SOCs run beneath their SIEM, and Cribl's SIEM offering is covered under Data Platform and Federated.

## What to test in a POC

- Run the existing detection library against 30 days of pre-reduction and post-reduction data. Every rule whose hit count changes needs a decision before the reduction rule ships.
- Apply each vendor's recommended reduction to the same three high-volume sources, then compare which fields each method dropped and which rules lost hits. Where the vendor offers a coverage check, compare its reported gaps with the replay results.
- During a simulated investigation, search and rehydrate data from the full-fidelity store, and record how long it takes to get a dropped field back into the SIEM.
- Get the data owner's sign-off on drop and re-tier rules before the POC starts, so it measures reduction rather than relocation.

# Data Platform Model




*Data Platform Model*

The Data Platform model runs security analytics on storage the customer controls. Owning the data does not by itself make it searchable or ready for detection. Vendors differ in how much detection content, case management, and response they run over that storage, and each step toward a full stack trades engineering load for dependence on the vendor's indexing and normalization pipeline. Before choosing, find out which indexes, normalized tables, and detections the SOC could still use from its own storage without that pipeline.

## Architecture analysis

*Part I* described two subtypes: Object-Storage-Native, which adds indexing, search, and detection to customer object storage, and Enterprise-Data-Platform-Native, which builds on an existing lakehouse.

At one end of the Object-Storage-Native range, the vendor provides indexing, search, and detection over the customer's bucket, and the SOC operates parsing, rules, cases, and response around it. At the other, the vendor provides normalization, entity resolution, detections, investigation workflows, and AI access, while the customer keeps custody of the raw data. Raw logs in the customer's bucket are portable. Indexes, normalized tables, enrichment, and detection logic may depend on the vendor's pipeline.

Customer-controlled storage also makes data engineering a SOC dependency. Enterprise-Data-Platform-Native deployments inherit the data organization's operating model. If security does not own the pipelines, tables, and schema changes, new sources and parser fixes compete with revenue analytics for the same engineering queue. Just Eat Takeaway's security team found that *the data engineering team becomes a bottleneck for security initiatives,* and in the *SANS 2026 State of Detection Engineering survey*, co-developed with Anvilogic, only 13% of detection engineers reported high software-engineering proficiency. Normalization to a common schema such as OCSF is an ongoing mapping program, not a one-time conversion. AWS tells Security Lake customers that for custom sources they are *responsible for any necessary transformation to OCSF*, and an Atlassian security logging engineer said *it took a year to map about one-third of the data* the team has mapped today.

Cheap retention is also not the same as fast investigation. Rippling's security team saw queries on large tables *occasionally taking up to an hour* until it added search optimization and query acceleration, and a Scanner case study reports that Ramp's searches over S3 archives *often timed out after 30 minutes*. Both are tunable, but the time and compute cost of a long-range hunt belong in the evaluation.

