ALLIED BizTech SOLUTIONS

FIELD GUIDE · 2026

Rebuild the data ops stack yourself. *The plan we follow.*

For SaaS teams of 50 to 500 engineers who pay Datadog, Sentry, Segment and Mixpanel by the host, the event and the tracked user, and want one open-source stack on their own AWS.

- → Is it worth it, at your size
- → What to rebuild and what to drop
- → One foundation, four modules
- → Migration and go-live checklists
- → The keys that link four tools
- → A 14-week build plan
- → Volume, retention, cutover order
- → Where teams get stuck

Four tools, four meters. *One stack replaces them.*

Datadog bills per host and per log line, sentry.io per event, Segment and Mixpanel per tracked user. Each meter grows on its own. A rebuild runs the open-source originals on your AWS, on one ClickHouse, one Postgres and one S3 bucket, and the bill follows bytes stored, not hosts or users.

STACK	FOUR TOOLS A YEAR	REBUILT: YEAR ONE / AFTER	SAVED IN YEAR ONE
50 engineers · 100 hosts · 100K MTU	~$120,000	~$76,800 / $32,000	~$43,200
100 engineers · 200 hosts · 500K MTU	~$264,000	~$98,800 / $54,000	~$165,200
500 engineers · 1K hosts · 5M MTU	~$700,000+	~$212,200 / $130,000	~$487,800

Tool prices are their public list prices. "Rebuilt" runs LGTM, self-hosted Sentry, RudderStack and PostHog on your AWS: direct billing plus hosting and maintenance ($22K + $10K, $36K + $18K and $98K + $32K a year), plus our fixed build price: $44,800 for all four modules, $82,200 (the top of the $61,000–$82,200 scale tier) from 500 engineers. If your own team builds it, put your engineering cost in that place.

THE RULE OF THUMB

At 50 engineers, year one saves ~$43K and ~$88K a year comes back after. **At 100 engineers, the rebuild saves ~$165K in year one and ~$585K over three years, and pays back in about three months.** At 500 engineers, year one crosses $410K.

Who you need

- **A senior platform engineer** at home with Kubernetes, Terraform, ClickHouse and S3: the foundation, Sentry and LGTM.
- **A senior data engineer** for the event side: SDKs, the tracking plan, RudderStack destinations, the PostHog import. The two tracks run side by side; our team ships all four modules in 10–14 weeks.
- **Your SRE lead and your product analytics lead**, two hours a week each: the dashboards, alerts and funnels that matter, and each pilot's sign-off.
- **An owner after launch**: upgrades, capacity, retention, and on-call for the stack itself.

Four modules. *Drop the rest.*

Each vendor sells a product for every signal; your engineers open a fraction of each, every day. Those are the project. On Splunk, New Relic or Amplitude instead? Same architecture: LGTM or PostHog in their place.

REBUILD ONE FOR ONE

- **Observability** (Datadog to LGTM): traces, log search, dashboards, host metrics, alerts, on-call routing, SLOs
- **Error tracking** (sentry.io to self-hosted Sentry): the same code and SDKs, issues, releases, source maps, Replay
- **Event routing** (Segment to RudderStack): identify, track, page and group calls unchanged; sources, destinations, transformations
- **Product analytics** (Mixpanel to PostHog): events, funnels, retention, cohorts, dashboards, SQL on ClickHouse
- **Shared identity**: SSO, one role model, one audit log across all four
- **Cross-module links**: an error opens its trace, its session and its events
- **Vendor configs moved**: dashboards, alerts, projects, audiences

REBUILD, SIMPLER

- **Watchdog**: alert rules, plus custom detection where a rule falls short
- **Audiences**: basic ones in RudderStack; complex ones from your warehouse
- **Mixpanel boards**: a new layout; 2–3 hours of retraining
- **Synthetics**: Grafana Synthetic Monitoring and k6

