

Pardinus AI Suite.

An AI operations layer embedded inside Checkmk. **Chat** talks to your infrastructure. **RCA** turns alerts into turns alerts into structured root-cause reports. **Cost intelligence** meters every token. One plugin, your data, plugin, your data, your premises.

× CRITICAL · DB01

CPU load — saturation 98%

14:22:08 → ongoing

EVIDENCE COLLECTED

17 items · 3 related services

livestatus + rest_api

RCA PIPELINE

collect → structure → synthesize

• 1.84s deterministic

PROBABLE ROOT CAUSE · 0.87

Backup job kernel I/O contention

• high confidence

SECTION 01 — WHY

Alerts are loud. Root causes are quiet.

Operations teams running Checkmk handle thousands of state changes a week. The signal is buried under acknowledgements, downtimes and copy-pasted runbooks — and the answer to "why did this break?" still lives inside someone's head.

73%

OF INCIDENTS

are diagnosed by tribal knowledge, not by tooling — tooling — average across mid-market enterprise ops enterprise ops teams.

4.2h

MEAN TIME TO RESOLVE

spent on triage and evidence-gathering before anyone anyone proposes a fix.

1 in 6

ALERTS REPEAT

because the previous resolution wasn't documented in a way the next operator can find.

PARDINUS SUITE

Two products. One AI operations layer.

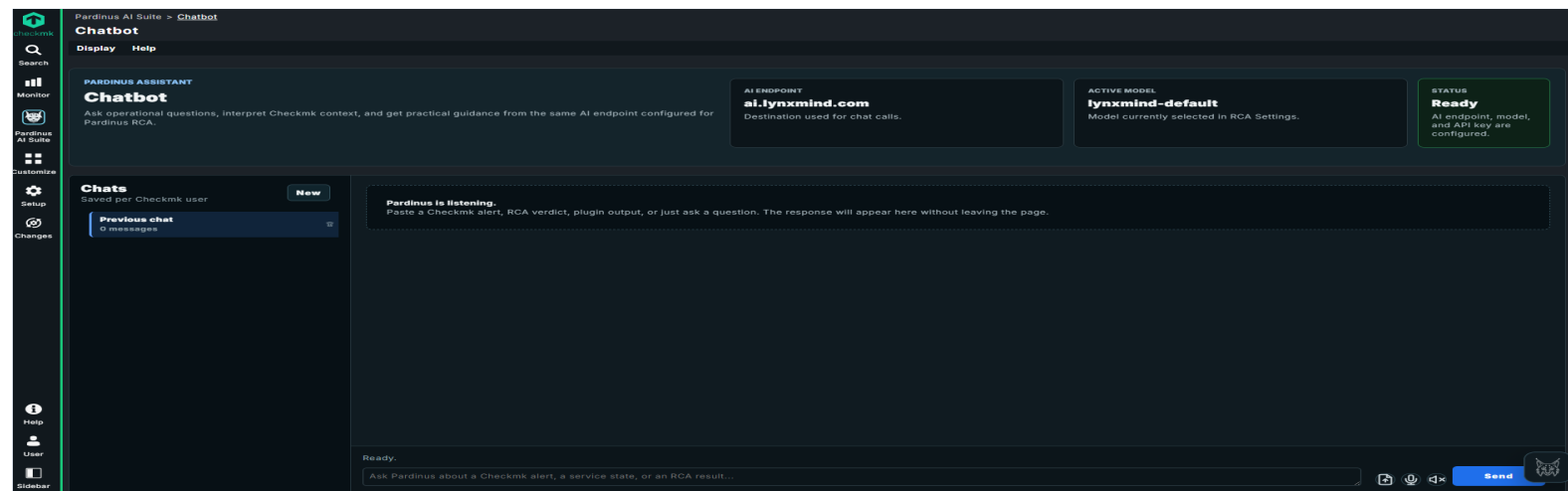
Both products share the same private AI stack, the same audit trail, and the same Checkmk integration. They install as a single MKP plugin and complement each other end-to-end.

PRODUCT 01

CONVERSATIONAL

Pardinus Chat

An AI operations assistant embedded in the Checkmk sidebar. Speaks English & Portuguese. Three modes: **Task** (act on the REST API), **Report** (investigate & synthesize), **Wiki** (learn the platform).



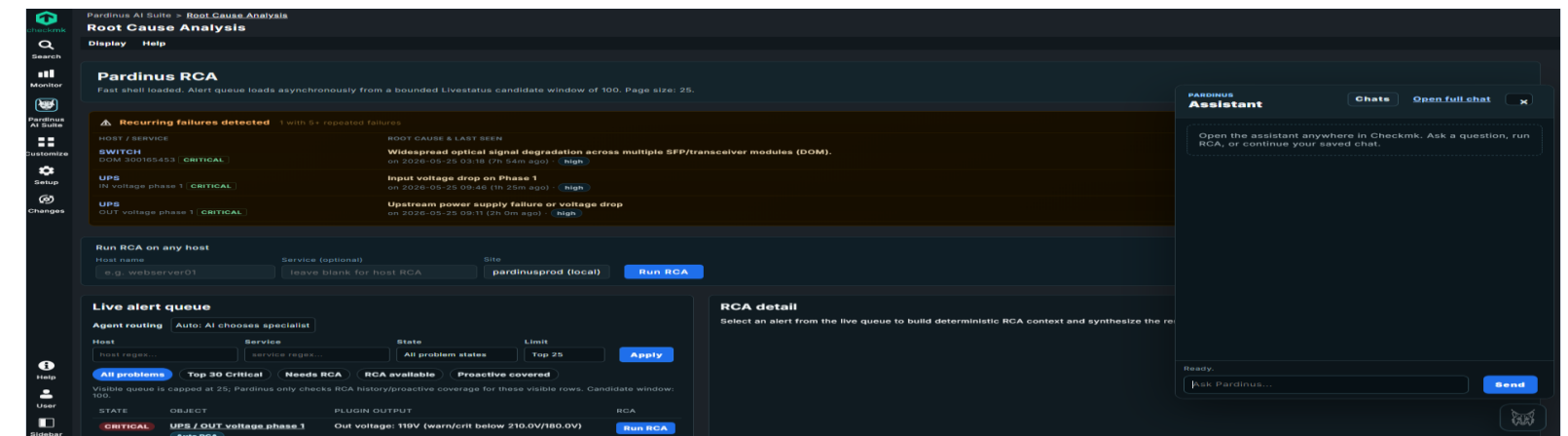
- Voice input + TTS replies
- Per-user chat history
- Confirms destructive ops
- Role-based access (Checkmk roles)

PRODUCT 02

DETERMINISTIC

Pardinus RCA

A dedicated module that runs root-cause analysis on any host or service. Evidence is built deterministically from Livestatus from Livestatus & the REST API; the LLM only explains it. Structured, auditable JSON output.

