



ASE 2026

OpsLens: Runbook-Free Root Cause Analysis via Multi-agent Collaboration and Tool Management

Haiyu Huang¹, Jiewei Lyu², Gou Tan², Pengfei Chen², Zhihan Jiang¹, Guangba Yu¹, Penghui Mi³, Fuqiang Wang³, Jian Han³, Chongkang Tan⁴, Michael R. Lyu¹

¹ The Chinese University of Hong Kong ² Sun Yat-sen University

³ Huawei ⁴ Independent Researcher



香港中文大學
The Chinese University of Hong Kong



“Service latency increased in region A.”

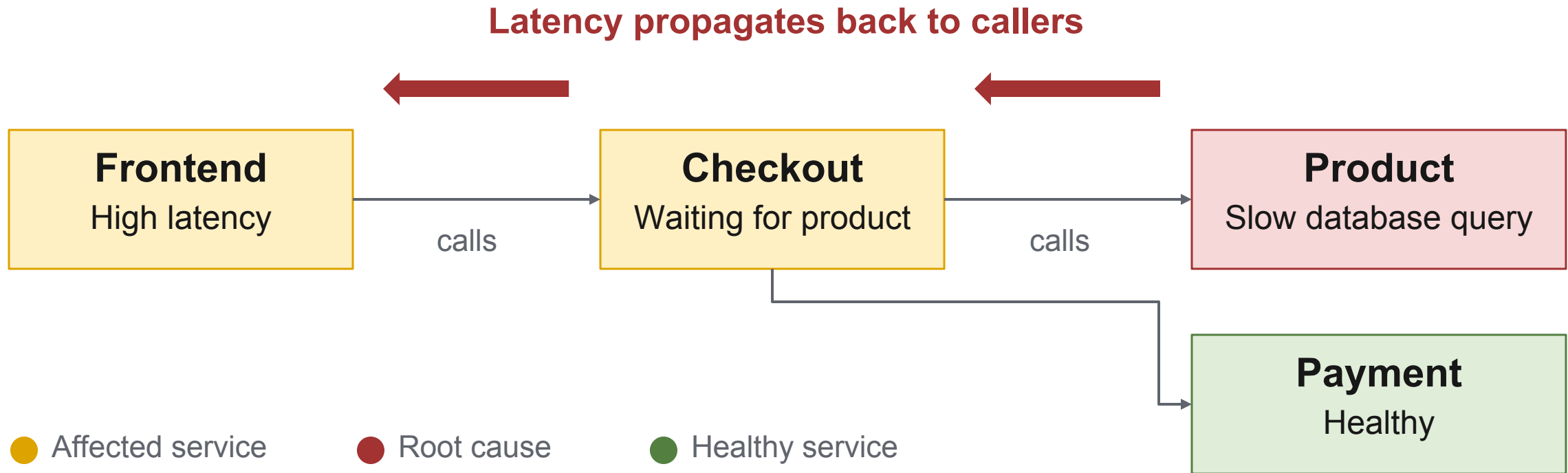
The alert describes the symptom.
The engineer needs to find the underlying cause.

RCA question	What the diagnosis should identify
When?	The onset of the failure
Where?	The component responsible for the failure
Why?	The underlying fault or operation

➤ Cascading symptoms across services

ASE 2026

One slow product query can delay an entire request.

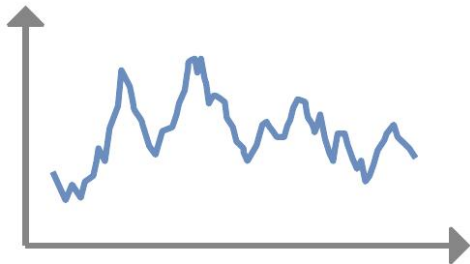


RCA separates the initiating fault from its propagated symptoms.

System-generated: collected automatically during operation

Once instrumentation is configured, the system produces evidence of its behavior.