DROP ON PURPOSE

- **Cron monitoring**: Healthchecks.io
- **Mobile attribution**: Adjust, AppsFlyer or Branch direct
- **Twilio Engage**: a separate project
- **Cloud security**: Falco and Wazuh, an add-on

ADD WHAT PER-UNIT PRICING DISCOURAGES

- New hosts, events and users at no fee
- Retention per signal: hot logs 7 days, cold 365 on S3
- Reverse ETL in place of Hightouch or Census

One foundation. *Four modules on top.*

Each vendor wraps an open-source original. Run the originals on one foundation and four tools become one stack: one ClickHouse, one Postgres, one S3 bucket, one sign-in, one Helm config.

LAYER 1 · COLLECT

Four ingest paths

- **OTel Collector** with its Datadog receiver, so dd-agents keep sending
- **Sentry Relay**: rate limits, sampling, PII scrubbing
- **RudderStack** data plane, Segment-SDK compatible
- **PostHog** capture, fed by RudderStack

LAYER 2 · STORE

The shared foundation

- **ClickHouse**: Sentry events and PostHog events, a database each
- **Postgres**: Sentry, PostHog, RudderStack and Grafana metadata
- **Mimir, Loki, Tempo**: metrics, logs, traces
- **S3**: long-term data, raw events, replays, source maps
- **Redis**: queues and caches

LAYER 3 · USE

Four UIs, one sign-in

- **Grafana** and Alertmanager to PagerDuty, OpsGenie, Slack
- **Sentry** at sentry.yourcompany.com
- **PostHog**: funnels, cohorts, replay, flags
- SSO, one role model and one audit log in front of all four

Decisions to make in week one

DECISION	OUR DEFAULT	WHY
Hosting	Your AWS account, EKS, multi-AZ	Your data, your residency, your bill
ClickHouse	One cluster, a database and user per module	Limits keep the modules apart (page 7)
Postgres	One instance, a database per module, pooled	One backup, one restore drill
S3	One bucket, a prefix and lifecycle rule per module	Retention is a setting, not a migration
Keys	One user ID, trace ID and label set (page 5)	Every link between modules needs them
Sign-in	Your SSO, one role model, one audit log	Offboarding happens once, for all four

Licenses: Grafana, Loki, Tempo and Mimir are AGPLv3, RudderStack Apache 2.0, PostHog MIT; Sentry's license allows self-hosting for your own company. Keep every Helm chart and Terraform module in Git from day one.

The tools bring their schemas. *You bring the keys.*

Sentry, PostHog, RudderStack and LGTM each ship their own tables. Your design work is the shared layer: the keys every module carries, so an error opens its trace, its session and its events. Agree it before the first collector, DSN or writeKey changes.

KEY OR RECORD · STORE	WHAT IT HOLDS
Resource labels · everywhere	service, env, region, cluster, team, version: the same names on metrics, logs, traces, Sentry tags and server events
user_id · everywhere	your database ID: Sentry's user.id, RudderStack's userId, PostHog's distinct_id after login
trace_id · everywhere	in every log line, on every Sentry event, as a property on server-side events
release · Sentry, Tempo, PostHog	the version CI builds, set on every error, span and event
session_id · PostHog, Sentry	PostHog's session, written into Sentry's event context, so an error opens its replay
Errors · ClickHouse	event_id, issue, project, release, environment, trace_id, user.id, stack trace
Events · S3, ClickHouse	messageId, event, userId, anonymousId, properties, timestamp; a raw copy in S3 before any destination
Retention · Git	module, signal, hot days, cold days, ingest limit, series limit
Access · Postgres	user, SSO identity, role, team; the audit log: who, module, action, object, when
Moved configs · Git	dashboards, alerts, Sentry projects, sources, boards, each with its vendor ID

TWO RULES THAT SAVE A REBUILD

One user ID, from your database, in every module: never an email and never a vendor's ID. Emails change; a database ID does not, and every link between modules matches on it. **Keep IDs out of labels and tags**: user, request and session IDs belong in the log line, the event body or the span. One unbounded label multiplies Mimir's series and slows every Sentry search.