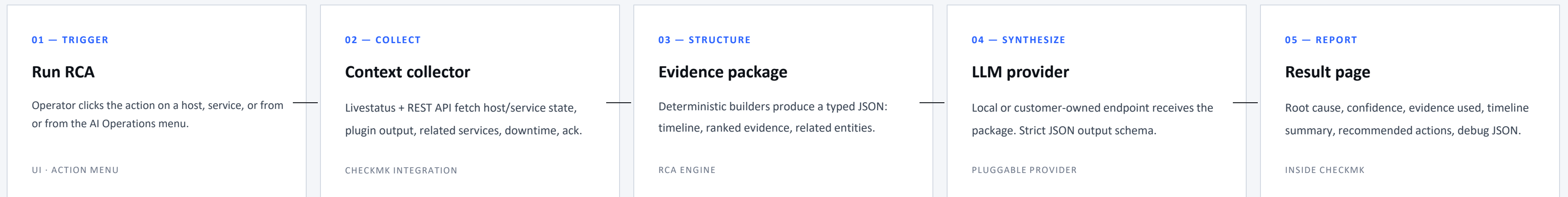


- One-click "Run RCA" action
- Specialist agents per domain
- Confidence + alternatives
- Auditable JSON post-mortems

DETERMINISTIC-FIRST PIPELINE

Evidence is built by code. The LLM only explains it.

We never let the model decide from raw alerts. Every RCA is a structured evidence package — Livestatus, REST API, topology, history — topology, history — assembled deterministically before any token leaves the site.



STRICT RULE

If evidence is missing, `missing_data` is reported. The model is forbidden from inventing it.

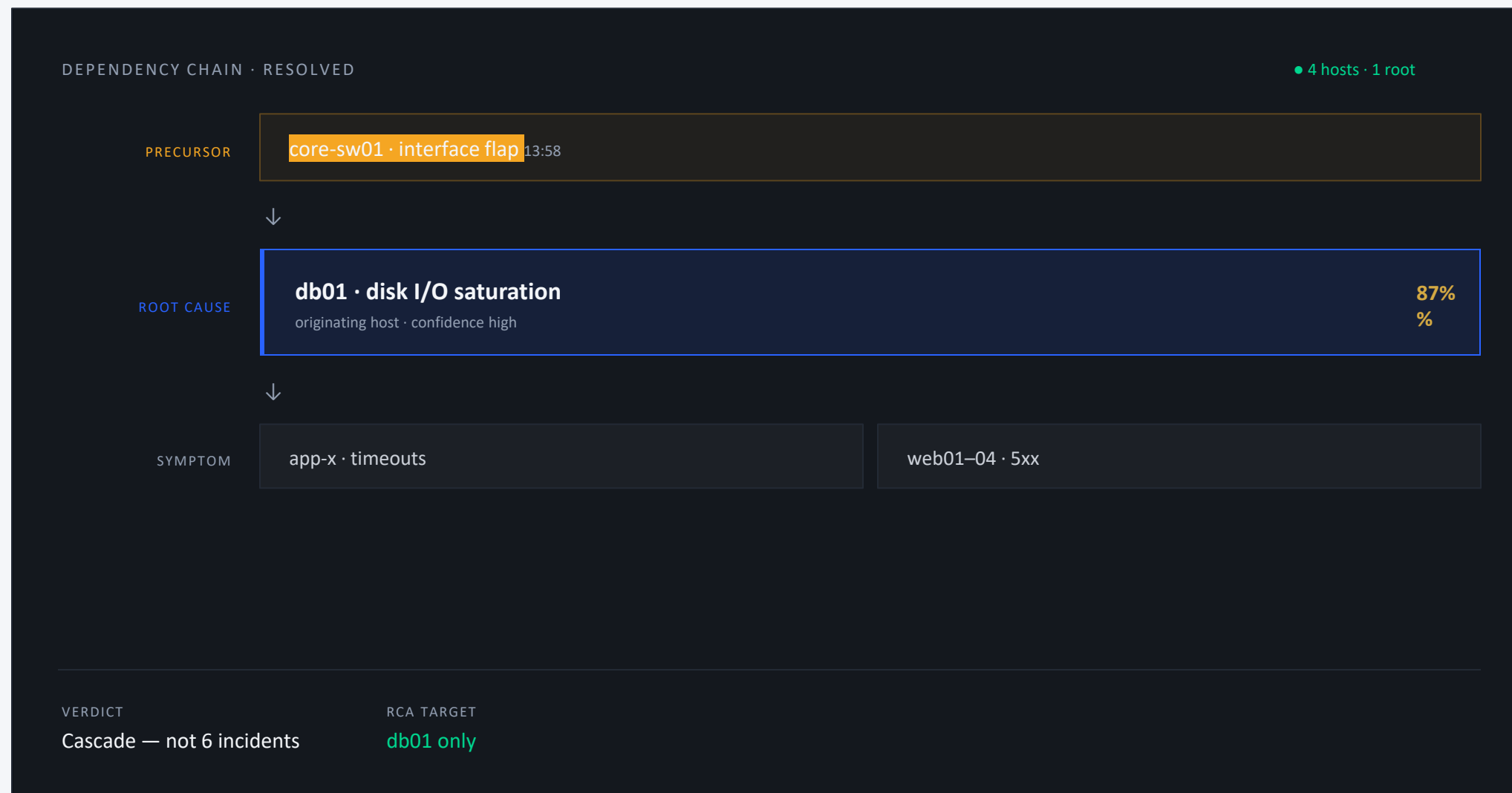
OUTPUT

Always valid JSON. `root_cause` · `confidence` · `evidence_used` · `timeline` · `recommended_actions` · `alternatives`

MULTI-HOST REASONING

200 alerts. One root. Found algorithmically.

When failure cascades across dependent hosts, Pardinus walks the dependency chain, isolates the originating host with a quantified confidence score, and classifies every event by its role — all before the LLM writes a word.



- EVERY EVENT IS CLASSIFIED
- **root_cause** — the originating failure
 - **symptom** — downstream consequence
 - **precursor** — early warning sign
 - **context** — supporting evidence

STORM VS CASCADE

Pardinus tells coincidental simultaneous incidents apart from a single cascade — and runs RCA only on the **root** the **root candidate**, with confidence scored **50–90%**.

MODULAR BY DESIGN

Clean separation. Replaceable parts.

