

系统组
微软亚洲研究院

电话: +8613810425457
邮箱: imzhenyu@outlook.com

摘要

郭振宇于 2006 年硕士毕业于清华大学，并加入微软亚洲研究院系统组，现任主任研究员，主要研究方向为分布式系统，操作系统，编译技术，以及测试运维工具等。他在 OSDI, NSDI, PLDI, EuroSys, USENIX ATC, OOPSLA, SoCC, HotOS 等系统顶级会议发表了很多论文，并承担相关的社区服务。他同时帮助产品组做架构设计和核心实现，包括交互式/流计算，大数据计算编译优化，分布式协议 (Paxos)，分布式存储，自动测试等。他目前主要工作是 rDSN (<https://github.com/Microsoft/rDSN>)，希望能够大幅改进工业界目前的分布式系统实践，快速构建高可靠高可用的云服务，该项目目前已经在微软 Bing/Office 和小米云等用于搜索和移动应用。

社区服务

OOSPLA 2015, SYSTOR 2013, BPOE-05, BPOE-04 等，程序委员会委员

中国计算机协会全国十佳优秀博士论文 2015 终评委员会委员

工作和教育履历

2006 到现在，微软亚洲研究院系统组，主任研究员

2003 到 2006，清华大学计算机系硕士

1999 到 2003，浙江大学计算机系学士

项目经历

分布式系统框架

- Parallel Paxos (Rex) – 解决传统 Paxos 实现中顺序执行（以达到多机确定性执行效果）和多核多线程不确定性之间的矛盾问题，实现了 Paxos 在多核上的 Scalability。
- Robust Distributed System Nucleus (<https://github.com/Microsoft/rDSN>) – 该项目用于快速开发鲁棒的云服务。项目来源于[对 service outage 的分析](#)，通过微内核架构，全局资源和模块配置，原生开发运维工具的支持，以及内置的高可用高可靠服务框架等，来改善鲁棒性。该项目已开源，并在公司内外都进行了实际的产品部署。
- Tron – 使用类 SQL 语言来构建云服务工作流，并自动优化和部署，希望能够更好地实现服务依赖的管理和优化，提高端对端云服务的鲁棒性和性能，以及开发和运维的效率。该项目同样源自于对 service outage 的分析，目前正在寻找合适的应用场景。

大数据计算，分析，和优化

- Cloud Meter – 一个和云服务一起部署的 streaming 和 interactive 分布式 query 引擎，主要用于监控云服务的健康状况，并在 streaming query 提示有问题的情况下提供交互式查询以查找 root cause。该项目得到了公司内 tech fair 的 golden volcano award。
- G2 – 一个和云计算一起部署的图计算分布式引擎，主要通过对函数级别的分布式执行流图（以及图上附属的相关信息如 log 等）进行 Slicing 和 Hierarchical Aggregation (Zoom In/Out) 操作，来有效帮助分布式系统的 root cause analysis。
- Program Analysis for Scope Script (PASS) – 通过静态分析来帮助寻找 scope job（类似 Spark）的正确性，隐私，和性能问题，帮助减少在 cluster 上运行失败的可能性，并降低相关的 IO cost。相关工作发表在 SoCC'10, NSDI'12, OSDI'12, OOPSLA'14, SoCC'14 等，其中正确性/隐私检查等集成到了相关产品中。

分布式存储

- Kirin, 用于 Bing 的网页和网页索引的存储。
- PDI Store, 基于 leveldb 和 rDSN 第一版的分布式存储，用于 Bing entity graph 的存储，并服务 office 365 等在线应用。
- Replicated rocksdb, 基于 rocksdb 和 rDSN 开源版本的分布式存储，用于小米云服务。