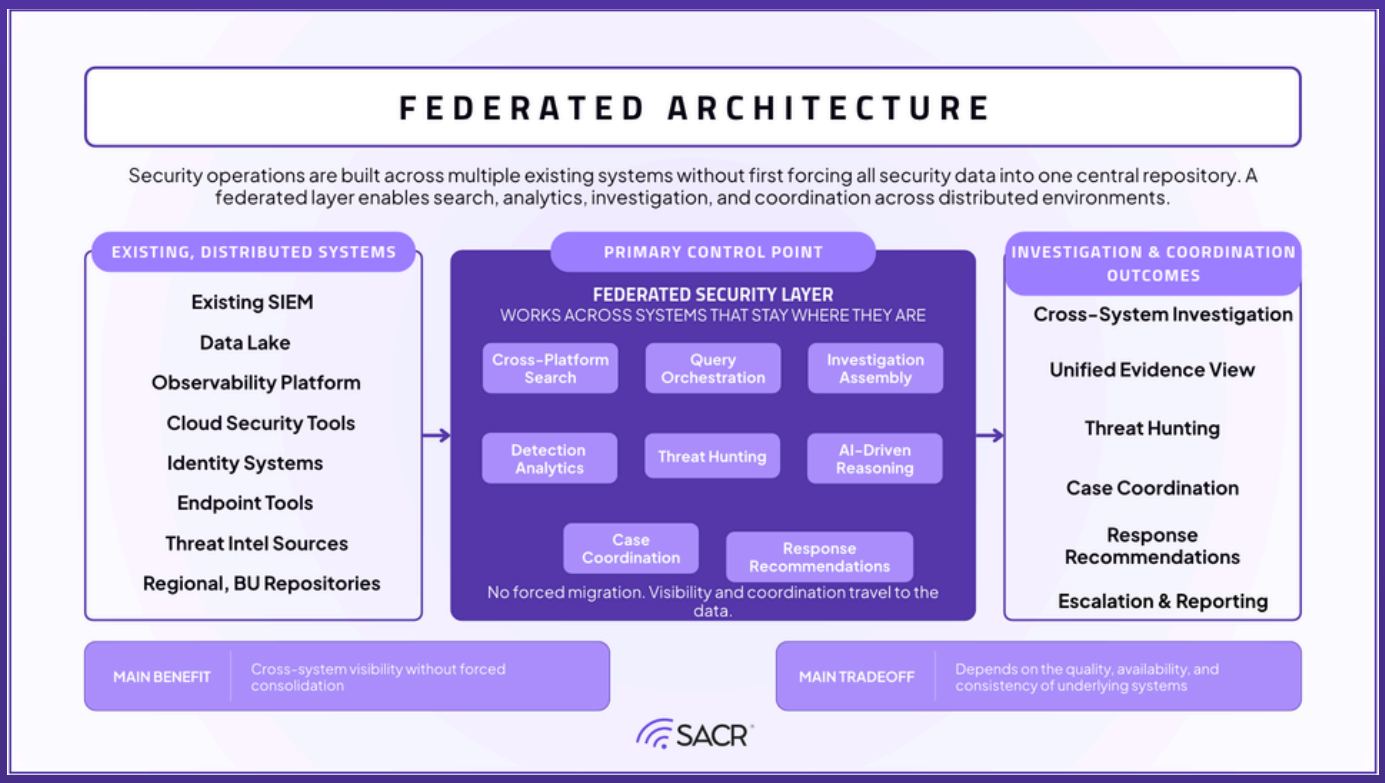
## How the vendors differ

Scanner sits at the lighter end: it *indexes logs in place* in the customer's own Amazon S3 buckets and runs search and continuous detection against them, leaving cases and response to existing tools. Mitiga adds normalization, entity resolution, detections, and investigation workflows over *bring-your-own-bucket (BYOB) storage* and can send a contextualized subset to an existing SIEM. Vega *builds a distributed index in the customer's own object storage*, or Vega provided, and runs analytics and detection over it. Cribl reaches the model from the pipeline, and *Cribl Detect is offered to Cribl.Cloud customers*, so buyers should confirm which storage, compute, and governance settings sit in the Cribl environment. Brava Security also reaches the model from the pipeline: it keeps the telemetry it does not route to the existing SIEM in lower-cost storage that Brava hosts in the customer's chosen AWS, Azure, or GCP region or writes to the customer's own storage, searchable through Aquarium. Buyers should confirm which detections run against that store and which are written for the downstream SIEM.

## What to test in a POC

- Run a 90-day and a 12-month hunt for one indicator of compromise across CloudTrail, identity provider, and proxy logs at production volume, and record wall-clock time, timeouts, and compute cost for that single query.
- Turn off the vendor's indexing and normalization pipeline in a test environment and list which indexes, normalized tables, and detections remain usable from customer storage alone.
- For vendor-managed deployments, list which data the SOC can read directly with its own tools, and which governance and retention settings exist only in the vendor's environment.
- Where the vendor sends a subset to an existing SIEM and keeps the rest in its own store, run the same hunt against both and record which detections execute on the store.
- For the top 20 sources, ask which OCSF version and classes are mapped, what share of fields lands in unmapped, and who maintains each mapping. Then change one upstream log format and time how long it takes to detect and repair the affected rules.

# Federated Model




*Federated Model*

The Federated model runs search, detection, and investigation across existing SIEMs, lakes, and log stores without migrating the data first. Its main risk, covered in Part I, is a query that returns results from four sources, times out on the fifth, and still looks complete. Partial-result reporting on every query and an accurate inventory of copied data decide whether analysts and agents can trust a cross-source answer, and any federated detection claim should be tested before it is relied on.

## Architecture analysis

Most federated vendors now copy some data. Indexes, cached metadata, hot storage for high-value sources, and managed collection paths all improve query performance, and they also change what the vendor holds. Buyers need a written inventory of what is queried live, what is indexed or cached, where those copies sit, and how long they are retained.

Detection on federated data also behaves differently from detection on ingested data. Rules that evaluate events as they arrive run close to real time. Scheduled queries against external sources run on an interval and depend on each source's API limits and availability. Detection latency and coverage need to be measured per source.