UI	GUI extension Sidebar entry "AI Operations → RCA Analysis". Host/service action "Run RCA". MKP-packaged.	CHECKMK PLUGINS
API	Backend job controller Accepts host_name, service, site, trigger. Job model: queued · running · completed · failed.	REST
RCA	Deterministic engine Context builder, timeline builder, evidence ranker. Pure Python. Fully tested.	CORE
DATA	Checkmk integration Livestatus client (Unix socket) + versioned REST client. No core-table writes.	READ-ONLY
LLM	Provider abstraction Local · OpenAI-compatible · Azure OpenAI · customer-owned. Swappable per-tenant.	PLUGGABLE
DB	Persistence PostgreSQL preferred, in-memory fallback. Jobs and results outside Checkmk core tables.	ISOLATED

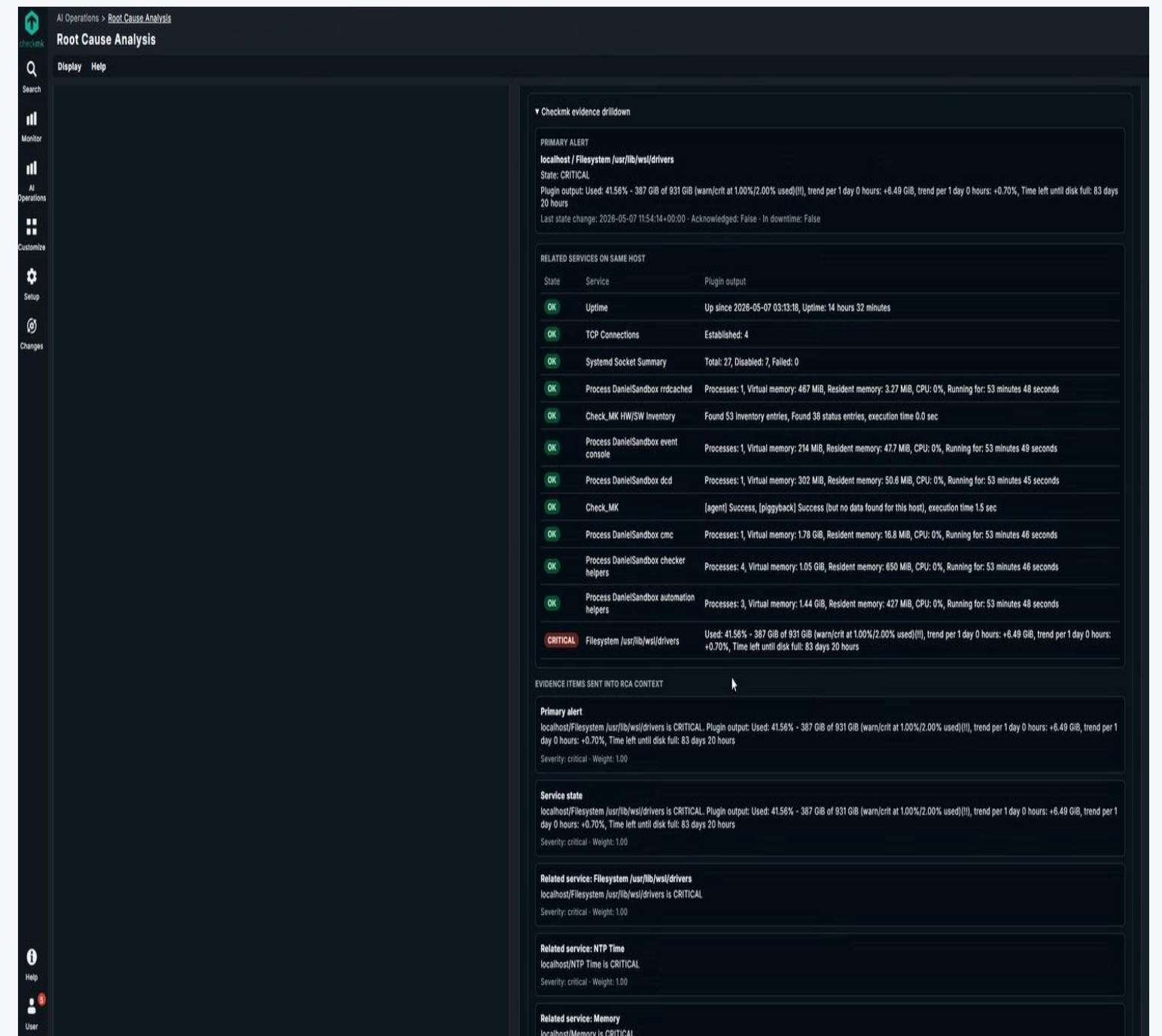


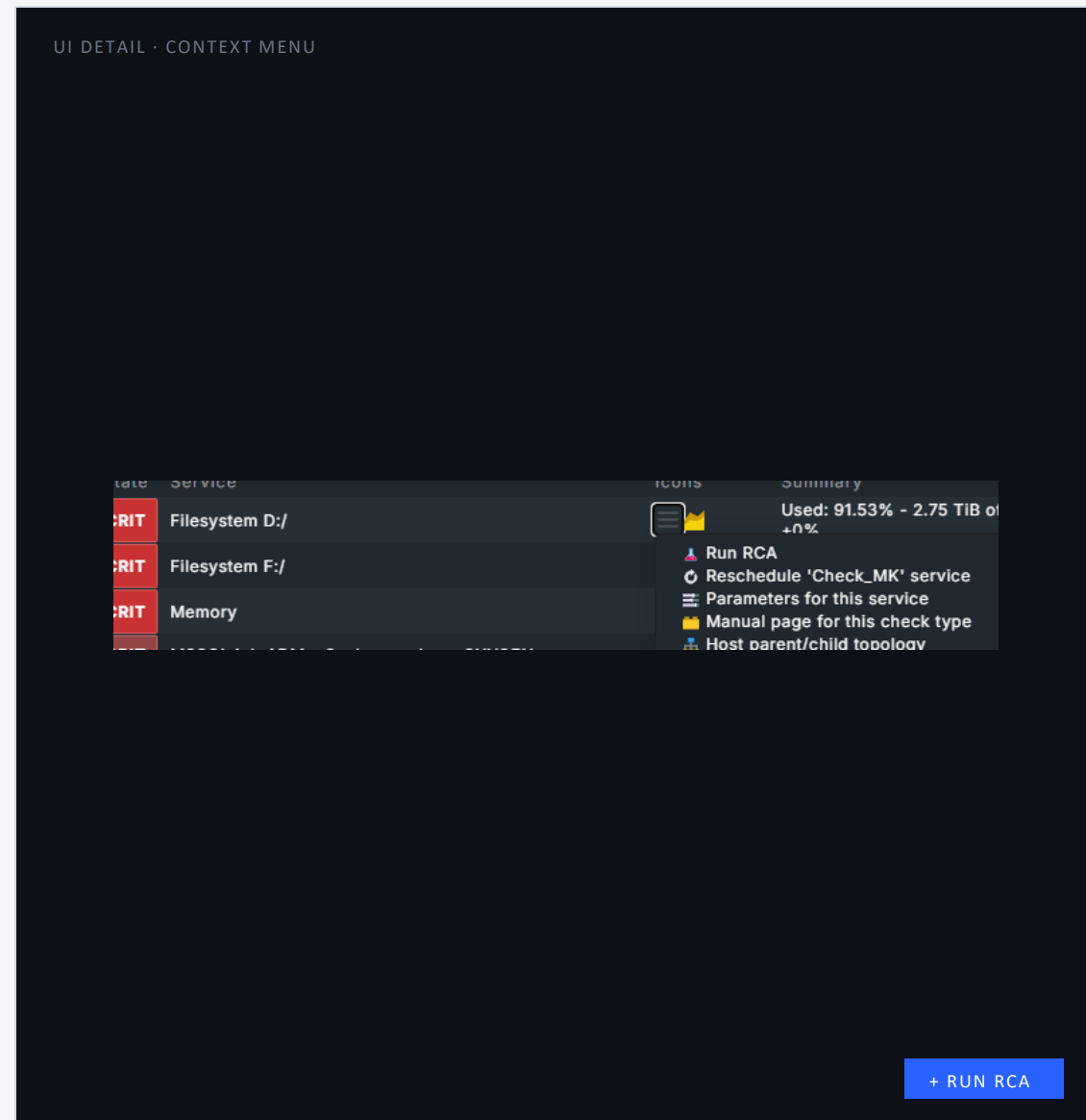
FIG.05 — Evidence package as collected from Checkmk Livestatus

INSIDE CHECKMK

Three surfaces, one workflow.

