



## The Extremely Large Databases Conference at Asia

June 22-23, 2012 , Beijing, China

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

In response to the exploding need for systems and tools capable of managing and analyzing "big data", and the quick embracement of the community, we are organizing the first Extremely Large Database Conference at Asia at Beijing, China on June 22-23, 2012. This conference will bring together people with highly demanding big data challenges, and researchers and solution providers who are developing systems to address such challenges. The conference will facilitate discussions on:

- The state of the art data handling technologies on big data;
- Practical use cases of current and anticipated data challenges;
- Lessons and innovations on building extremely large databases;
- Trends and strategies for surmounting current hurdles.

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.

### 一、会议日程大致安排

6月22日:

上午: 会议报到注册、开幕式、科学界大数据报告

**Panel Discussion: The Challenge and Requirements for Handling Extremely Large Scientific Data**

Moderator: Alexander Szalay

下午: 工业界大数据报告及讨论

**Panel Discussion: NoSQL: the Cure for Big Data?**

Moderator: Milind Bhandarkar



6月23日:

上午: 学术界大数据报告及讨论

**Panel Discussion: Evolution or Revolution: Database Research for Big Data**

Moderator: Laura Haas

下午: Lightning Talk 及讨论、闭幕式

## 二、会议注册费用

注册类型	2012年5月25前注册	2012年5月25后注册
非学生注册	人民币 1500 元	人民币 1800 元
学生注册	人民币 1200 元	人民币 1400 元

## 三、会议注册流程

会议提供美元账号注册和人民币账号注册。

美元账号注册（在线注册）：

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

人民币账号注册（在线登记，银行转账）：

第一步：登陆会议网站：[http://idke.ruc.edu.cn/xldb/www.xldb-asia.org/domestic\\_reg.html](http://idke.ruc.edu.cn/xldb/www.xldb-asia.org/domestic_reg.html) 完成网上注册获取“注册编号”；

第二步：进行银行转账。转账信息如下：

单位名称：科学国际旅行社

汇款帐号：11250101040005236

开户银行：北京农行科院南路支行

汇款附言：**XLDB ASIA + 姓名 + 注册编号**



第三步：扫描银行转账凭证（PDF 或 JPG 格式），并将其发送到  
[xldbasia12@gmail.com](mailto:xldbasia12@gmail.com)

**注：受场地限制，本次会议人数有限，有意参加者请尽快注册；  
只有成功转账后，注册信息方才有效。**

#### 四、会议地点

中国人民大学逸夫会议中心，北京

#### 五、会务组联系方式

联系人：马如霞，慈祥（办公电话：010-62512334）；

联系邮箱：[xldbasia12@gmail.com](mailto:xldbasia12@gmail.com)。

XLDB Asia 2012 组委会

2012 年 5 月 3 日



## XLDB Asia 2012 Program

---

Day 1: Friday, June 22, 2012

08:00 AM	Registration	
8:30-8:40 AM	<b>Welcome by the Chair</b>	Xiaofeng Meng (Renmin University, XLDB Asia 2012 chair)
8:40-8:45 AM	<b>History of XLDB</b>	Kian-Tat Lim (SLAC National Laboratory, Stanford University)
<b>REFERENCE CASES FROM SCIENTIFIC COMMUNITIES</b>		
SESSION CHAIR: KIAN-TAT LIM		
8:45-9:10 AM	<b>Extreme Data-Intensive Scientific Computing</b>	Alexander Szalay (John Hopkins University)
9:10-9:35 AM	<b>Toward Derivation, Management, and Analysis of Exascale Feature Sets</b>	Joel Saltz (Emory University)
9:35-10:00 AM	Coffee Break	
10:00-10:25 AM	<b>Extremely Large Databases and the Large Synoptic Survey Telescope</b>	Kian-Tat Lim (Stanford SLAC)
10:25-10:50 AM	<b>Data Intensive Astronomy and Astroinformatics</b>	Chenzhou Cui (National Astronomical Observatories, Chinese Academy of Sciences)
10:50-11:15 AM	<b>Data Intensive Study in Geoinformatics</b>	Guoqing Li (The Center of Earth Observation and Digital Earth, Chinese Academy of Sciences)