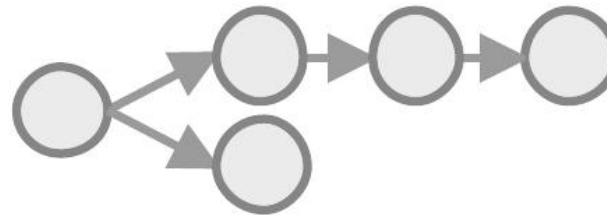
Metrics



Measurements over time

Narrow the abnormal
time window

Traces



Request paths across
services

Identify suspicious
services and spans

Logs

INFO [host-1]
Directory
Deployment has
finished in 9 ms

Events within a component

Inspect errors and
operations in a service



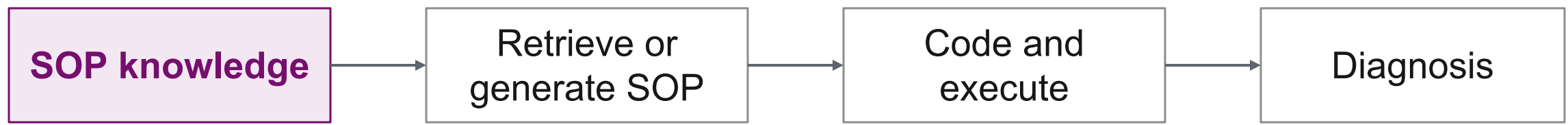
A runbook records troubleshooting steps for a recurring operational problem.

	System-generated telemetry	Human-generated knowledge
Origin	Collected during operation	Curated by engineers
Purpose	Evidence of what happened	Guidance on what to check
Upkeep	Instrumentation and collection	Authoring and manual updates

Illustrative latency runbook

Check metrics, inspect traces, then read the relevant service logs.

Curated knowledge can become a dependency of the investigation workflow.



Flow-of-Action ablation with GPT-3.5-Turbo

Setting	Localization accuracy
Full method	54.22%
Without SOP knowledge	8.56%

Removing the knowledge also makes its SOP flow ineffective.

The agent falls back to ReAct.

The reported degradation is specific to this method and ablation setting.

➤ Human-generated data is not always available

ASE 2026

Huawei's Platform A: 23 systems in two regions

Region	Complete telemetry	Maintained RCA artifacts
A	15 / 17	3 / 17
B	5 / 6	1 / 6
Overall	20 / 23 (87.0%)	4 / 23 (17.4%)

Telemetry: metrics, traces, and logs.

Artifacts: structured incident reports or prebuilt diagnostic tools, as assessed by SREs.

Runbook-free RCA must work from available telemetry.

Creating the knowledge takes SRE time

Runbook creation for 50 historical incidents averaged 3.42 person-hours per incident.

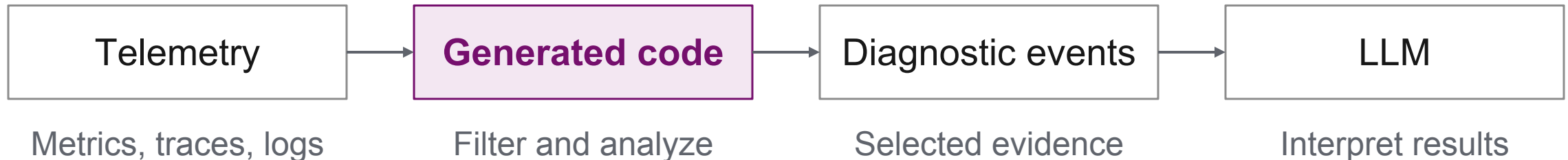
Engineers collect evidence and organize reusable troubleshooting guidance.

Directly prompting with raw data

Large volumes and specialized structures can exceed the context window.
Truncation may discard relevant evidence.