Federation also moves entity resolution into every investigation. Each source keeps its own schema, so the analyst or agent has to resolve users, hosts, and addresses at query time. Superhuman's security team built a context graph for its threat-hunting agent because host data across its device management, EDR, and network logs has *no single key linking it*, and Corelight's Black Hat NOC write-up found that *field semantics, version drift, and vendor-specific log extensions still trip up models*.

## How the vendors differ

Vega *queries datasets in place with Kusto Query Language (KQL) or natural language* as though they shared one OCSF schema, and its distributed index in customer object storage places it in the Data Platform model as well. 7AI *normalizes query results to OCSF*, offers hot storage for data customers choose to collect, and separates rules on collected events from scheduled queries against external sources, so buyers can see which detections run in real time and which run on an interval. *Cribl Search* queries data in place alongside Cribl's own lakehouse engine, and as of October 2026 Cribl had not yet made federated detections generally available.

## What to test in a POC

- Request the vendor's inventory of what is queried live, indexed, cached, or stored, then check it against the storage and collection activity observed during the trial.
- Run the same detection as a rule on collected data and as a scheduled query against an external source, then compare detection latency and missed events.
- Query as a restricted analyst, a service account, and an AI agent, and confirm each sees only what its permissions in the source system allow. Confirm that each agent query runs under its own identity and that every agent finding traces to the exact query, time range, and results returned.
- Give an agent ten hunts across three sources it has not seen, and count failed queries, queries that ran against wrong or nonexistent fields, and whether the API returns bounded, paginated, entity-shaped results or a raw row dump.

# Practical Recommendations

*Part I*'s recommendations set out the baseline and the four architecture decisions that apply to any SIEM evaluation. Once a starting model is chosen, five steps matter most when comparing vendors within it.

1. Map detection eligibility before any storage decision. For each critical source, record which tier it lands in and whether that tier can trigger real-time detections, then replay a known attack against it. Record which sources were excluded and which retention windows were shortened because of the license metric, and price the dataset a breach investigation needs rather than the steady-state dataset.
2. Test parity on the SOC's own estate. Onboard the non-native sources the SOC depends on and compare detections, entity mapping, and response actions with native telemetry.
3. Inventory what each vendor holds. Ask for a written account of what is queried live, indexed, cached, copied, or rehydrated, and check it during the POC.
4. Estimate switching cost before signing. Export detections, case history, and a representative dataset, and record what would have to be rebuilt, including what the SOC can read from its own storage without the vendor's pipeline. Migrations that work prune before they port, run both platforms in parallel, move the least-queried sources or aged data first, and use the renewal date as the forcing function. Etsy's security team, for example, recommended *focusing on the best detection rules and log sources instead of migrating everything wholesale*.
5. Assign a system of record for raw data, detections, cases, response actions, and agent access before a second platform goes in. Start with the case of record and where response playbooks run. The SIEM, the security orchestration, automation, and response (SOAR) platform, and the AI SOC layer now all ship agent builders, and each will offer to be both. When those choices accumulate through individual purchases, the SOC ends up with the same fragmentation the modernization was meant to remove.

## Questions Every CISO Should Ask

- Which of our sources can trigger real-time detections in each storage tier, and what happens to that coverage when we change retention? Which sources did we exclude, and which retention windows did we shorten, because of the license metric?
- Which out-of-the-box detections and response actions apply to our third-party sources, and which require add-on content? What share of our projected volume is non-native to this vendor, and what does it cost at year-three growth?
- What data do you copy, index, or cache, where does it sit, and how long is it kept?
- If we stop using your indexing and normalization pipeline, which indexes, normalized tables, and detections remain usable from our own storage?
- Which capabilities in this evaluation are generally available today, and which are in preview or on the roadmap?
- If you are acquired, merged, or discontinue this product, what happens to our detection content, case history, and integrations, what does the contract commit you to, and how quickly can we export them?
- Does each agent query run under its own identity with scoped, read-only access, and can we trace every agent finding to the exact query, time range, and results returned, separately from human users and service accounts?
- If an MSSP or MDR provider operates the platform, which of our detections will it act on, where do its rules and playbooks live, and do we keep them at exit?
- Where do AI gateway, AI detection and response (AIDR), and agent or Model Context Protocol (MCP) activity logs land today, and does any detection or playbook consume them?

