

YIMENG WANG

mobile: (+86) 13169927376 · email: emon100@qq.com

github.com/emon100 · emon100.com · linkedin.com/in/yimeng-wang-4a31b6194

EDUCATION

Northeastern University, School of Computer Science and Engineering

Sep. 2018 - Jun. 2022

Bachelor of Computer Science; GPA: 3.94 / 5.0; Rank: 29 / 272

Shenyang, Liaoning, China

PROFESSIONAL EXPERIENCE

ByteDance: Object Storage - Storage Engineer

Jul. 2022 - Present

Overview: TOS is ByteDance's large-scale distributed object storage system, supporting dozens of exabytes of data and EB-level daily throughput for global products such as TikTok and Douyin.

- **Led Key Stability and Performance Improvements for TOS Gateway Service:** Designed and implemented 5 critical enhancements for the TOS gateway service—including connection management, rate limiting, circuit breaking, and memory optimization—significantly improving scalability, performance, and reliability. Achieved a service level agreement of **99.995% uptime**, with zero P0/P1 incidents and mean time to recovery (MTTR) under 30 minutes.

- Developed middleware to proactively manage slow/idle connections and traffic spikes, increasing per-node connection capacity from **10k to 100k**. Integrated this solution across gateway components, preventing over **10 incidents** post-launch.
- Enhanced service degradation middleware to prevent cascading failures (e.g., from cache penetration). Implemented error throttling and connection rejection during overload, and provided backoff/retry guidance to clients. Resulted in an **80% increase in successful QPS** during overload scenarios.
- Integrated RPC circuit breaker middleware to mitigate single-point failures, enabling automatic failover and retries. **Prevented 8+ alarms per month**, significantly increasing reliability.
- Built object pooling and memory monitoring middleware, reducing garbage collection CPU usage by **50%** and average end-to-end latency by **10%**. Proactively prevented OOM issues under high concurrency, further improving system stability.
- Designed advanced stability solutions, including fine-grained resource management and traffic isolation via request tagging. Defined gateway rules and integrated with the operations platform to enable semi-automated emergency handling, further enhancing scalability and reliability.

- **Enhanced TOS Service Quality:** Reduced user-reported issues by improving the functional testing pipeline and resolving security vulnerabilities and user problems, ensuring high service quality and security.

- Built a multi-cloud functional testing pipeline, systematically verifying complex scenarios (e.g., cross-cloud conditions). **Discovered over 10 multi-cloud related bugs**, enhancing TOS quality.
- Investigated and fixed external vulnerability related to parsing escaped object names in URLs for GCS buckets, improving security.
- Improved Node.js SDK crc64 performance using WASM, accelerating speed from **1MB/s to 200MB/s** and resolving user complaints.
- Diagnosed and fixed a long-lasting Python SDK error response loss by ensuring graceful TCP disconnection on the server, addressing issues with error reporting.

- **Improved TOS User Platform:** Refactored the TOS control plane service, addressing legacy issues and developing new features to enhance efficiency and functionality.

- Standardized control plane deployment by eliminating 90% of hardcoded logic with configuration management, reducing deployment effort and code changes from 900 to 100 lines of config, supporting rapid multi-region scaling.
- Simplified development and debugging workflow by integrating a new framework and building a dedicated test cluster, reducing deployment time from 10 minutes to 5 seconds.

- **Provided Spark SQL support for Log Storage:** Developed offline query features with rate limiting for large-volume log data analysis.

- Developed distributed offline analytics features using Spark's extensibility (Datasource API extension, SQL parsing, pushdown, partition logic), enabling Spark SQL queries and export to Hive. **Increased exported entries from tens of thousands to tens of millions.**
- Systematically resolved storage engine overload by implementing read/write splitting, dedicated compute/storage clusters, and distributed rate limiting. **Reduced average weekly overload alarms by 5 after launch.**

Tencent, WeChat Group - Value-added Business Department

Jul. 2021 - Oct. 2021

- **Internship**

- **Designed and implemented a crawler system:** Developed a crawler system to help operation and product teams gather game product data from social media platforms. This system complied with the robots.txt protocol and could extract relevant information from specified platforms.

Alibaba Cloud, Cross-platform Technology and Basic Services

Jun. 2021 - Jul. 2021

- **Internship**

- **Designed and implemented a static analysis tool for Spring applications:** Developed a static analysis tool using the ASM library to parse bytecode of compiled Spring applications. This tool extracted information about exposed HTTP APIs without requiring application execution or modification. Based on this, I also built an API compatibility detection tool comparing APIs across different application versions. Utilizing a prefix tree algorithm, it accurately identified API additions, deletions, and parameter changes.

OPEN SOURCE PROJECTS

HADSS (github.com/emon100/HADSS)

Overview: A high-availability distributed storage system, suitable for use as an object storage service backend.

- **Design Philosophy:** Achieves high availability with no single point of failure and remains responsive during cluster scaling.
- **Architecture:** Composed of Gateway, Index, and Storage Layers. Nodes in each layer can be deployed across different regions, data centers, or racks.
- **Key Features:** Supports Quorum write/read; Storage Layer supports automatic storage usage balancing based on NodeStatus; Includes E2E, integration, and unit tests.
- **Technologies:** Rust, Go, Shell, JavaScript.

SKILLS

Programming Languages and Related Skills

- **Go, Python, C/C++:** Proficient and familiar with toolchains and language features.
- **SQL, Scala, Java:** Experienced.
- **Certifications:** TOEFL 98. Certified as a Tencent Cloud Architecture Engineer.