Fourteen weeks, *four pilots, four cutovers.*

This is the order we build in: one foundation, then the lowest-risk module first. Each module runs a one-week pilot next to its vendor and cuts over before the next one's pilot starts. The event work begins in week three, beside the infrastructure work.

WEEKS	BUILD	DONE WHEN
1–2	**Foundation.** Inventory all four tools: agents, monitors, dashboards, DSNs, projects, sources, destinations, events, boards, renewal dates. EKS, ClickHouse, Postgres, Redis and S3 in Terraform; SSO, roles, the audit log; keys and retention agreed.	The inventory is signed off; engineers sign in once, to everything
3–4	**Sentry.** Relay with rate limits and PII scrubbing; projects, teams, ownership and alert rules; source maps in every CI pipeline. Pilot week, then one new DSN per project, swapped in CI.	Every alert sentry.io sent, your Sentry sent too
5–7	**LGTM.** The OTel Collector in front of the dd-agents; Mimir, Loki and Tempo on S3; dashboards converted from their JSON exports, the monitors that fire translated. Pilot week, then agents point only at the collectors.	On-call works a full shift from Grafana
8–9	**RudderStack.** Sources, destinations with fresh credentials, transformations, the tracking plan validated at ingest. Pilot week, then a new writeKey per source.	Counts match Segment's, per destination
10–12	**PostHog.** PostHog added as a RudderStack destination; 12 months of Mixpanel history imported; funnels, cohorts and boards rebuilt. Pilot week, then Mixpanel's destination off.	Leadership's funnels match, number for number
13–14	**One stack.** Cross-module links on one Grafana home, monitoring of the stack itself, a restore drill from S3, the runbook and training.	An error opens its trace, session and events in one click

AFTER EACH CUTOVER

Stop sending to the vendor and keep the account **until its renewal date**, then cancel. Four vendors means four renewal dates: the week-one inventory lists them, and none is cut early.

One store, four workloads. *Keep them apart.*

Installing four open-source tools is the easy part. The work is running them on one foundation at mid-stage volume (200 hosts, 300GB of logs a day, 50M Sentry events a month, 500K tracked users) without one module slowing the others.

1. **Budget the volume first.** For each module, write down series, log gigabytes, error events and analytics events a day, and how long each is kept. Size ClickHouse, disk and S3 from that sheet, and check it monthly.
2. **Give each module its own room.** In ClickHouse, a database and a user per module, with memory limits and query quotas per user, so a year-long PostHog funnel never delays Sentry's inserts. Pin a ClickHouse version both Sentry and PostHog support, and upgrade when both do.
3. **Cap cardinality at the door.** Drop or rewrite unbounded labels in the collector and tags in Relay. Set series and ingest limits per tenant, and review the top series by team every week.
4. **Set retention per signal.** Hot logs 7 days, cold 365 on S3; Sentry projects at 7, 30 or 90 days. Sample traces at ingest: keep every error and slow trace, a share of the rest.
5. **Keep Postgres calm.** RudderStack queues its jobs in Postgres and writes all day. Give it its own database, watch its size, and pool connections for all four modules.

The cutover order

Lowest risk first. **Sentry** goes first: no user-facing change, and the fastest payback. **Datadog** second: engineers adopt Grafana fastest. **Segment** third: the SDK is a drop-in, and RudderStack keeps feeding Mixpanel while PostHog fills. **Mixpanel** last: the analytics team retrains on new boards, so it moves once clean events are flowing.

Watch the watcher

Run a small, separate monitor of the stack itself: ingest drops, ClickHouse disk and merges, Postgres size, S3 errors. When monitoring fails, it fails silently, unless something else is watching it.