11:15-12:00 AM	<b>Panel Discussion: The Challenge and Requirements for Handling Extremely Large Scientific Data</b>	Moderator: Alexander Szalay
12:10-1:30 PM	Lunch	
<b>REFERENCE CASES FROM INDUSTRY</b> <b>SESSION CHAIR: TOMASZ NYKIEL</b>		
1:30-1:55 PM	<b>The Case of Hadoop and BigData in Rakuten, the Largest e-Commerce Company in Japan</b>	Masaya Mori (Rakuten Institute of Technology, Japan)
1:55-2:20 PM	<b>Distributed Online Machine Learning Framework for Big Data</b>	Shohei Hido (Preferred Infrastructure, Inc., Japan)
2:20-2:45 PM	<b>Towards Industry Standard Benchmarks for Big Data</b>	Milind Bhandarkar (EMC Greenplum Labs, USA)
2:45-3:10 PM	<b>Google Storage Architecture and Challenges</b>	Xuemei Gu (Google Site Beijing)
3:10-3:40 PM	Coffee Break	
3:40-4:05 PM	<b>Scalability Challenges: Hadoop Distributed File System at Facebook</b>	Tomasz Nykiel (Facebook, USA)
4:05-4:30 PM	<b>Oceanbase: a Scalable Relational Database</b>	Zhenkun Yang (Taobao, China)
4:30-5:15 PM	<b>Panel Discussion: NoSQL: the Cure for Big Data?</b>	Moderator: Milind Bhandarkar

**Day 2: Saturday, June 23, 2012**

08:00 AM	Registration	
8:30 AM	<b>Announcements</b>	



## RESEARCH ON BIG DATA MANAGEMENT

**SESSION CHAIR: MARTIN KERSTEN**

8:30-8:55 AM	<b>Integrating Extremely Large Data is Extremely Challenging</b>	Laura Haas (IBM Almaden Research Center)
8:55-9:20 AM	<b>A scale-out Model for Big Data Software Development in Distributed Systems</b>	Xiaodong Zhang (Ohio State University)
9:20-9:45 AM	<b>Managing and Mining Billion-Node Graphs</b>	Haixun Wang (Microsoft Research Asia)
9:45-10:15 AM	Coffee Break	
10:15-10:40 AM	<b>Arrays in Database Systems, the Next Frontier?</b>	Martin Kersten (CWI, Netherlands)
10:40-11:40 AM	<b>Panel Discussion: Evolution or Revolution: Database Research for Big Data</b>	Moderator: Laura Haas
12:00-1:30 PM	Lunch	

## LIGHTNING TALKS

1:30-2:20 PM	<p><b>Lightning Talks I</b> (Session Chair: Weisong Shi)</p> <p><b>Streamlining Processing for Big Data (SPBD): a Case Study on Metagenomics Software</b></p> <p>Weisong Shi (Wayne State University, USA)</p> <p><b>Discovering Events from Satellite Data</b></p> <p>Hideyuki Kawashima (University of Tsukuba, Japan)</p> <p><b>SciQL, Bridging the Gap Between Science and Relational DBMS</b></p> <p>Jennie Zhang (CWI, Netherlands)</p> <p><b>Jacqueline: JSON/JAQL for MonetDB</b></p>	
--------------	---	--



	<p>Fabian Groffen (CWI, Netherlands)</p> <p><b>Massive Data Collection for Your BigData with Fluentd (A Ruby Tool)</b></p> <p>Abhishek Parolkar (viki.com, Singapore)</p>
2:20-3:30PM	<p><b>Lightning Talks (II)</b> (Session Chair: Jan-Jan Wu)</p> <p><b>A Highly Scalable Cloud Database for Massive Multi-User Query Processing</b></p> <p>Jan-Jan Wu (Academia Sinica, Taiwan)</p> <p><b>TaijiDB: A Titanic and Just-in-Time DB</b></p> <p>Yunpeng Chai (Remin University, China)</p> <p><b>COLA: A Cloud-based On-Line Aggregation System</b></p> <p>Yingjie Shi (Remin University, China)</p> <p><b>LUMOS: An Extensive Data Cloud Platform for Big Data Analytics</b></p> <p>Jidong Chen (EMC, China)</p> <p><b>Lustre, Scalable Storage for Exascale</b></p> <p>Liang Zhen (Whamcloud, China)</p>
3:30-4:10 PM	<b>Coffee Break and Poster Discussion</b>
4:10-4:20 PM	<b>Closeout and Future</b>
4:20-5:00 PM	<b>Program Committee Meeting</b>



## 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: 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

---



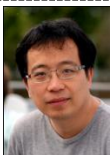
**Title: Google Storage Architecture and Challenges**  
Ms. Xuemei Gu, Director, Google Beijing

---



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

---



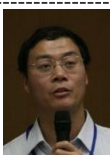
**Title: Managing and Mining Billion-Node Graphs**  
Dr. Haixun Wang, Senior Researcher, Microsoft Research Asia

---



**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