Analyzing data through code

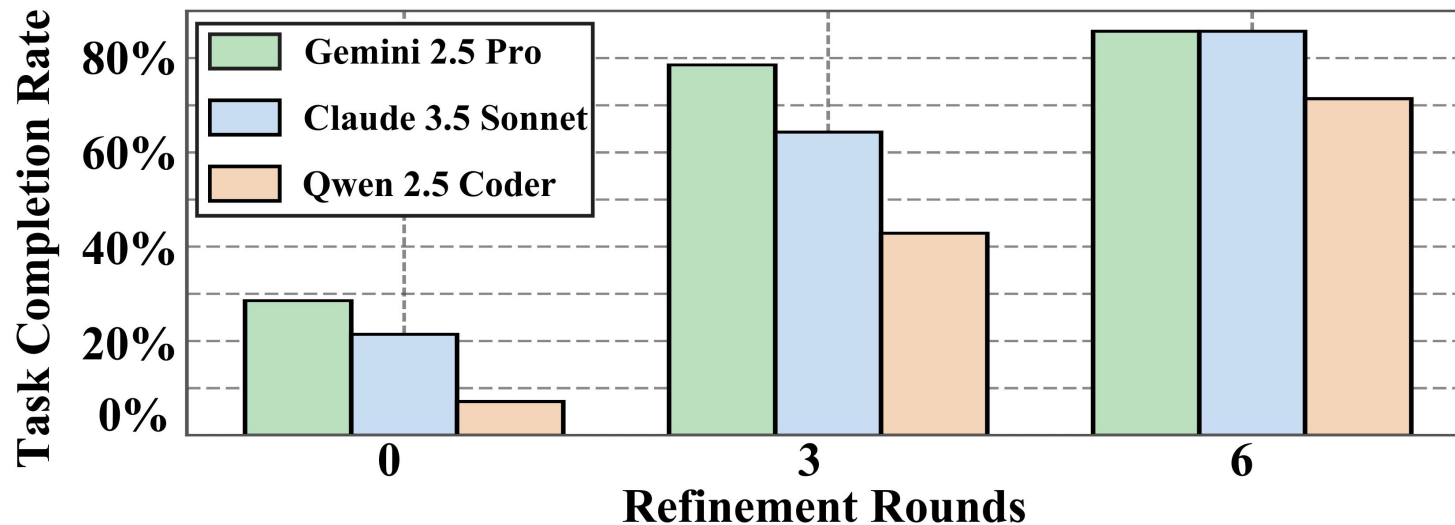
Code filters and analyzes data outside the prompt. The LLM interprets compact diagnostic results.



The next question: does the generated analysis find useful evidence?

➤ Generated tools need quality management

A tool can run successfully and still miss the diagnostic signal.