01 · TRIGGER

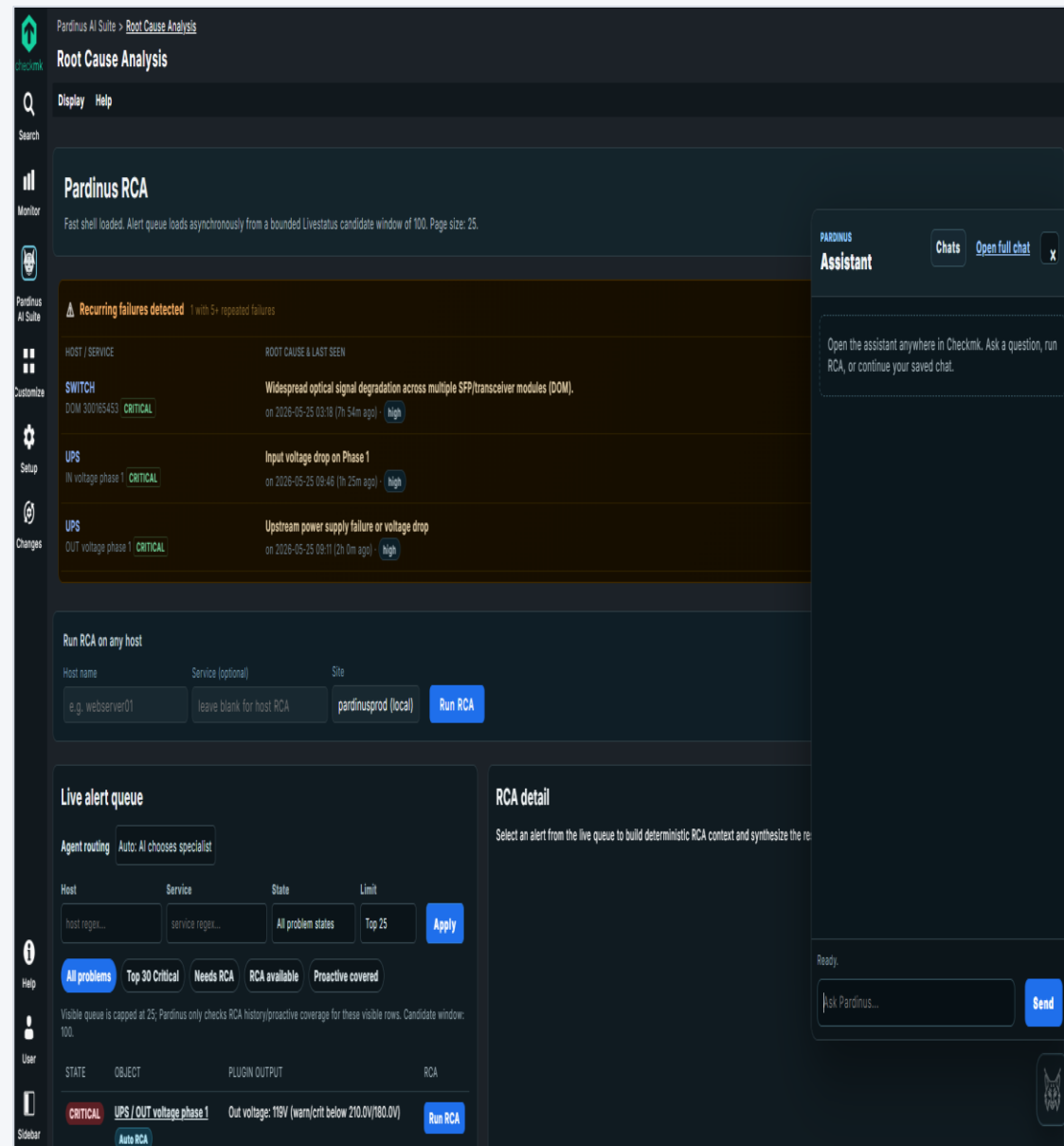
Action menu on any service



One click on a host or service opens the RCA job page. No new tooling to learn.

