Advanced Data Management Technologies
Projects for Fun, Profit and Social Good

Alexandros Labrinidis

Advanced Data Management Technologies Lab
Department of Computer Science
University of Pittsburgh
About the ADMT Lab

• Advanced Data Management Technologies Laboratory

• Directed by:
  - Panos K. Chrysanthis
  - Alexandros Labrinidis

• Established in 1995
About the ADMT Lab

• Currently (Fall 2016):
  o 6 PhD students
  o 1 MSc student (CS BSc, 2015)
  o 4 undergraduate students (Junior - Senior)

  o Our “slogan”:
    User-centric data management for network-centric applications
Style of research

• Emphasis on **systems** and **algorithms**
• Building **real systems**
  o Often based on academic prototypes or on top of well-known open-source software
• **Experimenting** using real systems and simulation
  o Often have real users (e.g., astronomers, biologists)
• **Evaluating alternatives**
  o Should we do grouping of queries in way A or way B?
  o What is the relative benefit of each alternative?
  o In which cases would a certain algorithm be better than another?
Enter Moore’s Law

Moore’s law is the observation that, over the history of computing hardware, the number of transistors in a dense integrated circuit doubles approximately every two years. The law is named after Gordon E. Moore, co-founder of Intel Corporation, who described the trend in his 1965 paper.

Source: http://en.wikipedia.org/wiki/Moore's_law
Bezos’ law is the observation that, over the history of cloud, a unit of computing power price is reduced by 50% approximately every 3 years.

Source: http://blog.appzero.com/blog/futureofcloud

Storage capacity increase

HDD Capacity (GB)

Insert other exponentially increasing graphs here
(e.g., data generation rates, world-wide smartphone access rates, Internet of Things, …)

(c) 2016 Alexandros Labrinidis

Wikipedia Data

Oct 7, 2016
But

- Human processing capacity remains roughly the same!
We refer to this as the:

Big Data – Same Humans Problem
<table>
<thead>
<tr>
<th>Systems point of view</th>
<th>Human point of view</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Response time</td>
<td>• Making sure humans do not get lost in a</td>
</tr>
<tr>
<td>• Throughput</td>
<td>sea of data!</td>
</tr>
<tr>
<td>• Scale-up</td>
<td></td>
</tr>
<tr>
<td>• Scale-out</td>
<td></td>
</tr>
</tbody>
</table>
Rethinking Scalability

Example:
Data Stream Management System processing 1,000,000 events per second

<table>
<thead>
<tr>
<th>Systems point of view</th>
<th>Human point of view</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Fantastic!</td>
<td>• Terrible!</td>
</tr>
</tbody>
</table>
Looking at the entire data lifecycle
Current projects for undergrads

• Mobile app for **indoors way-finding**
  (for Carnegie Museum / CS Department)
• **Location-based** virtual bulletin boards
• Virtual queue management
• Zombie-run game (flu-sim)
Upcoming projects for undergrads

• Experiment with Twitter Heron platform (realtime analytics platform) [w/ Lee+Chrysanthis]

• Sports venues analytics [w/ Pelechrinis]

• Real-Time Transit Information Pilot Study [w/ Pelechrinis]

• Bluetooth Beacon Games and Apps
Current Team

Current Undergraduate and Master’s Students:

Anthony Sicilia
David Tsui
Devansh Desai
Kevin Zhang
Mark Silvis

Recent Alumni:  Alec Fox  Chris Meier  John Linahan
                Clint Wadley  David Neiman
Alexandros Labrinidis

Web: http://labrinidis.cs.pitt.edu
Email: labrinid@cs.pitt.edu
Office: 6105 Sennott Square Building

Office Hours:
Mondays: 2:15pm – 3:30pm
Wednesdays: 2:15pm – 3:00pm