其他还有一些自动测试，分布式监控，自动错误回放等测试运维工具，以及编译相关的项目，主要发表在国际会议以及相关的开源项目中（如 [open64](#)）。

发表论文

- [1] **Zhenyu Guo**, Cheng Chen, Haoxiang Lin, Sean McDirmid, Fan Yang, Xueying Guo, Mao Yang, and Lidong Zhou, Arming Cloud Services with Task Aspects, MSR Tech Report, 2014
- [2] Chang Liu, Jiaying Zhang, Hucheng Zhou, Sean McDirmid, **Zhenyu Guo**, and Thomas Moscibroda, Automating Distributed Partial Aggregation, in SoCC, 2014
- [3] Tian Xiao, **Zhenyu Guo**, Hucheng Zhou, Jiaying Zhang, Xu Zhao, Chencheng Ye, Xi Wang, Wei Lin, Wenguang Chen, Lidong Zhou, Cybertron: Pushing the limit on I/O reduction in data-parallel programs, in OOPSLA, 2014
- [4] Tian Xiao, Jiaying Zhang, Hucheng Zhou, **Zhenyu Guo**, Sean McDirmid, Wei Lin, Wenguang Chen, and Lidong Zhou, Nondeterminism in MapReduce Considered Harmful? An Empirical Study on Non-commutative Aggregators in MapReduce Programs, in ICSE SEIP, 2014
- [5] **Zhenyu Guo**, Chuntao Hong, Mao Yang, Lidong Zhou, Li Zhuang, and Dong Zhou, Rex: Replication at the Speed of Multi-core, in EuroSys, 2014
- [6] **Zhenyu Guo**, Sean McDirmid, Mao Yang, Li Zhuang, Pu Zhang, Yingwei Luo, Tom Bergan, Peter Bodik, Madan Musuvathi, Zheng Zhang, and Lidong Zhou, Failure Recovery: When the Cure Is Worse Than the Disease, in HotOS, 2013
- [7] **Zhenyu Guo**, Xuepeng Fan, Rishan Chen, Jiaying Zhang, Hucheng Zhou, Sean McDirmid, Chang Liu, Wei Lin, Jingren Zhou, and Lidong Zhou, Spotting Code Optimizations in Data-Parallel Pipelines through Periscope, in OSDI, 2012

- [8] Jiaying Zhang, Hucheng Zhou, Rishan Chen, Xuepeng Fan, **Zhenyu Guo**, Haoxiang Lin, Jack Y. Li, Wei Lin, Jingren Zhou, and Lidong Zhou, Optimizing Data Shuffling in Data-Parallel Computation by Understanding User-Defined Functions, in NSDI, 2012
- [9] **Zhenyu Guo**, Dong Zhou, Haoxiang Lin, Mao Yang, Fan Long, Chaoqiang Deng, Changshu Liu, and Lidong Zhou, G2: A Graph Processing System for Diagnosing Distributed Systems, in USENIX ATC, 2011
- [10] Cheng Zhang, **Zhenyu Guo**, Ming Wu, Longwen Lu, Yu Fan, Jianjun Zhao, and Zheng Zhang, AutoLog: facing log redundancy and insufficiency, in APSys, 2011
- [11] Ming Wu, Fan Long, Xi Wang, Zhilei Xu, Haoxiang Lin, Xuezheng Liu, **Zhenyu Guo**, Huayang Guo, Lidong Zhou, and Zheng Zhang, Language-Based Replay via Data Flow Cut, in FSE, 2010
- [12] Bingshen He, Mao Yang, **Zhenyu Guo**, Rishan Chen, Bing Su, Wei Lin, and Lidong Zhou, Comet: Batched Stream Processing for Data Intensive Distributed Computing, in SoCC, 2010
- [13] Bingsheng He, Mao Yang, **Zhenyu Guo**, Rishan Chen, Wei Lin, Bing Su, Hongyi Wang, and Lidong Zhou, Wave Computing in the Cloud, in HotOS, 2009
- [14] Ruini Xue, Xuezheng Liu, Ming Wu, **Zhenyu Guo**, Wenguang Chen, Weimin Zheng, Zheng Zhang, and Geoffrey M. Voelker, MPIWiz: Subgroup Reproducible Replay of MPI Applications, in PPOPP, 2009
- [15] Xi Wang, **Zhenyu Guo**, Xuezheng Liu, Zhilei Xu, Haoxiang Lin, Xiaoge Wang, and Zheng Zhang, Hang analysis: Fighting responsiveness bugs, in EuroSys, 2008
- [16] Xuezheng Liu, **Zhenyu Guo**, Xi Wang, Feibo Chen, Xiaochen Lian, Jian Tang, Ming Wu, M. Frans Kaashoek, and Zheng Zhang, D³S: Debugging Deployed Distributed Systems, in NSDI, 2008
- [17] **Zhenyu Guo**, Xi Wang, Jian Tang, Xuezheng Liu, Zhilei Xu, Ming Wu, M. Frans Kaashoek, and Zheng Zhang, R2: An Application-Level Kernel for Record and Replay, in OSDI, 2008
- [18] Xi Wang, Zhilei Xu, Xuezheng Liu, **Zhenyu Guo**, Xiaoge Wang, and Zheng Zhang, Conditional Correlation Analysis for Safe Region-based Memory Management, in PLDI, 2008
- [19] **Zhenyu Guo**, Xi Wang, Xuezheng Liu, Wei Lin, and Zheng Zhang, BOX: Icing the APIs, no. MSR-TR-2008-03, January 2008
- [20] Shiding Lin, Aimin Pan, Zheng Zhang, Rui Guo, and **Zhenyu Guo**, WiDS: an Integrated Toolkit for Distributed System Development, in HotOS, 2005