Move it, launch it, *run it.*

Migration

- ☐ Inventory all four tools, with each renewal date
- ☐ Keep each vendor's ID on every dashboard, alert, project and board
- ☐ Put the OTel Collector's Datadog receiver in front of existing agents
- ☐ Translate the monitors that fire; retire the rest
- ☐ Recreate Sentry projects and teams with the same slugs
- ☐ Export the tracking plan into Git; fresh credentials per destination
- ☐ Import 12 months of Mixpanel history through the Export API
- ☐ Run each module next to its vendor for its pilot week

Go-live

- ☐ Engineers signing in to all four through your SSO
- ☐ Alert routes match today's PagerDuty, OpsGenie and Slack
- ☐ Every agent, DSN and writeKey pointed at the new stack
- ☐ Key funnels match Mixpanel, number for number
- ☐ A restore from S3 and the Postgres snapshot proven
- ☐ Rollback ready per module: one config change

Running it

- ☐ **Always on:** monitoring of the stack itself, security patches
- ☐ **Weekly:** top series, log volume and noisy alerts by team; tracking-plan violations
- ☐ **Monthly:** upgrades through staging, one module at a time
- ☐ **Quarterly:** cost review, the AWS bill against the four bills you replaced; retention review
- ☐ **Yearly:** a restore drill and an access review

Who owns it

Name one person before launch. This stack is how you learn everything else is broken, and how product decisions get made; someone has to answer when the stack itself is the thing that broke.

Five ways a rebuild *stalls.*

1. Four projects instead of one

Four separate rebuilds mean four ClickHouse and Postgres installs, four SSO setups and four systems to monitor, upgrade and restore, with no data flowing between them: about 22 weeks run in sequence. One foundation ships all four in 10–14.

2. Switching vendors and calling it a saving

New Relic, Splunk, Bugsnag, mParticle and Amplitude are the same hosted model under another brand. The per-host, per-event and per-user meters move with you.

3. The hosted open source, at scale

Grafana Cloud suits teams below 200 hosts and approaches Datadog's pricing above 500. RudderStack Cloud and PostHog Cloud approach Segment's and Mixpanel's above ~500K MTU. Self-hosting the same code is far cheaper at that point.

4. A demo install in production

Each tool installs in an afternoon and looks finished. ClickHouse tuning, Mimir limits, retention, upgrades and on-call are the real work: 2–6 weeks of senior time per module, from Segment at the low end to Datadog at the high end.

5. Cutting over all four at once

A pilot week per module finds the monitor nobody listed, the DSN in a forgotten app and the destination that goes quiet. Four small cutovers beat one big one.

And one that is not a mistake: starting with one module

Many teams start with Datadog or Sentry, where the math is sharpest, and add the rest within six months. Build the foundation for all four on day one, and each later module is a smaller project.

RATHER HAVE IT BUILT?

We build this exact plan. *Fixed price.*

The same foundation, four modules and cutover order, delivered by a senior team that has shipped 250+ products.

$44,800 fixed for all four modules, $6,600 less than four separate rebuilds; $61,000–$82,200 from 500 engineers, with multi-region, sharding and SOC 2 documentation

Live in 10–14 weeks, with a one-week pilot per module before each cutover

All four vendor configs moved: dashboards, alerts, projects and audiences

90-day warranty, four quarterly check-ins through year one, and the Helm and Terraform config in your GitHub from day one

We can run it for $14,000–$22,000 a year, all four modules: hosting, monitoring, upgrades, SLA

alliedbiz.tech/rebuild/data-ops-stack-rebuild

Allied BizTech · 15+ years · ISO 9001:2015 · 250+ products shipped. Datadog, Sentry, Segment, Twilio, Mixpanel, Grafana, Loki, Tempo, Mimir, k6, RudderStack, PostHog, ClickHouse, Splunk, New Relic, Amplitude, Bugsnag, mParticle, Hightouch, Census, Healthchecks.io, Adjust, AppsFlyer, Branch, Falco, Wazuh, PagerDuty, OpsGenie, Slack and AWS are trademarks of their owners, named here for identification only; this guide is independent and not endorsed by any of them. Figures as of November 2026.