02 · JOB

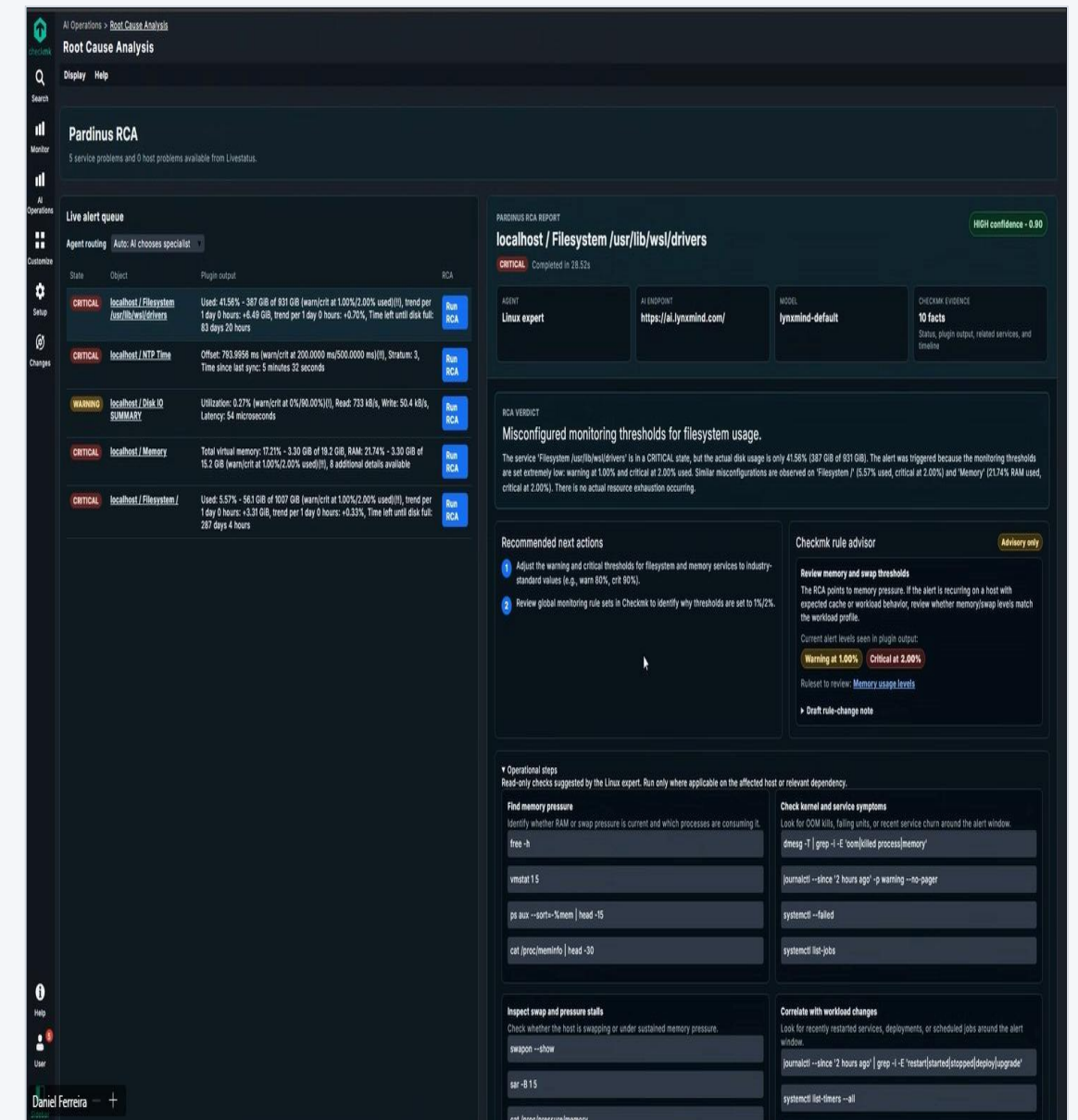
Live alert queue



Queued, running, completed, failed. Recurring failures clustered automatically.

03 · RESULT

Structured RCA report



Root cause, confidence, timeline, evidence, recommended actions, raw JSON.

ENTERPRISE-GRADE CONTROLS

Your data. Your rules.

Pardinus RCA ships with first-class data-governance controls. Configure your AI endpoint, redact endpoint, redact sensitive values before they leave the site, and set hard limits on what the LLM the LLM ever sees.

GUARDRAIL

Redact sensitive values

IP addresses, tokens, secrets stripped pre-flight.

GUARDRAIL

Deterministic fallback

If AI fails, evidence-only output.

GUARDRAIL

Provider test before RCA

Endpoint must return valid JSON.

GUARDRAIL

Confidence threshold

Below 0.75? Operator is warned.

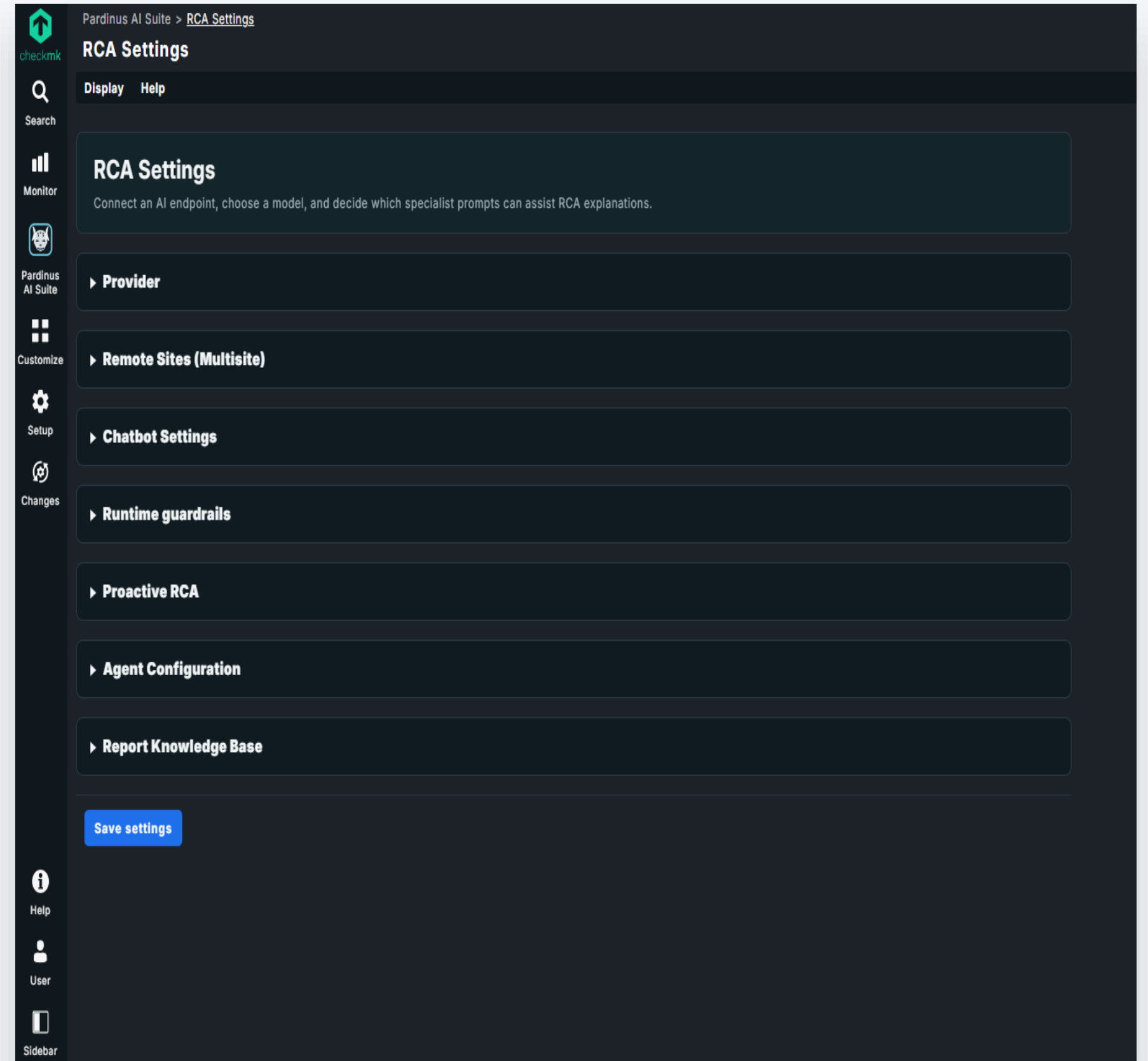


FIG.07 — RCA Settings panel inside Checkmk

TALK TO YOUR INFRASTRUCTURE

Three modes of intelligence.

Pardinus Chat handles natural language in English & Portuguese and routes each request to the right capability. No API documentation, no command syntax — just conversation.

TASK Act

Calls the Checkmk REST API on the user's behalf. Confirms destructive ops before executing. Creates hosts, moves folders, schedules downtimes, schedules downtimes, manages users.