## Success Metrics

Track these through the POC and the first year in production:

- Validated coverage: the share of prioritized MITRE ATT&CK techniques for which an emulated or replayed attack produced the expected alert in the last 90 days, on sources whose parsing and freshness checks passed. Rule counts mapped to ATT&CK overstate coverage: CardinalOps data from production SIEMs found *21% average technique coverage and 13% of rules non-functional*.
- Detection gaps found per quarter, by cause: missing rule, missing source, or storage-tier or retention exclusion.
- Mean time to detect (MTTD) per source, false-positive rate per rule after any pipeline or retention change, and the share of alert volume from the noisiest 5% of rules.
- Cost per GB ingested and searched, and query and token cost per agent investigation.
- The share of agent findings that reproduce when the logged query is re-run.
- For pipelines, net savings after license, infrastructure, and staffing costs, tracked alongside retention of detection-relevant fields.
- The engineering hours needed to onboard a new source and port a custom detection.

# Conclusion

SACR's SIEM research has followed the security data platform breaking apart in 2024, control moving into the security data path in 2025, and, in Part I, the SIEM decision becoming an architecture decision. Part II tests that decision vendor by vendor. Feature lists look alike across most of the 19 vendors here. The differences sit in retention design, third-party coverage, data copies, how much of the stack runs on customer storage, and how the platform meters and scopes agent workloads, and a POC is usually the first place they show up.

Evaluate SIEM vendors on where their control point actually sits and what it would take to move away from it, and design the POC to expose both before the contract does.

Overlap between the models will keep growing as converged platforms add data lake tiers and federated search, pipelines add storage and detection, and federated vendors add their own storage. SACR expects vendor positions within the models to change faster than the models themselves.


Cribl®

# Vendor Profile

## Cribl Detect Puts a SIEM on Top of a Telemetry Data Platform

*Cribl Detect* extends the Cribl AI Platform for Telemetry into a packaged SIEM, available to Cribl.Cloud customers. Cribl says it covers detection, investigation, response, detection posture management and AI-assisted triage across the data the platform can reach, including federated sources and in-stream detections. The CardinalOps and Radiant technology Cribl acquired is built into Detect and isn't sold separately.

Cribl is positioning Detect as a full SIEM replacement. In practice, many enterprises will find it easier to start by running Detect alongside existing infrastructure than to replace their SIEM in one move. Either way, buyers should verify supported sources, response integrations, licensing and what their own workload will cost before adopting it.

Detect is built on mature telemetry products, which makes it a credible entry, but it only launched at CriblCon 2026 and SACR has not yet seen independent evidence of customers running it in production. Cribl's near-term roadmap adds federated detections by the end of 2026, EU compliance reporting, and an agentic SOC in Q4 2026.

## Analytics decoupled from a single storage tier

Detect doesn't tie security analytics to one proprietary storage tier. Detections can fire as data moves through the pipeline, and investigations can span Cribl-managed and external stores. Detect runs on Cribl's native Lakehouse Engine and federated Search, with an embedded pipeline powering ingestion (Cribl Stream is sold separately and is not required for Detect). That design is the core of Cribl's replacement argument, since it gives customers a way off a centralized, high-cost SIEM architecture.

Cribl says detections will also run against federated sources. How that works will drive cost, latency and load on the source system, so buyers should test it against their own stores once it ships.

## Built on the App Framework

Cribl built Detect on its Agentic Development Environment and runtime for apps, workflows and agents. According to Cribl, the framework is generally available, more than 150 apps have been built on it, and a new Marketplace distributes apps, packs and agents from Cribl, partners and the community. If that plays out, Cribl could ship new security apps faster, and partner or community apps running on the same telemetry would make the platform more than one vendor's SIEM. Neither has happened yet, the app count is Cribl's own figure, and most of the apps demoed at CriblCon were observability use cases rather than security.

## Detection posture and coverage

Cribl's acquisition of *CardinalOps*, announced in July 2026, brings detection engineering, coverage analysis and MITRE ATT&CK-oriented capabilities into Detect.

A pipeline that only moves data is easier to pull out of the SOC than one that can also show whether the available telemetry supports the detections the organization needs. With CardinalOps, Cribl is starting to tie what data gets collected to whether that data is enough to detect specific threats.

For buyers, the test is practical: import representative detection content, check how well it translates across environments, and work out how much ongoing tuning it needs and who owns it.

## AI-assisted triage and investigation

Cribl's acquisition of *Radiant Security technology assets*, which followed CardinalOps, adds AI-assisted triage, investigation and response to Detect. Cribl frames it as the next step in pushing the platform into AI-driven security operations.

With Radiant, Detect covers more of the path from telemetry to detection, triage and investigation, which makes it easier for Cribl to argue that Detect can replace an incumbent SIEM outright.