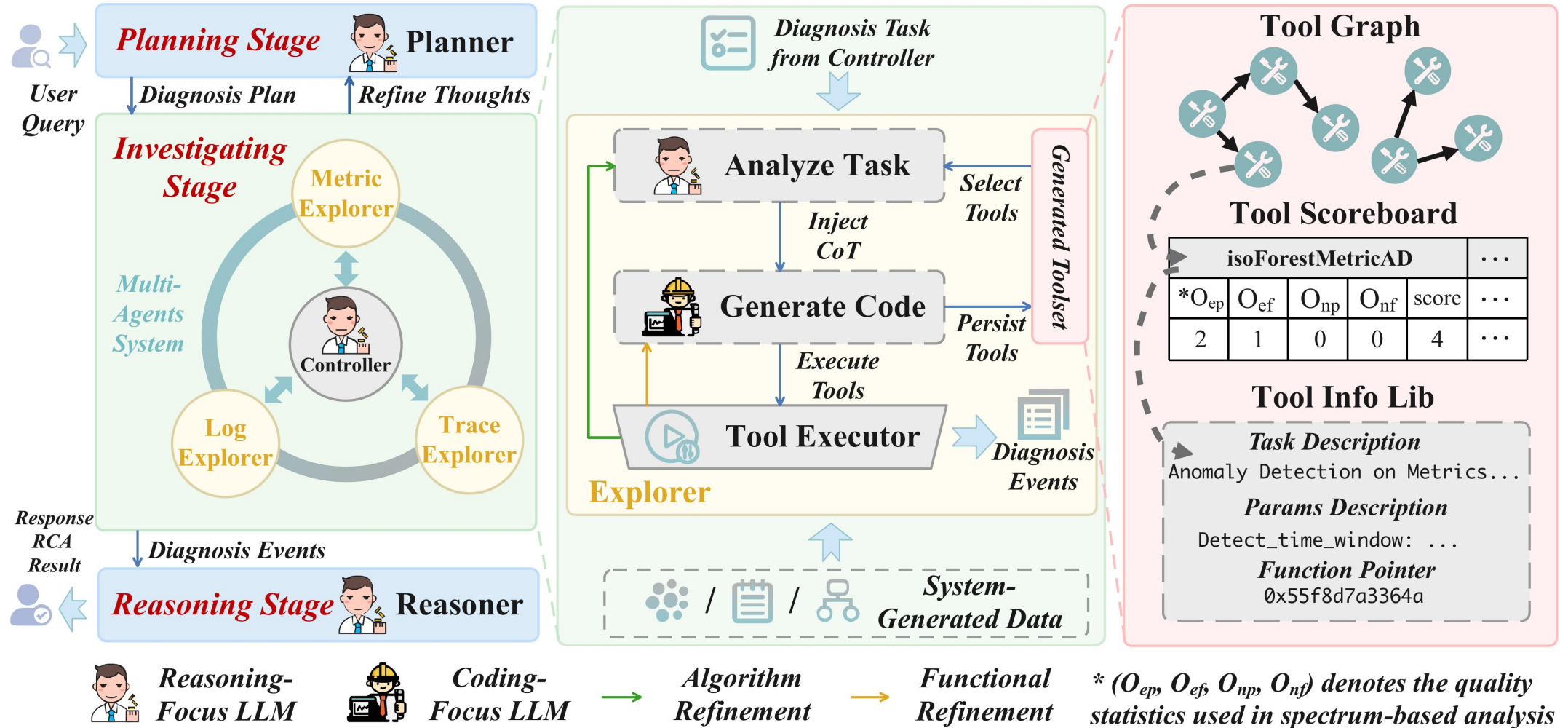


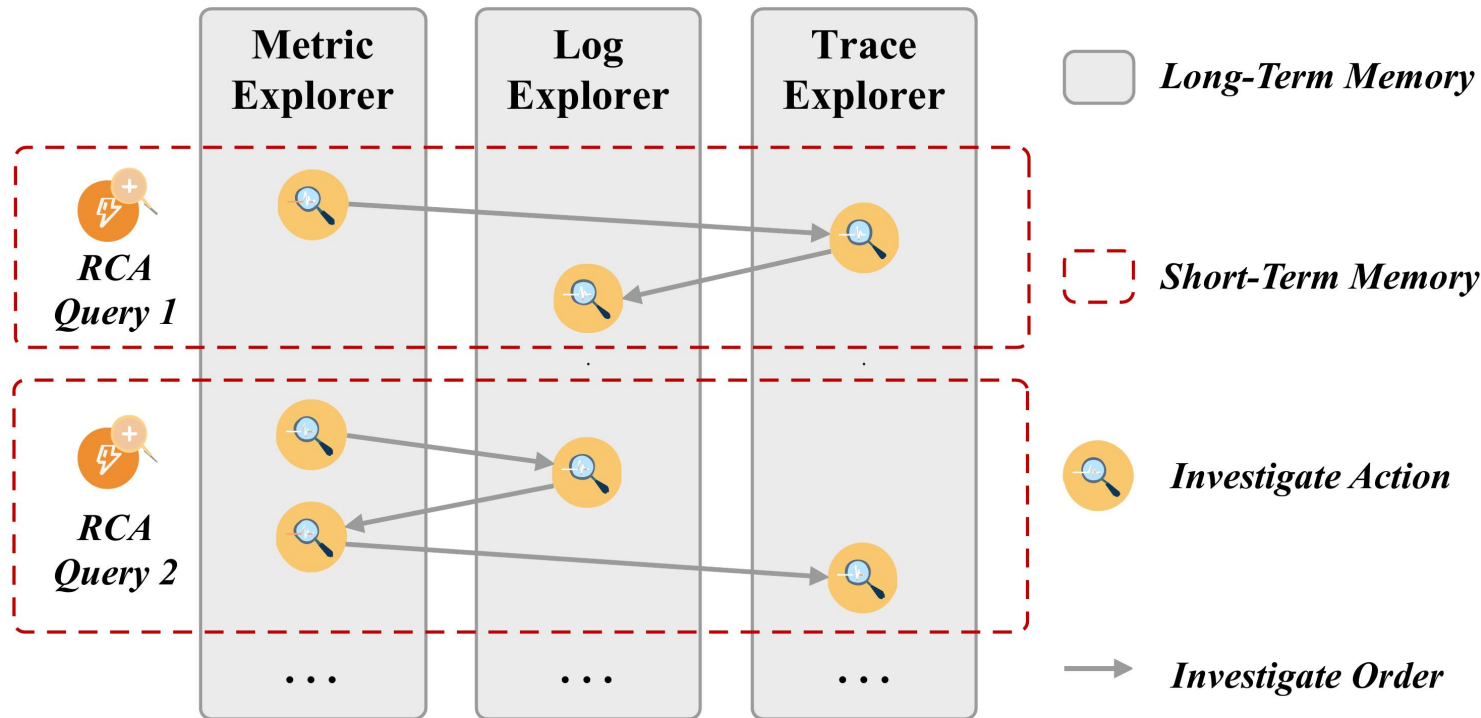
In the 14-subtask study

- First runnable tools completed fewer than 30% of tasks.
- After six SRE-guided refinement rounds, completion reached 80.9%.

System changes can also make previously useful tools unsuitable.

