

Di Xu

Kirkland – WA

• +1 (425)-615-3057 • xudifsd@gmail.com • github.com/xudifsd

Professional Summary

Principal Software Engineer with 10 years of experience, specializing in **distributed systems** and **backend infrastructure** (Kafka, Raft, C++). Proven track record of leading multi-year technical strategies at Microsoft and Baidu, delivering **99.95%+ availability** and reducing system latency by **5%+**.

Working Experience

Microsoft Corporation

- Principal Software Engineer** Redmond, WA
Sep 2025 - Present
- Recommendation platform*
- Selected by leadership to **intervene in a critical cross-organizational initiative** for the Tally Counting Service. Diagnosed and remediated performance bottlenecks in partner-owned code, re-architecting the read path to achieve a **5% reduction in E2E latency** and unblocking the release.
 - Led the final phase of the Content Builder modernization, retiring legacy protocols and standardizing the ingestion stack for 100% of Microsoft News traffic.
 - Ported Feature Extraction Library to Linux for **LLM Grounding**, unified technical stack for Recommendation and Grounding service.

- Senior Software Engineer** Redmond, WA
Sep 2021 - Sep 2025
- Recommendation platform*
- Content Builder: Architected an event-driven ingestion platform (Kafka) to replace legacy point-to-point pipelines.
 - Overcame organizational resistance from partner level manager to deprecate legacy protocols by designing a backward-compatible shim, unblocking a multi-year technical debt initiative.
 - Orchestrated a multi-quarter migration strategy for 2 major partner teams and migrated 10+ other teams at scale.
 - Feature Extraction Library: Designed & implemented a static C++ library to replace legacy C# feature calculation.
 - Reduced P95 E2E latency by **3ms** and CPU usage by **15%** in the first migration experiment.
 - Led the migration of 50+ feature sets, **designing a config-driven interface that allowed Data Scientists to deploy changes without engineering intervention**.

- Software Engineer II** Bellevue, WA
Dec 2019 - Sep 2021
- Deep Learning Training Service*
- Optimized job management, reducing 95th percentile job creation time from 400s to 46s.
 - Optimized init process, reducing job initialization time from 45s to 2s.

- Software Engineer II** Beijing, China
Jun 2018 - Dec 2019
- Microsoft (Search Technology Center Asia)*
- Designed & implemented job runtime for OpenAI platform. Unified runtime for Hadoop and **Kubernetes** backend.
 - Collaborated cross-functionally with the Bellevue team to build Deep Learning Training Service.

- Baidu Inc.** Beijing, China
Mar 2016 - Jun 2018
- Software Engineer**
- Infrastructure*
- Re-architected** the "Matrix" cluster management system (similar to Kubernetes/Borg) by replacing the MySQL storage backend with a **Raft-based replicated state machine**.
 - Improved system availability from 99.9% to **99.95%** and reduced leader failover time from 30 minutes to **<1 minute**, eliminating the need for manual SRE intervention during master failures.

Education

- M.S. in **Software Design** @ University of Science and Technology of China Sep 2013 - Mar 2016
- B.S. in **Computer Science** @ Dalian Polytechnic University Sep 2009 - Jun 2013

Skills

Languages: C++, Python, Java, C#

Infrastructure: Linux, Raft, Kafka, Kubernetes