“Create a downtime for all hosts in Lisboa for the next 2 hours.”

REPORT Understand

Autonomously investigates: queries host status, service health, performance metrics. Synthesises a structured answer in plain language.

“How are my Linux machines performing today?”

WIKI Learn

Answers Checkmk concept & configuration questions with session context. New operators ramp without reading the manual.

“What is a host tag group?”

ALSO

Voice input · per-user history · dark mode · multilingual (EN/PT) · tamper-evident audit trace on every interaction.

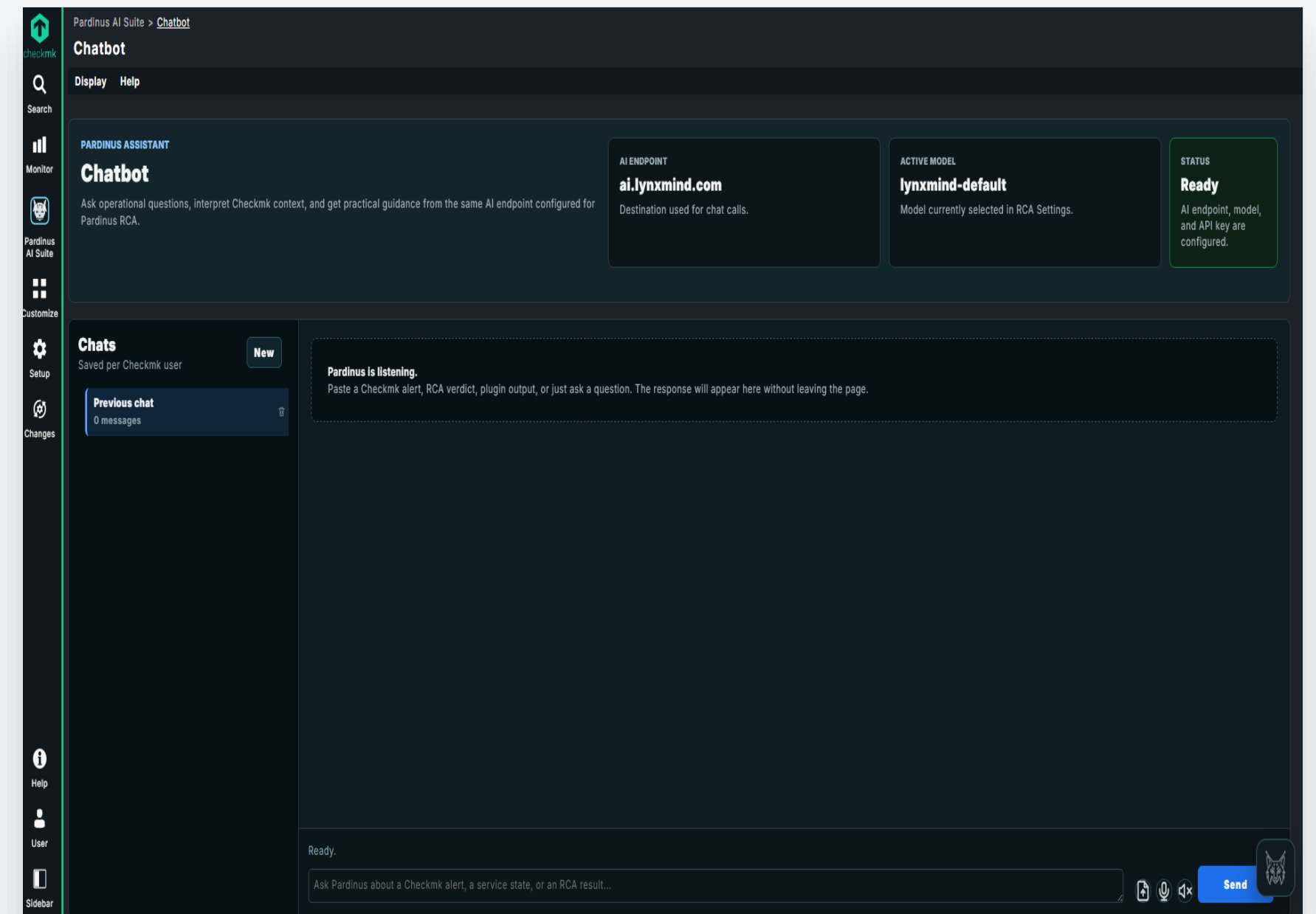


FIG.08 — Pardinus Chatbot · dedicated page inside Checkmk

RIGHT EXPERT FOR THE RIGHT ALERT

7 specialists. Auto-routed.

Each RCA is dispatched to a domain specialist by label match, keyword match, or AI routing. Lynxmind-managed prompts you don't have to prompts you don't have to maintain — plus unlimited custom agents for your bespoke services.

<p>PRESET</p> <p>Linux / Unix</p> <p>cpu · memory · filesystem · systemd</p>	<p>PRESET</p> <p>Windows</p> <p>service · eventlog · wmi · hyper-v</p>
<p>PRESET</p> <p>Networking</p> <p>ping · tcp · dns · interface · ntp</p>	<p>PRESET</p> <p>Database</p> <p>mysql · postgres · oracle · mssql</p>
<p>PRESET</p> <p>SAP</p> <p>hana · netweaver · basis · abap</p>	<p>PRESET</p> <p>Middleware</p> <p>jvm · tomcat · kafka · rabbitmq</p>
<p>PRESET</p> <p>Cloud</p> <p>aws · azure · gcp · k8s</p>	<p>CUSTOM</p> <p>Your own</p> <p>unlimited · import/export</p>

ROUTING DECISION

label match → keyword match → **AI routing** → fallback. The reason is logged on every RCA.