Quality management must assess outputs, refine the analysis, and guide which tools to select and reuse.





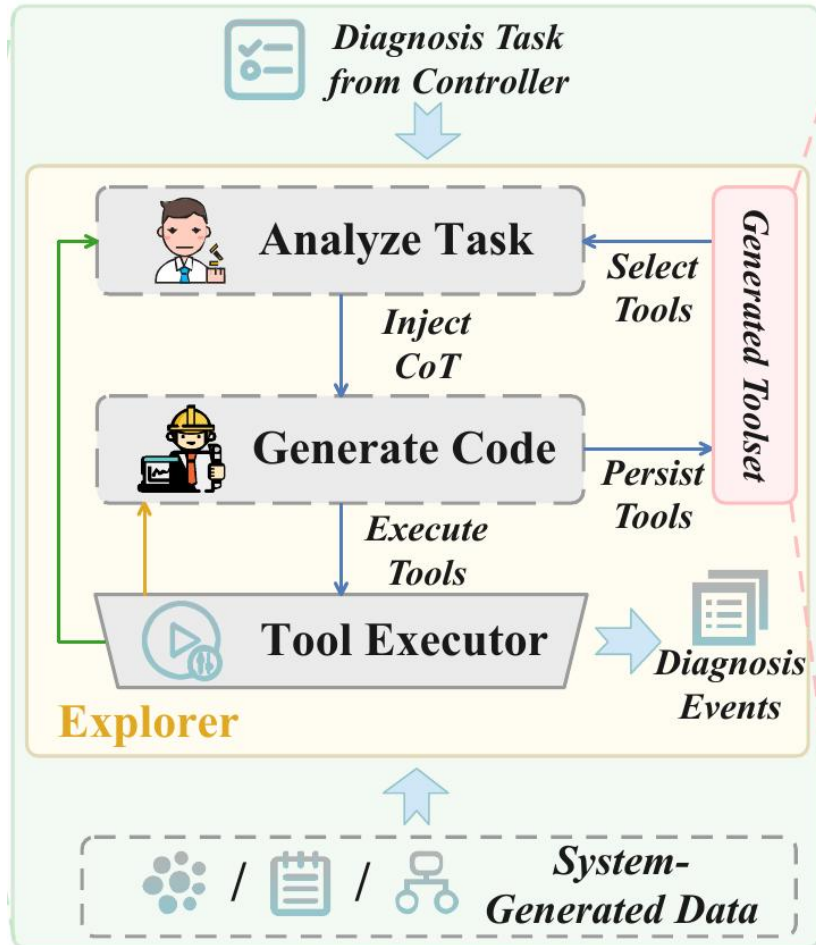
Within an incident

Explorers share final investigation outcomes.

Across incidents

Each modality keeps its own investigation experience.

Shared event fields: source, timestamp, component, and description.

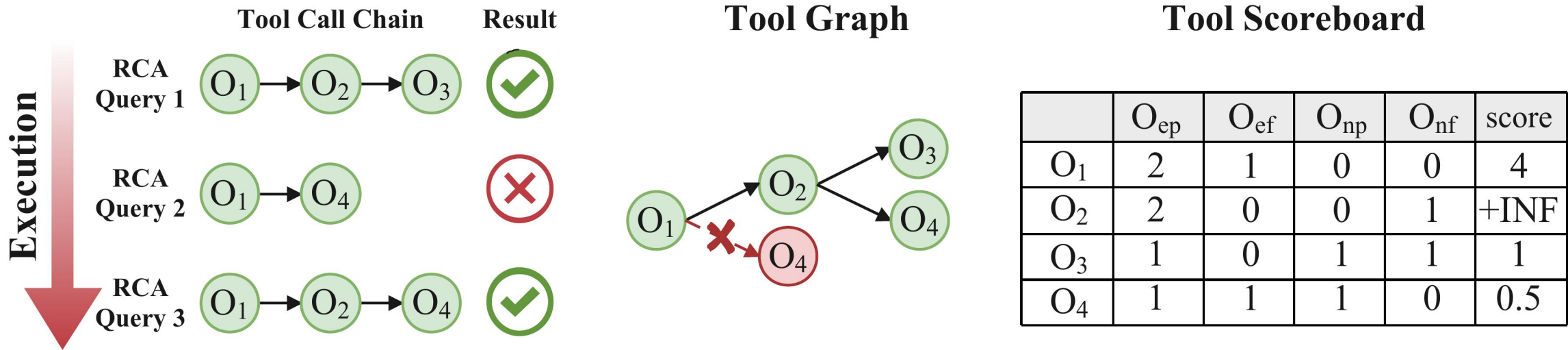


Generation

- A reasoning role designs the analysis.
- A coding role implements it.

