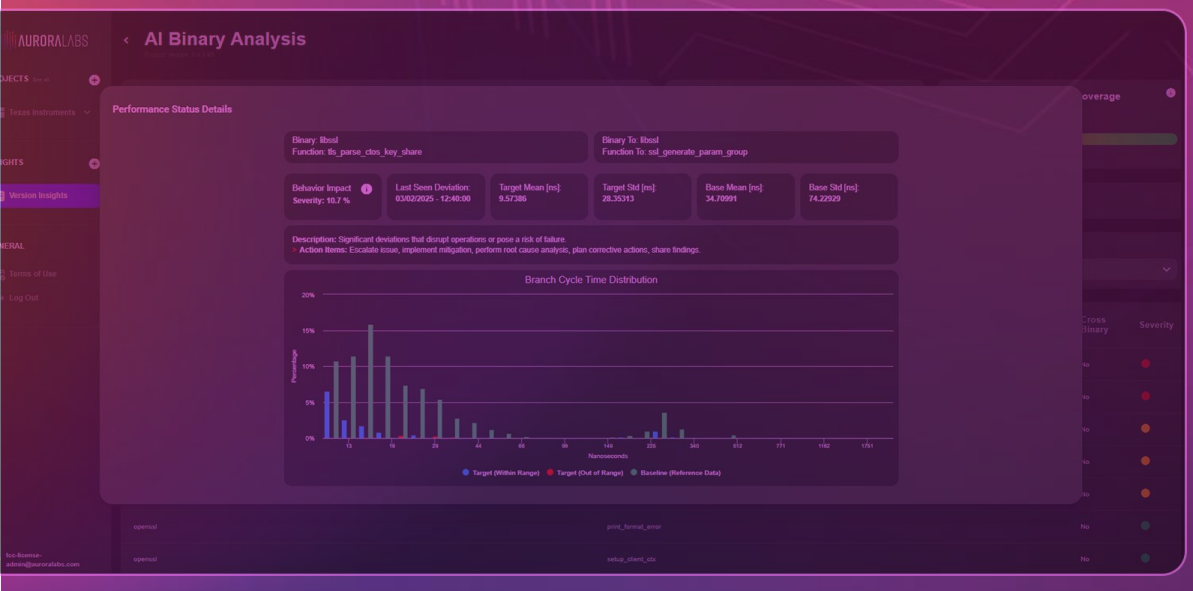
AURORALABS

Executive Solution Brief

LOCI - Autonomous Performance & Reliability Optimization



Published Q4 2025

Executive Summary

LOCI is an autonomous system that predicts, diagnoses, and optimizes GPU/CPU workloads before they run. It eliminates the need for manual profiling, accelerates development cycles, and prevents costly performance regressions — improving engineering velocity, infrastructure efficiency, and product reliability.

LOCI analyzes compiled binaries; not source code, not runtime traces - to surface latency, CPU load, energy inefficiencies, memory risks, and version-to-version regressions, directly after the build.

No flashing. No instrumentation.
No testbeds. No hardware.

For engineering leaders under pressure to deliver predictable performance and shorten cycles, LOCI provides early visibility into what traditionally only appears late in testing or production.

Business Value

- 30–60% reduction in engineering hours spent debugging performance issues
- 20–50% fewer test and QA cycles due to predictive workload analysis
- 2–10× fewer regressions impacting customers or production
- 5–15% lower compute and power costs
- Faster delivery of models, features, and releases

ROI Snapshot (Mid-Size Engineering Team)

Assumptions: 15–25 engineers working on performance-sensitive workloads (AI, data, kernels, backend systems).

- Typical debugging time: 4–8 hours/week per engineer
- Total monthly cost wasted on performance issues:
- 300–700 hours × ~$140/hour ≈ $42k–$98k per month
- LOCI impact (30–60% reduction):
- $12.6k–$58.8k saved monthly

Additional Recurring Value

- Avoided customer incidents: $50k–$250k per major regression
- Infrastructure/power optimization: 5–15% compute savings
- Release acceleration: 5–20% throughput increase

Projected annual benefit: $300k–$1.2M

Problems LOCI Solves

- High cost of diagnosing GPU/CPU bottlenecks
- Long feedback cycles between development and validation
- Regressions escaping to customers or production
- Overprovisioned compute due to unknown inefficiencies
- Fragmented performance workflows

How LOCI Works (with differentiators)

1. Input & Goal
- Natural-language intent
- Works on compiled workloads - Shift Left
- Zero instrumentation

2. Predict
- No-run performance model
- The First , Microarchitecture-aware
- Instant bottleneck detection

3. Diagnose
- Early regression detection
- Explains what's slow and why
- Basic-block insight
-

4. Optimize
- HW-aware, The First no illusions model for coding
- Code & kernel rewriting
- Runtime/config tuning

5. Validate Safety
- Shadow-fork builds
- Objective measurement
- Zero-risk integration

6. Continuous Improvement
- Closed-loop learning
- Stronger predictions
- Compounding ROI

Key Differentiators

- Zero-run predictive modeling
- Microarchitecture-aware reasoning
- Autonomous code rewriting and optimization
- Safe, CI/CD-ready shadow-fork deployment
- Full closed-loop system that improves every iteration

Recommended Next Steps

Have your Platform/Infra/ML team run a 1–2 week pilot on a known bottleneck or regression. LOCI will return a quantified optimization report showing time saved, performance impact, and projected annual ROI.

LOCI supports binaries built for ARM32/64, AURIX, and (soon) x86_64 and RISC-V.

📩 Book a session with our team
📩 info@auroralabs.com
🌐 www.auroralabs.com