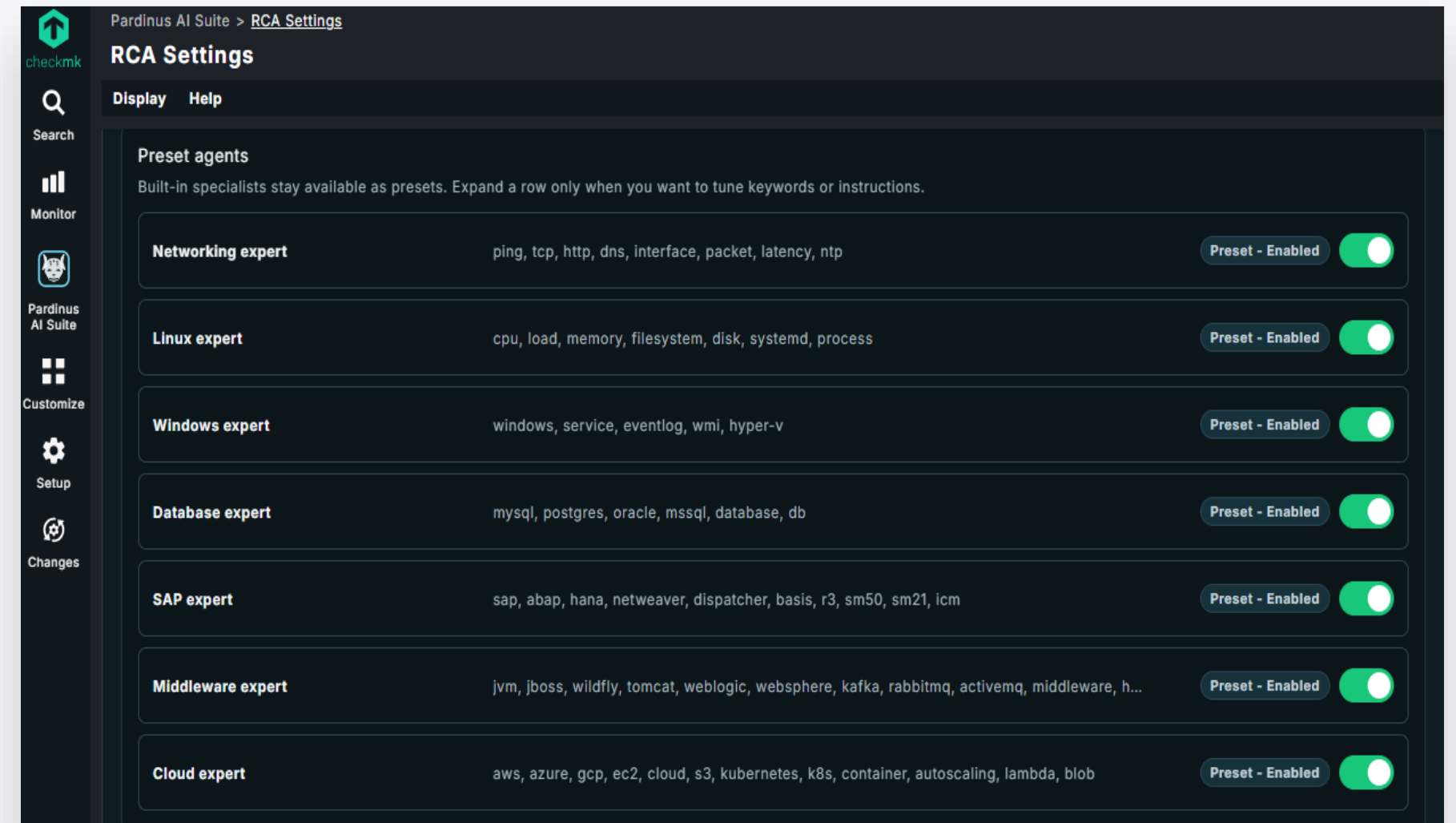


FIG.09 — Preset specialists · Lynxmind-managed prompts

ALWAYS-ON, NEVER NOISY

RCA fires before the on-call wakes up.

A background runner watches every alert and dispatches RCA when configured rules match. Rate-limited, flapping-aware, storm-protected aware, storm-protected — so AI usage stays bounded even when production isn't.

01 · RULES

Match what matters

Per-rule filtering on host, service, severity, plugin output, labels and tags. Cooldown windows
Cooldown windows prevent duplicate analysis.

host regex	service regex
severity	labels / tags

02 · RATE LIMITS

Bounded cost

Global and per-site caps, hourly and daily. AI consumption can never spike beyond the budget you set.

Global / hour	20
Global / day	100
Per-site / hour	10
Per-site / day	50

03 · STORM PROTECTION

One root cause, not 200.

When a storm hits, Pardinus picks the **earliest-state-change** alert as the probable root cause and runs RCA only on that one. Everything else is held with a reason.

PLUS

Flapping suppression · hold log with reason · amber storm banner on the alert queue

EVERY TOKEN IS METERED

You see the cost. Before it shows up.

Token Usage is a first-class page inside the suite. Provider-neutral metering, RCA reuse savings, model-level breakdowns, and a 24h / 7d / 30d window. The same data that powers the commercial model later in this deck.

TRACKED

1.3M tokens

per Checkmk site, last 7 days

AI QUERIES

170

across RCA + chatbot

SAVED BY REUSE

465.6k tokens

15 repeated alerts · zero extra cost

AVG LATENCY

22.9s

fresh RCA synthesis only

Token Usage

CONNECTED AI ENDPOINT: **http://94.62.2.149:8001**
Destination used for RCA routing and synthesis calls.

ACTIVE MODEL: **lynxmind-default**
Model currently selected in RCA Settings.

PROVIDER STATUS: **Ready**
Latest provider test passed for this exact endpoint and model.

Provider capabilities

- Token usage reported: OK ? Exact provider usage is available.
- JSON mode verified: OK ? Provider test returned valid JSON for the current configuration.
- Redaction enabled: OK ? Outbound AI payloads are sanitized.
- RCA reuse enabled: OK ? Repeated unchanged alerts can reuse recent RCA results.

Estimated cost

€0.6819

Mock pricing is enabled for product planning. Replace these placeholder rates later with your real internal LLM allocation or customer-specific model prices.

- INPUT RATE: €0.5000 / 1M tokens
- OUTPUT RATE: €1.50 / 1M tokens
- INPUT TOKENS PRICED: 1.2M
- OUTPUT TOKENS PRICED: 44.8k
- PRICED RECORDS: 75
- NEEDS TOKEN DATA: 95

Time window: 24h | **7d** | 30d

AI QUERIES: 170 (in selected 7d window)

TOKENS TRACKED: 1.3M (reported by provider when available)

TOKENS SAVED BY RCA REUSE: 465.6k (15 repeated alerts avoided another AI call)

AVG RCA LATENCY: 22.9s (fresh synthesis only)

FIG.11 — Token Usage dashboard · provider-neutral metering

OPEN ARCHITECTURE

API-first. Wire it into **anything**.

Every Pardinus capability is exposed through a versioned REST API. Trigger RCA from your incident pipeline, fetch reports into your ticketing tool, attach evidence to a Slack thread — without touching the Checkmk UI.

REST ENDPOINTS		
POST	/api/v1/rca/run	trigger an RCA job
GET	/api/v1/rca/{job_id}	poll status + result
GET	/api/v1/rca/{job_id}/report.json	structured report
GET	/api/v1/rca/{job_id}/report.md	markdown for tickets
POST	/api/v1/chat/query	conversational ops
HOOK	webhooks · job.completed	push, don't poll

DROP-IN FOR THE TOOLS YOU ALREADY RUN			
n8n	PagerDuty	ServiceNow	Jira
Slack	MS Teams	Opsgenie	Zapier
Ansible	Webhooks	Email / SMTP	your stack

```

TRIGGER RCA FROM ANYWHERE

curl -X POST https://checkmk.lynxmind.ai/site/pardinus/api/v1/rca/run \
-H
"Authorization: Bearer $SPARDINUS_TOKEN" \
-H
"Content-Type: application/json" \
-d
'{
  "host_name": "db01",
  "service_description": "CPU load",
  "trigger_source": "n8n",
  "callback_url": "https://n8n.lynxmind.ai/webhook/rca-done"
}'
→ 202 Accepted { "job_id": "rca_a8f3...", "status": "queued" }
    
```

EXAMPLE · N8N WORKFLOW ALREADY IN PRODUCTION

Auto-attach RCA to every new incident

Checkmk CRIT → n8n trigger → **POST /api/v1/rca/run** → await webhook →

create ServiceNow incident + RCA report as attachment

TIME TO SET UP	LINES OF CODE	AUTH
~15 min	0	Bearer

CHANGE INTELLIGENCE

It knows when you ship.