Execution feedback

- Runtime checks find code errors.
- Task-specific constraints check result quality.
- Failed checks trigger code or analysis refinement.



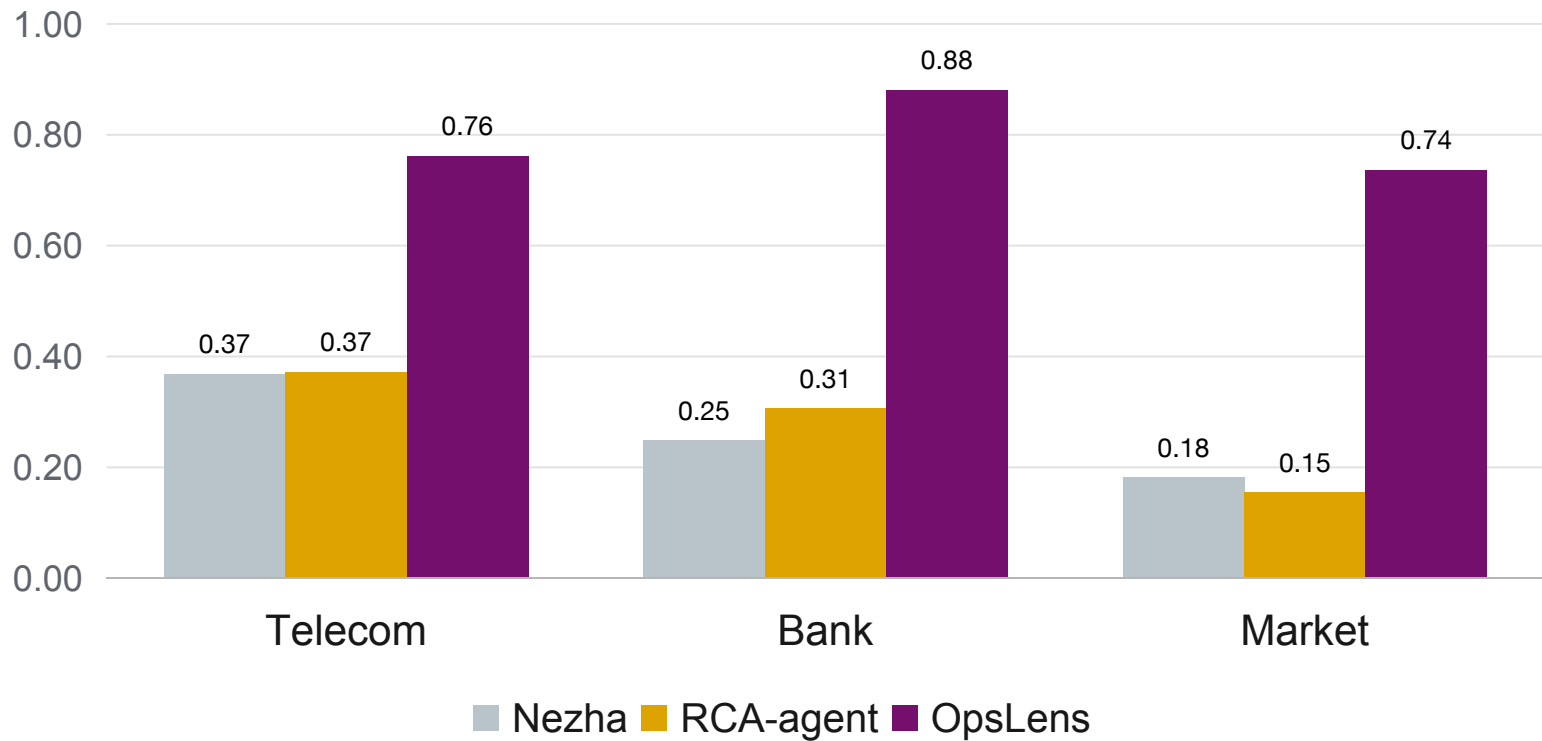
Stored information	How it helps select the next tool
Tool description	Match the analysis task and required parameters
Quality scoreboard	Rank tools by historical RCA outcomes
Invocation graph	Suggest tools used next in successful investigations

- OpenRCA: 335 failure cases from Telecom, Bank, and Market.
- Queries require root cause elements such as time, component, and reason.

