

# scanprobe

## A read-only first-pass GPU evidence scan

Use it when a GPU node has become a suspect and you need obvious local NVIDIA evidence before rerunning, draining, or filing a support ticket.

```
python3 scanprobe.py
python3 scanprobe.py --json
```

### What It Checks

- **nvidia-smi** GPU visibility and basic GPU state
- ECC counters exposed by **nvidia-smi**
- temperature and hardware throttle flags
- current-boot NVIDIA Xid/kernel-log evidence when visible

It prints a primary issue, visible evidence, next action, and one conservative triage label: CLEAR, WATCH, DRAIN, or UNKNOWN.

### What It Does Not Check

No silent data corruption, NCCL/fabric health, application correctness, cloud-provider state, storage, scheduler behavior, or performance regressions without visible local evidence.

No DCGM, NCCL tests, CUDA samples, PyTorch, matmul tests, bandwidth tests, thermal stress tests, daemons, network calls, telemetry, resets, clock changes, persistence changes, drains, or workloads.

### DCGM Boundary

This does not replace NVIDIA DCGM. DCGM is the serious datacenter GPU management and diagnostics stack. **scanprobe** is the smaller first-pass scan before rerun, drain, escalation, or deeper diagnostics.

### Why It Exists

In GPU incidents, the first pass often scatters across shell commands, screenshots, Slack threads, and support-ticket fragments. **scanprobe** puts that first pass into one read-only command and one pasteable report.

Question: is this node/GPU obviously weird from local NVIDIA evidence, or should I keep looking elsewhere?

### Alpha Ask

- real outputs from normal and weird nodes
- wrong verdicts or confusing wording
- whether it saved time versus the manual first pass
- missing visible local evidence that would have changed the next action
- where the command ran: host, container, Slurm job, notebook, or Kubernetes pod

Success means it saves a few minutes, catches obvious local evidence, and makes the first response easier to paste into Slack, Jira, or a provider ticket.

Use it for the first look after something feels off. Do not use it as a routine heartbeat or a health certificate.

Repo: <https://github.com/cv700/scanprobe>

Alpha hypothesis: roughly 2-10 minutes saved per first-pass incident, depending on access and familiarity with the GPU stack.