Yunfan Zhang

Gothenburg, Sweden · +46 73 346 28 18 · yunfanteo@outlook.com
linkedin.com/in/yunfanz123 · github.com/44sunsetsf · yunfanteo.world

Profile

I am an MSc student at Chalmers looking for a spring 2027 thesis in distributed systems or systems software. I wrote my own Raft-based key-value store in C++, and at Huawei I helped build an agent that 130+ engineers now use. I am patient, curious and easy to work with, and I like knowing how things work under the hood.

Experience

Software Development Engineer Intern — Huawei Technologies **Jun 2026 - Sep 2026**
Outstanding Intern award • *Python, FastMCP, Pydantic, SQLite, YAML, C++*

- Implemented notification-type switching for OpenHarmony background tasks in C++ on an RK3568 development board, so live progress and normal notifications convert when a task's type changes; unit-tested all 9 transitions.
- Built core modules of PerfAgent (team project), an **MCP** server with a declarative skill system for LLM agents to analyse HarmonyOS traces beyond their context window; it grew from 20 to **130+** internal users.
- Designed its skill engine to reduce hallucination: the framework fills in tool parameters and checks each conclusion against the raw data; file-based context lets subagents run in parallel, and all 17 skills can resume after a failure.
- Built its query engine, which merges 5 parameter levels, keeps spikes in downsampling and orders 80+ metrics by dependency, so composite metrics are computed from their inputs.
- Built its testing and observability: decorators catch 4 agent failure modes, 27 golden samples catch silent regressions, and structured logs trace each session.
- Cut a dashboard API from 11.7 s to **0.93 s** by adding composite indexes and rewriting slow SQL.

Education

Chalmers University of Technology — MSc in Computer Systems and Cybersecurity **Sep 2025 - Jun 2027**
IPOET Scholarship (75% tuition fee waiver)
Courses: Distributed Systems (5/5), Compiler Construction (5/5), Applied Machine Learning, Computer Security

Northwestern Polytechnical University — BEng in Information Security **Sep 2020 - Jun 2024**
GPA 86.4/100, ranked 10/94 (top 11%)

Projects

VParser | *Go, Pydantic, gRPC, MySQL, Redis, Kafka, Qdrant, Vue* **May 2025 - Jul 2025**
Long-video analysis and Q&A agent, deployed and live *Demo: yunfanteo.world/vparser | GitHub*

- Built a Planner–Executor–Critic agent that answers questions about long videos with **RAG** and checks each timestamped claim against the transcript and on-screen text before showing it.
- Offloaded long tasks to a **Kafka** pipeline with tiered retries and a dead-letter queue, so task submission returns in under 50 ms instead of blocking for 60 s+.
- Built a multimodal video context by running ASR and key-frame OCR in parallel and merging speech with on-screen text along the timeline.
- Built hierarchical summaries, keyword matching and embedding-based semantic retrieval for long videos, so only the most relevant passages reach the model and irrelevant context stays out.
- Enforced content-level deduplication and idempotent consumption with distributed locks and MD5 fingerprints, so the same video is never analysed twice.
- Implemented resumable uploads from the last chunk in Redis, and capped model spend with token-bucket rate limiting.

Distributed key-value store | *C++, Protobuf, Linux, CMake, GitHub Actions* **Apr 2025 - Jun 2025**
Linearizable KV store on Raft that survives minority node failures, deployed and live *Demo: yunfanteo.world/raft | GitHub*

- Built a custom RPC framework on Protobuf so that nodes can make remote calls and exchange data.
- Built a key-value storage engine on a skip list for fast reads and writes.
- Implemented **Raft** from scratch, with leader election, heartbeats, log replication, persistence and snapshots.
- Built a MapReduce framework on the same RPC layer and ran a distributed word count on it.

Skills and Languages

Programming	C++ with the STL, move semantics, coroutines and other modern features; Python, Go
Systems & networking	Linux multithreading and memory management; TCP/IP, HTTP, sockets, I/O multiplexing
Backend & middleware	gRPC, Protobuf, FastAPI, Gin; MySQL, Redis, Kafka; async, idempotency, rate limiting
Distributed systems	Raft consensus, RPC frameworks, storage engines
Other	MCP and multi-agent LLM development; Git, CI/CD, Docker, CMake; Claude Code
Languages	Chinese – native; English – fluent