Measure	Meaning
C@1	All required elements in the top-ranked answer
Score rate	Fraction of available element points earned

Online evaluation protocol

Stores start empty, then retain tools and memory across cases.
Completed-case labels update scores only for later cases.



Overall benchmark results

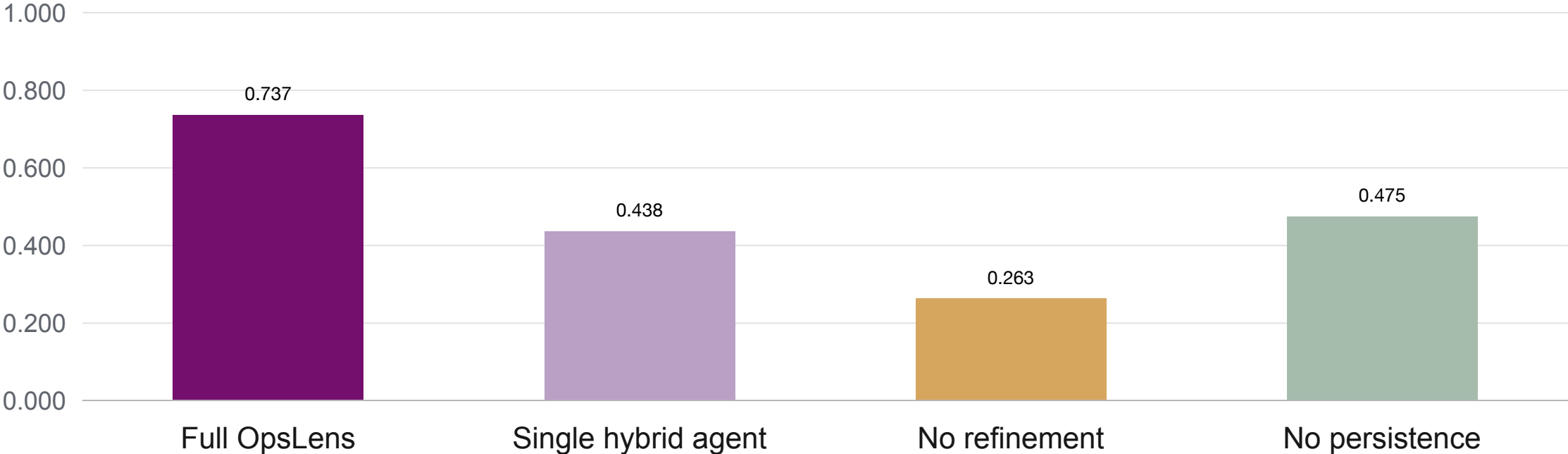
Metric	OpsLens
Score rate	0.80
C@1	49.85%

Score rate gives credit for individual root cause elements.

With the same Claude 3.5 Sonnet backbone, OpsLens achieves 1.79× score rate on 55 sampled cases. The full comparison uses different backbones.

Score rate: Table 3. Main comparison uses different backbones. Matched model: §4.4

Removing either collaboration or tool management reduces score rate.



The largest reduction on Market occurs when refinement is removed.

	General fault diagnosis	Network diagnosis
Dataset	31 diverse cases	259 alerts, 10 categories
Model	Gemini 2.5 Pro	GPT-4o
C@1	64.52%	100%
Additional result	83.87% C@3	29.6 seconds per case

The network result comes from a bounded orchestration workflow.

It uses internal inspection tools and SRE-provided category guidance.

Experience from four systems at Huawei's Platform A

Latency budgets

Warm up and reuse tools. After 5–6 cases, steady-state diagnosis averages 173–220 s.

SRE trust

Expose the investigation evidence so engineers can verify the conclusion before taking action.

Data sensitivity

Use internal models for sensitive tasks and external models for reasoning-heavy tasks.

System migration

Standard telemetry helps transfer. Configure data paths and field schemas for each system.

- Runbook-free RCA starts from the telemetry a system already produces.
- Modality-specialized agents connect investigation evidence.
- Execution feedback and quality-aware reuse improve diagnostic tools.

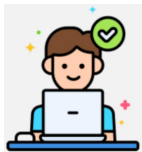
Thank you

Contact me: hyhuang25@cse.cuhk.edu.hk

Questions?

Alert: service latency increased in region A, 08:00–12:00.