Pipe CI/CD events — or any third-party signal — into Pardinus through an inbound webhook. Every deploy, config change or release becomes context, so a symptom spike is traced straight back to the change behind it.

```
POST /pardinus/events
```

- CI/CD pipelines
- config changes
- feature flags
- any 3rd-party event

PARDINUS · EVENT CORRELATION

DEPLOYMENT RECEIVED
app-x v2.4.1
via inbound webhook · 14:02 · source: CI/CD

CORRELATED SIGNALS · POST-DEPLOY

HTTP sessions	▲ +320%
User experience · p95	4.8s
Memory usage	▲ 94% 94%

Correlated to deploy **app-x v2.4.1** — not a coincidental incident.

DIFFERENTIATORS

Four things nothing else gives you.

01

Native to Checkmk.

Not a side dashboard. Lives inside the menu, the action bar and the views the operator already uses. MKP-packaged, version-pinned per site.

02

Deterministic before AI.

Evidence is built by code, not by the model. The LLM only narrates a verified verified package — so confidence scores are real, not vibes.

03

Bring your own model.

Local LLM, OpenAI-compatible, Azure, or customer-owned. Data sovereignty is the sovereignty is the default, not an upsell.

04

Auditable output.

Every RCA emits typed JSON: evidence used, alternatives considered, missing considered, missing data, timeline. Compliance and post-mortem ready. ready.

USE CASES

From 3am pages to weekly reviews.

A — On-call

First-responder triage

RCA fires automatically on critical state. The on-call engineer wakes up to a structured report, not a wall of plugin output.

B — Tier-1

Service-desk escalation

L1 attaches the RCA report to the ticket before escalating. L2 gets evidence, not "it's not "it's broken".

C — Capacity

Recurring incident clustering

Pardinus surfaces patterns across past RCAs — "this CPU spike has happened 12 times happened 12 times on db01 in 30 days".

D — Compliance

Post-mortem evidence

Every RCA is auditable JSON. Feeds directly into ITIL post-incident reviews and SOC 2 and SOC 2 evidence collection.

E — MSP

Multi-tenant RCA

Per-customer LLM provider, per-customer billing. MSPs can resell RCA as a managed managed service line.

F — Knowledge

Living runbook

Every resolved RCA becomes searchable context for the next one. Tribal knowledge knowledge becomes structured data.

OPEN BY DESIGN

Your model. Your sites. Your rules.

Pardinus runs against any OpenAI-compatible endpoint and analyses every Checkmk site you operate — from a single socket to a socket to a distributed, multi-site estate. Nothing is locked to one vendor.

SUPPORTED LLM PROVIDERS - ALL PRODUCTION

OpenAI	gpt-4o · gpt-4-turbo	● PROD
Azure OpenAI	enterprise deployment	● PROD
Ollama	LLaMA · Mixtral · local	● PROD
vLLM	standalone inference server	● PROD
Open WebUI	self-hosted OpenAI wrapper	● PROD
Any OpenAI-compatible	your endpoint, your keys	BYO

DISTRIBUTED MONITORING

RCA across every site

Cross-site analysis over Livestatus — Unix socket and TCP. A per-site collector reaches remote sites; the Remote Sites panel lives in Settings.

- livestatus · unix
- livestatus · tcp
- per-site collector

PRIVACY GUARDRAIL

Redacted before the LLM

IPs, hashes, UUIDs and secrets are masked pre-flight — whichever provider you point at. Your sensitive data never reaches the never reaches the model.

SECTION 02 — COMMERCIAL

Annual subscription. Sized to your estate.

One predictable annual fee, priced by the number of monitored Checkmk services. You bring your own LLM endpoint — Pardinus never sees your traffic, and your data never leaves your premises.

WHAT'S INCLUDED

Everything, in one license.

- + Pardinus Chat — three modes
- + Specialist agents + custom agents
- + Integration API + webhooks
- + Updates & support included
- + Pardinus RCA — deterministic engine
- + Proactive RCA + storm protection
- + Cost-intelligence dashboard
- + Unlimited operators

BRING YOUR OWN LLM

Data sovereignty by default.

from €3,200

PER YEAR · SIZED TO YOUR CHECKMK

- + Any OpenAI-compatible endpoint · local, on-prem, or your cloud
- + Data never leaves your premises
- + Fits regulated & air-gapped environments

SIZED TO YOUR CHECKMK

Pricing scales with you. Same model as Checkmk.

Annual subscription priced by the number of monitored services. Bring your own OpenAI-compatible LLM endpoint — local, on-prem, or your own cloud.

MONITORED SERVICES	ANNUAL SUBSCRIPTION BYO LLM ENDPOINT	TYPICAL ESTATE
7,000	€1,600 per year	Small team · single site
12,000	€2,400 per year	Growing infrastructure
18,000 POPULAR	€3,200 per year	Mid-market · ~3,000 hosts
30,000	€4,700 per year	Large enterprise
60,000	€7,800 per year	Multi-datacenter
100,000	€11,500 per year	Global estate
200,000	€19,200 per year	Service provider · MSP

* > 200,000 services — custom quote. Multi-site, multi-tenant, and MSP partner discounts on request.

ROI · 18,000-SERVICE SITE · 12-MONTH HORIZON

A typical mid-market customer.

~3,000 hosts · ~40 critical incidents/month with deep RCA · auto-triage on every CRIT · chatbot for daily ops.

COST — ANNUAL

License · Pardinus Suite (BYO LLM)	€3,200
AI endpoint	customer-owned
Onboarding (one-off)	€2,500
Year 1 total	€5,700

VALUE — ANNUAL

MTTR reduction	-38% %
Engineer hours saved	~720 h
Engineer hour value	€85
Time savings	€61,200
Avoided major-incident escalation	€18,000
Net value · Year 1	€73,500

PAYBACK PERIOD

≈ 1 month · **13.9x ROI** in year one

● EARLY ACCESS

30% off your first-year license — founding customers only.

Sign before **30 June 2026** and lock the discount for the full initial term. Pilot fee fully credited.

YEAR-1 LICENSE

€3,200

€2,240

LET'S RUN THE FIRST RCA ON YOUR STACK.

knowing

30-day pilot on your Checkmk site. We deploy the MKP, run RCA on your top 10 noisiest services, and hand you a written report. Pilot fee credited against year-one license.

TALK TO US

TRY THE DEMO

PILOT TIMELINE