Welcome!

Jacek Becla
XLDB4 Chair
SLAC National Accelerator Laboratory
History

XLDB1, SLAC, CIRCA 2007
Our Secret Recipe for XLDB Events

• Ingredients
  • Big science
  • Data-intensive industries
  • Vendors
  • Academia

• Directions
  1. Mix together
  2. Repeat yearly
Laboratories and Agencies

- Los Alamos National Laboratory
- CERN
- Pacific Northwest National Laboratory
- NASA
- Lawrence Livermore National Laboratory
- ESA
- Argonne National Laboratory
- Fermilab
- Berkeley Lab
- Oak Ridge National Laboratory
- European Molecular Biology Laboratory
- SLAC National Accelerator Laboratory

and more…
Data-Intensive Industries

- eBay
- Nokia
- Google
- Yahoo!
- Chevron
- Facebook
- LinkedIn
- Visa
- ExxonMobil
- MITRE
- Hyves
- AT&T
- AOL

and more…
Vendors

- Teradata
- Oracle
- Vertica
- Microsoft
- Greenplum
- SciDB & Zetics
- IBM
- Netezza
- Aster Data
- MySQL
- Objectivity
- Cloudera
- PostgreSQL
- and more…
Researchers

[Logos of various institutions]

and more…
XLDB Goals

1. Identify trends, commonalities and major roadblocks related to building extremely large databases

2. Bridge the gap between users trying to build extremely large databases and database solution providers worldwide

3. Facilitate development and growth of practical technologies for XL data stores

XLDB4, Jacek Becla
### Workshop vs Conference

<table>
<thead>
<tr>
<th><strong>Workshop</strong></th>
<th><strong>Conference</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>It is all about discussions</td>
<td>It is all about communication</td>
</tr>
<tr>
<td>- Ad-hoc</td>
<td>- Mostly one-way</td>
</tr>
<tr>
<td>- In-depth</td>
<td>- Plenaries only</td>
</tr>
<tr>
<td>- Highly focused</td>
<td>- Time for socializing</td>
</tr>
<tr>
<td>Invitational</td>
<td>Open</td>
</tr>
<tr>
<td>- Carefully selected group of experts</td>
<td>- Publicized inside db and data-intensive communities</td>
</tr>
<tr>
<td>- Open report published</td>
<td>- All talks published</td>
</tr>
<tr>
<td>“Small”</td>
<td>“Big”</td>
</tr>
</tbody>
</table>

Focus: extreme scale & complex analytics & practical issues
### Attendance - Rough Breakdown

<table>
<thead>
<tr>
<th>workshop</th>
<th>conf</th>
</tr>
</thead>
<tbody>
<tr>
<td>xldb1</td>
<td>43%</td>
</tr>
<tr>
<td>xldb2</td>
<td>41%</td>
</tr>
<tr>
<td>xldb3</td>
<td>45%</td>
</tr>
<tr>
<td>xldb4</td>
<td>37%</td>
</tr>
<tr>
<td>xldb4</td>
<td>24% Data-intensive scientific users</td>
</tr>
<tr>
<td></td>
<td>21% Data-intensive industrial users</td>
</tr>
<tr>
<td></td>
<td>19% Vendors, incl. startups</td>
</tr>
<tr>
<td></td>
<td>6% Academic, db research &amp; programmers</td>
</tr>
<tr>
<td></td>
<td>0% DOE, NSF, EC</td>
</tr>
<tr>
<td></td>
<td><strong>Total count</strong></td>
</tr>
</tbody>
</table>

- **XLDB4 attendees:**
  - North & South Americas, Europe, Asia, Australia
That Was Yesterday...

• “New” communities:
  – Oil/gas
  – Finance
  – Medical / bioinformatics

• Science benchmark
• Statistical analytics
• Emerging challenges
• Writable databases
• Next steps

Report will be published later this year
XLDB Community

- XLDB events
- XLDB-L
- http://xldb.org
- XLDB wiki
- YOU

Use #xldb when tweeting
SLAC Tour

- Klystron Gallery, MCC
  - Wed 3:15pm
  - Wed 4:15pm
  - Thu 9:00am
- Vis Lab
  - Wed 3:00pm
  - Wed 3:30pm
  - Wed 4:00pm

60 min

30 min
Stanford Tour

- Bus(es) leave at 5:30pm
- Semi-guided or self-guided
- Get your iPhone ready
Dinner

- At the Stanford Faculty Club
- Reception 6:30pm
- Dinner ~7:40pm
- Buses:
  - To: 5:30pm, 6:00pm
  - From: 9:00pm, 9:30pm
- Or drive on your own
  - Free parking
- Have your dinner ticket with you
Dinner Location
Miscellaneous

• Wear your name tag at all times
• Wireless – use “visitor” network
• Power outlets available throughout auditorium
• Please turn off your cell phones
• No sitting in the isles
• Earthquake drill tomorrow
Organizing Committee

- Magdalena Balazinska, Univ. of Washington
- Jacek Becla, SLAC (chair)
- Jeff Hammerbacher, Cloudera
- Peter Fox, Tetherless World Constellation
- Kian-Tat Lim, SLAC
- Raghu Ramakrishnan, Yahoo!
- Arie Shoshani, LBNL
BIG Thanks to Our Sponsors

1. SciDB & Zetics
   Big Data. Big Analytics.

2. IBM, Greenplum, eBay

3. LSST, aster data
   Large Synoptic Survey Telescope
   more data. big insights.
Agenda

http://xldb.org/4/Program.asp