Investigation	Evidence collected
Metrics	First execution returns no diagnosis events. One refinement finds the 08:15–08:20 anomaly window.
Traces	Candidates: svc-728 and svc-041. Symptoms may also appear in affected services.
Logs	A full table scan on svc-728: <code>SELECT * FROM t_collect</code>



Diagnosis: slow SQL on svc-728

The SRE reviews the timing, component, and log evidence.

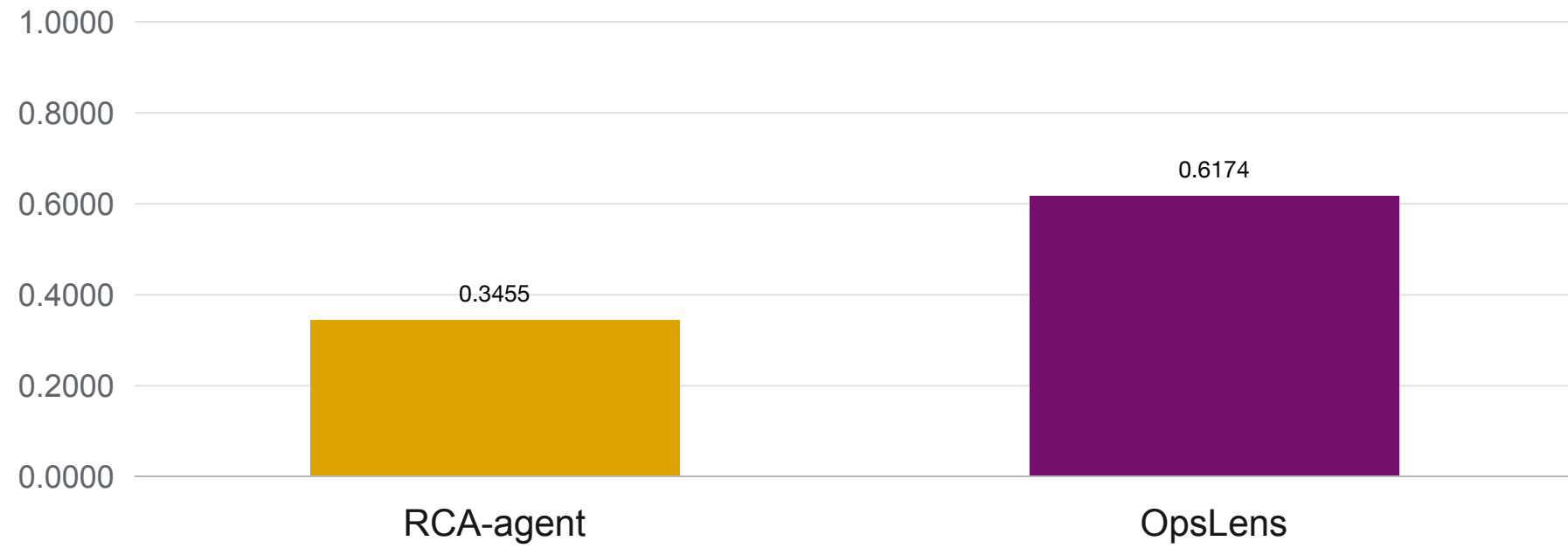
➤ Backup: scope and assumptions

Setting	Assumptions and retained knowledge
Core RCA workflow	Telemetry access, data schemas, general modality knowledge
OpenRCA evaluation	Empty initialization, then persistent tools and memory
Scoreboard updates	Completed-case labels guide later cases
Network workflow	Internal tools plus SRE category and remediation guidance

Runbook-free initialization still requires data access and feedback.

➤ Backup: same-backbone comparison

Both systems use Claude 3.5 Sonnet on the same 55-case sample.



OpsLens achieves 1.79× score rate in this matched-backbone study.

➤ Backup: tool selection and quality checks

Similarity ≥ 0.6 , top 5 by quality, plus graph successors

score(O) = $ep^2 / (ef + np)$

Count	Historical cases
ep	Successful cases involving the tool
ef	Failed cases involving the tool
np	Successful cases without the tool

The score measures association with successful RCA outcomes.

➤ Backup: latency, cost, and limits

Measure	Reported result
Steady-state diagnosis	173–220 seconds average
Generating a new tool	260 seconds average
LLM calls per task	14.5 average
Tokens per task	52.67k average

- Initial tool generation can exceed a strict latency budget.
- Evidence comes from one provider. Incomplete telemetry and heuristic checks remain limits.