## The Data Platform Under Detect

Detect sits on a data platform that combines a native optimized store with federated access. Cribl says customers can point a source at a Lakehouse or Cribl Lake endpoint and send data straight in. Detect's embedded streaming works in the background, recognizing the data type and normalizing it (to OCSF, for example) without anyone configuring a pipeline. Teams that want more control can still build their own Stream pipelines and land the output in the same store, or other destinations (e.g. existing object stores or security data lakes).

## Stream: the policy boundary

*Cribl Stream* is the control point to filter, reduce, enrich, mask, transform and route telemetry before it reaches an analytics platform. It can route the same source to a SIEM, a data lake, an observability platform or another destination, and it can redact PII and other sensitive data along the way. That makes the pipeline a **policy boundary**: a transformation or routing decision made in Cribl affects every downstream consumer.

A bad parser, an overly aggressive reduction rule or a misrouted source can drop evidence before a detection, an analyst or an AI agent ever sees it. That puts pipeline governance, testing, rollback and replay on the buyer's list of security requirements.

## Search: federation and owned analytics

*Cribl Search* supports two modes: federated search over data that stays in external systems, and lakehouse engines that ingest and store data for direct analysis. Federation keeps existing storage choices and cuts data movement, but investigations inherit the source system's latency, permissions, availability and limits. Lakehouse engines give Cribl more control over query performance, and in exchange storage, compute, governance and lifecycle responsibility move into the Cribl environment.

Cribl says federated search V2, released in March 2026, is three to four times faster and now completes queries that used to fail. Cribl also acknowledges it still isn't fast enough for operational use cases, which is why it built the Lakehouse Engine as a native optimized store. The two paths aren't interchangeable, so a POC should run the same investigation across federated and lakehouse data and measure **speed, completeness, cost, and behavior when the underlying source is unavailable**.

## Buyer Tradeoff: A Predictable Cost Model, Delivered Cloud-First

Cribl prices Detect on the infrastructure it runs on instead of ingest or search volume, and says customers will cut costs by up to 50% compared with their existing SIEM. That figure is Cribl's, and SACR hasn't seen a measured customer comparison to back it up.

Two things decide whether the cost case holds up. The first is federated search across months of data without rehydration delays, which a Fortune 50 customer quoted by *TechTarget* called the claim that matters most. The second is agent-driven query volume. Federated performance is capped by the source system's APIs, throttling and query charges, and those limits bite harder when agents test many hypotheses in parallel. Cribl's view is that agents make query language and speed benchmarks matter less, since an agent will wait for results. If that's right, source limits and query cost are what buyers should be testing.

Deployment is the other constraint. Detect is Cribl.Cloud only for now, with no on-premises or self-managed option, which the same customer called a showstopper for regulated industries. Cribl says it is considering both. Migration is a real project too, and that customer expects most large enterprises to keep their SIEM at first.

## SACR Analyst Take

Cribl started as a pipeline company and now spans the data platform and federated categories. Detect is its entry into the SIEM market, combining CardinalOps detection engineering, Radiant's AI-assisted triage and investigation, and a security analytics layer across the Lakehouse Engine, federated Search and in-stream detection, all on the data foundation Cribl already sells.

The move changes what Cribl asks of buyers. As a pipeline, Cribl sat under whatever SIEM the customer chose. With Detect, it competes for the same budget and workflow as the SIEM vendors it routes data to, and those vendors and their customers may start to look at the pipeline decision differently.

Cribl's established telemetry products give Detect credibility, but buyers should expect the packaged SIEM to keep maturing and should test it on their own data before treating it as a drop-in replacement.

That POC should cover failure and recovery as well as throughput and data reduction, including what happens when a federated source goes down and how operations continue outside Cribl.

## What security practitioners should validate

- Run Detect alongside the incumbent SIEM on the same data and compare detection coverage, investigation time and cost.
- Replay known attacks before and after pipeline reduction and verify that detection-relevant evidence survives.
- Run the same hunt through federated and lakehouse engines; compare latency, completeness, cost and source-failure behavior. When federated detections ship, test them the same way and measure query load on the source system.
- Import representative detection content into Detect and test translation, ATT&CK coverage and maintenance requirements.
- Test AI-assisted triage on an investigation that requires a non-Cribl response, and confirm that evidence and outcomes remain portable.

Software Analyst®
Cyber Research

business personal

softwareanalyst.io

SACR®

Trusted research. Sharp insights. Real conversation.

| | |
|---|---|
| CISO | VENDOR |
| SECURITY TEAMS | INVESTORS |