Jin Li

Bellevue, WA 98004, USA
Tel. +1 (425) 443-6988
+1 (425) 270-8263
Email: jinli.ccs@gmail.com

Career Highlights

Partner Research Manager, Cloud Computing and Storage, Microsoft Research, Redmond, WA.

IEEE Fellow

Microsoft Gold Star Service Award x4
(1999: for contribution in founding Microsoft Research Asia.
2001: for contribution to scalable audio compression.
2006: for contribution to P2P Video on Demand and P2P folder sharing
2010: for contribution to Deduplication in Windows Server.)

Microsoft Technical Community Network (TCN) Storage Technical Achievement Award 2013

Established a highly productive research team with broad and in-depth technical contribution to Microsoft Products, with financial impact in the order of hundreds of millions dollars per annum: WMA9 Lossless (Reversible Transform), Live Messenger (NAT traversal, sharing folder), Live Mesh (NAT traversal), Windows 7 (Teredo), Lync (Bandwidth Estimation & Management, FEC, Media Gateway, DiffServ, QoS monitoring), Windows 8 (RemoteFX for WAN, BranchCache, Miracast), Windows 8 server (Primary Data Deduplication, Erasure Coding in Storage Spaces), Azure (Local Reconstruction Coding), Bing (Global Traffic Management, Bing object store), Xbox Live (Low Delay Message Protocol), DL Workspace (AI Infrastructure).

Extensive contribution to multimedia compression standards: JPEG 2000 (sub-bitplane scanning and rate-distortion optimization, visual weighting and progressive visual coding, JPEG Interactive Protocol), MPEG 4 (arbitrary shape wavelet transform), H.264 SVC (motion compensated temporal filtering).

95 issued US patents.

IEEE ComSoc Distinguished Lecturer, 2011-2012.

Extensive Community Service and Organization Committee Involvement
ICME 2011 Lead Program Chair.
ICME Steering Committee Chair, 2014-2015.
ACM Multimedia 2016 Program CoChair

Ph.D. with honor in Electrical Engineering, Tsinghua University, 1994.

Affiliated Professor, Tsinghua University, from 2000

Demoed to Xiaoping Deng in 1984. The event brought forth the quote "Computer literacy should start with children" (计算机普及要从娃娃抓起), an iconic event in China. The event photo and the computer used are in display at Shanghai Science Museum.

Selected Recent Projects

Deep Learning workspace: A turnkey infrastructure for AI scientists (2016-current)

- An open source toolkit that builds a turnkey cloud computing infrastructure (either for public cloud or for on-perm clusters)
- Support multiple deep learning toolkits (TensorFlow, CNTK, Caffe, MxNet, etc.) and big data analytical tools (Hadoop/Spark)
- Daily production use within Microsoft (Microsoft Cognitive Service, SwiftKey, Bing Relevance)

- Manage AI training, interactive exploration, inference, and analytics on the cluster without installation
- Seamless setup/collaboration among AI scientists
- Modularly built with heavy use of open source components (CoreOS or Ubuntu with PXE deployment, Kubernetes + docker for orchestration, network file share via NFS, glusterFS, HDFS or Azure file share, ASP.Net/flask WebUI, OpenID for authentication)
- Open sourced at https://github.com/Microsoft/DLWorkspace (released: Sept. 2017)

Erasure coded storage (2006-2017) [press]

- [with Azure] developed Local Reconstruction Code (LRC), which reduced storage overhead from 1.5x (Reed-Solomon) to 1.29x. The work went into production around 2012.
- LRC receives a number of awards, include:
  - The best paper at USENIX ATC 2012
  - 2013 Microsoft TCN Storage Technical Achievement Award
- It saves Microsoft hundreds of million dollars per annum for Azure alone.
- A slight variation of the code is deployed in Windows Storage for Windows 8 and Windows Server 2012.
- Owned the implementation of a number of erasure coding implementation in Microsoft, include the code used in Windows Media Server, Skype/Skype for Business, RemoteFX for WAN.

High performance SSD (Flash) based storage (2007-current)

- Developed "FlashStore", which is a SSD optimized, low RAM footprint key-value store that organizes storage on flash in a log-structured manner.
- It was tech transferred to Bing Object Store in Microsoft backend. SkimpyStash has been further developed to implement an ultra-low RAM footprint key-value store. The storage layer design of SkimpyStash has been incorporated into BW-Tree, a joint project with MSR Database group, and has shipped to SQL Server 2014 (Hekaton) and Azure DocumentDB.

Deduplication (2007-2012) [press]

- [with Windows File Server group] architected and implemented the Primary Data Deduplication feature in Widows Server 2012 [paper] and End-to-End Deduplication for Storage Virtualization in Windows Server 2012 R2. Key contributions include a new data chunking algorithm, a low RAM footprint indexing data structure to detect duplicate data (based on ChunkStash), and a data partitioning and reconciliation technique, the latter two for scaling index resource usage with data size. Deduplication led to major saving to customers (20-82%), and is among top 3 features for Windows File Server introduced at Windows Server 2012. The feature has received rave reviews (The Register, IT Pro, Arts Technica, IT World, Tech Republic), and some customers upgrade to

Windows Server 2012 for the primary data deduplication feature.

Prajna: Cloud Computing Platform, http://msrccs.github.io/Prajna/ (2013-2016) [Fortune press]

- (Spark on .Net) A Distributed Functional Programming Platform for Interactive Big Data Analytics and Cloud Service Building
- Open sourced at https://github.com/MSRCCS/Prajna (released: Dec. 2015)

Honors and Awards

Gold Star service award x4 (1999, 2001, 2006, 2010), Microsoft
(1999: for contribution in founding Microsoft Research Asia.
2001: for contribution to scalable audio compression.
2006: for contribution to P2P VoD and P2P folder sharing
2010: for contribution to Deduplication in Windows Server.)
Microsoft Technical Community Network (TCN) Storage Technical Achievement Award 2013.
Microsoft Member Bench Program, 2007.
Best paper award, USENIX ATC 2012.
Best paper award, ICME 2009.
The Young Investigator Award from SPIE/IS&T, 1998
The Best Ph.D. Thesis Award, Tsinghua University, 1994
Various prestigious scholarships of Tsinghua Univ. during year 1987-1994, such as "Tsinghua Ten Stars", the Supreme Guanghua Scholarship(1993), the Supreme Scholarship of Tsinghua (1991), the Best M.S. Thesis Award (1991), etc..
Champion (#1) of National Youth Computer Programming Competition, China, 1987. (precursor of IOI)