



The Extremely Large Databases Conference at Asia

June 22-23, 2012, Beijing, China

<http://idke.ruc.edu.cn/xldb/>

<http://www.xldb-asia.org>

The Extremely Large Database Conference (XLDB) is founded by people with highly demanding data challenges, and researchers and solution providers who are developing systems to address such challenges. There is no strict definition of "XLDB", and in many cases it represents a major increasing trend of data sizes and the associated complexity and challenges on managing and analyzing the amount of data. The scales of data may also vary from domain to domain. For example, a social network company may handle petabytes of data, and a biomedical study with 30TB of spatially derived data from microscopy images may run a complex spatial query for days or weeks with traditional RDBMS. The conference will provide a meeting place for database researchers, for businesses with advanced solutions, and for people from many research disciplines, industries and organizations who need to urgently address real data challenges.

“No papers. Only talks and discussions by big data owners from science and industry!”

Different from traditional technique conferences, XLDB conference is based on invited premium talks by pioneers and leaders in the field. This year we have brought together world class speakers, who are owners of real extremely large data from industries and scientific research, practitioners who are handling the real data, or DBMS researchers who are researching new solutions. Speakers include Dr. Alexander Szalay, a pioneer in astronomic data management, who builds one of the largest scientific databases in the world together with Jim Gray from Microsoft; Dr. Joel Saltz, a pioneer in biomedical informatics, founder of the virtual microscopy technology, who works extensively on high performance data analysis and queries of biomedical data; Dr Kian-Tat Lim from SLAC National Accelerator Laboratory, Stanford University, who works on designing



and building the petabyte-scale data management systems for the Large Synoptic Survey Telescope, one of the coming largest scientific databases in the next decade. Industry speakers include Dr. Milind Bhandarka, the Chief Architect of Greenplum Labs in EMC, who will present the proposed Big Data Benchmarking framework; Dr. Hairong Kuang, who leads the development of Hadoop Distributed File System at Facebook, and a long time contributor and committer to the Apache Hadoop project; Dr. Martin Kersten, one of the founders of MonetDB, and a pioneer of column store and array database. Other example industry speakers include Google, eBay, TaoBao, Rakuten. Academic speakers include Laura Haas, director of IBM massive data, analytics and modeling research at IBM Almaden Research Center, IBM fellow, ACM fellow, member of National Academy of Engineering, who has invented many technologies and database products for DB2 and Infosphere; Dr. Xiaodong Zhang from Ohio State University, Dr. Haixun Wang from Microsoft Research Asia, and others. Besides invited talks, the conference also accepted a set of high quality lightning talks and poster presentations about XLDB related research and systems.

1. Preliminary Program

Date	Time Period	Activities
June 22	Morning	Registration, Opening, Talks and Discussions from Science
	Afternoon	Talks and Discussions from Industry
June 23	Morning	Talks and Discussions from Academia
	Afternoon	Lightning Talks and Discussions, Closing

2. Registration

Registration Type	Earlybird Registration (Before 25 May 2012)	Late or On-site Registration (After 25 May 2012)
Full Registration	US\$250	US\$300



Student Registration	US\$200	US\$240
----------------------	---------	---------

Registration Website: <https://www.hkws.org/registration/xldb-asia2012/reg.html>

3. Venue

Yifu Conference Center, Renmin University of China, Beijing, China

4. Contacts

Contact: Ruxia Ma, Xiang Ci (Office Phone: 86-10- 62512334)

Email: xldbasia12@gmail.com

XLDB Asia 2012 Organizing Committee

May 4, 2012



XLDB Asia 2012 Tentative Speakers and Talks



Title: Integrating Extremely Large Data is Extremely Challenging

Laura Haas, IBM Fellow and Director, Institute for Massive Data, Analytics and Modeling, IBM Almaden Research Center, USA



Title: Exascale Data Analytics for Biomedicine and Healthcare (TBD)

Dr. Joel Saltz, Chair of Department of Biomedical Informatics, Director of the Center for Comprehensive Informatics at Emory University, Chief Medical Information Officer at Emory Healthcare, USA



Title: Title: A Scale-out Model for Big Data Software Development in Distributed Systems

Dr Xiaodong Zhang, Chair of the Computer Science and Engineering Department at the Ohio State University, USA



Title: Arrays in Database Systems, the Next Frontier?

Dr. Martin Kersten, Professor at CWI, founder of MonetDB, Netherland



Title: Extreme Data-Intensive Scientific Computing

Dr. Alexander Szalay, Alumni Centennial Professor of Astronomy, Professor of Computer Science, Johns Hopkins University, USA



Title: Towards Industry Standard Benchmarks for Big Data

Milind Bhandarkar, Chief Architect, Greenplum Labs, Data Computing Division, EMC



Title: Extremely Large Databases and the Large Synoptic Survey Telescope

Kian-Tat Lim, SLAC National Accelerator Laboratory, Stanford University, USA



Title: Scalability Challenges: Hadoop Distributed File System at Facebook

Hairong Kuang, facebook, USA



Title: Scalability Challenges: Hadoop Distributed File System at Facebook
Tomasz Nykiel, facebook, USA



Title: Data Intensive Study in Geoinformatics
Dr. Guoqing Li, Director of the Center of Earth Observation and Digital Earth Science, Chinese Academy of Sciences, China



Title: The Case of Hadoop and BigData in Rakuten, the Largest e-Commerce Company in Japan
Masaya Mori, founding director of Rakuten Institute of Technology, Japan



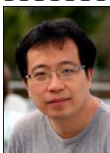
Title: Distributed Online Machine Learning Framework for Big Data
Shohei Hido, co-leader of Jubatus, Preferred Infrastructure, Inc., Japan



Title: Google Storage Architecture and Challenges
Ms. Xuemei Gu, Director, Google Beijing



Title: Extreme Analytics at Ebay
Mr. Eddy Cai, Manager of Data Platform Engineering, eBay



Title: Managing and Mining Billion-Node Graphs
Dr. Haixun Wang, Senior Researcher, Microsoft Research Asia



Title: Data Intensive Astronomy and Astroinformatics
Dr. Chenzhou Cui, Chief Information Officer, National Astronomical Observatories, Chinese Academy of Sciences, China



Title: Oceanbase: a Scalable Relational Database
Zhenkun Yang, Senior Scientist, Taobao, China



XLDB Asia 2012 Tentative Lightning Talks

Massive Data Collection for Your BigData with Fluentd (A Ruby Tool)	Abhishek Parolkar (viki.com, Singapore)
SciQL, Bridging the Gap Between Science and Relational DBMS	Jennie Zhang (CWI, Netherlands)
Jacqueline: JSON/JAQL for MonetDB	Fabian Groffen (CWI, Netherlands)
LUMOS: An Extensive Data Cloud Platform for Big Data Analytics	Jidong Chen (EMC, China)
TaijiDB: A Titanic and Just-in-Time DB	Yunpeng Chai (Remin University, China)
COLA: A Cloud-based On-Line Aggregation System	Yingjie Shi (Remin University, China)
A Highly Scalable Cloud Database for Massive Multi-User Query Processing	Jan-Jan Wu (Academia Sinica, Taiwan)
Streamlining Processing for Big Data (SPBD): a Case Study on Metagenomics Software	Tung Nguyen (Wayne State University, USA)
Lustre, Scalable Storage for Exascale	Liang Zhen (Whamcloud, China)
Discovering Events from Satellite Data	Hideyuki Kawashima (University of Tsukuba, Japan)
.....