Windows Azure for Research

Guobin Wu, Ph. D.

Research Program Manager
Microsoft Research Asia

guobinwu@Microsoft.com

Microsoft Research Connections
微软亚洲研究院·学术合作

Nearly every product Microsoft ships includes technology from Microsoft Research

Microsoft Research (MSR)

- Around 900 people (1%)
- Research lab locations
  - Redmond, WA (09/1991)
  - Cambridge, UK (07/1997)
  - **Beijing, China (11/1998)**
  - Silicon Valley, CA (07/2001)
  - Bangalore, India (01/2005)
  - Boston, MA (09/2008)
  - New York City (05/2012)

MSR Redmond

MSR Asia

MSR India

MSR Cambridge

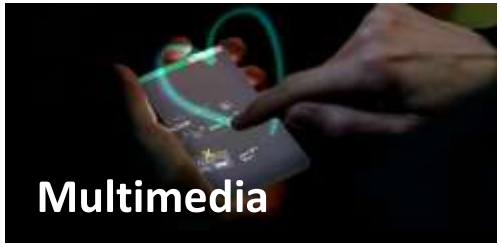
Microsoft Research Asia

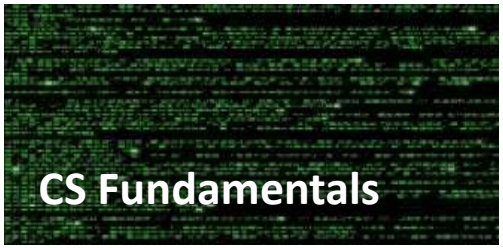
3000+ papers & 40+ best paper awards

300+ technologies transferred to MS products

20+ technologies licensed to other companies

Research Interests

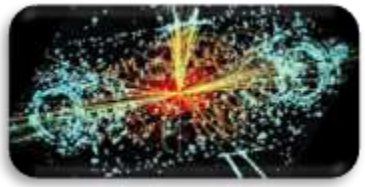




The Future: an Explosion of Data

Experiments

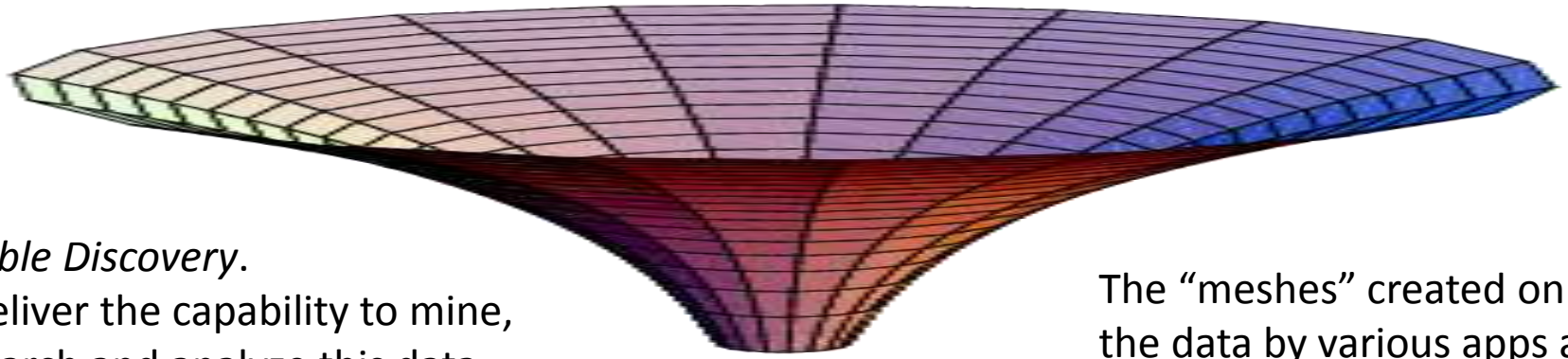
Simulations



Archives

Literature

Instruments



Enable Discovery.
Deliver the capability to mine, search and analyze this data in near real time

Enhance our Lives
Participate in our own heath care. Augment experience with deeper understanding

1PB=1,125,899,906,842,624(1024^5)B

Petabytes
Doubling every
2 years

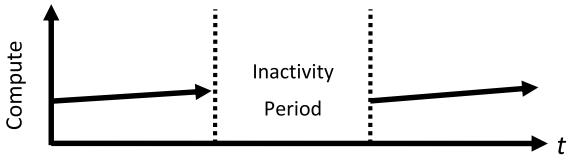
The "meshes" created on top of the data by various apps and further "interconnect" and form "intelligence inside" for future apps and services

Data becomes a strategic asset

Empower Data-driven Businesses

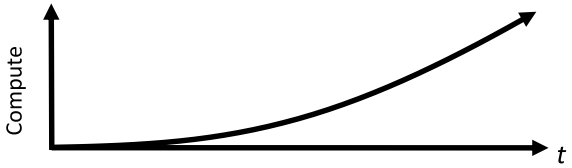
- Search
- Location services: maps, driving directions, traffic, attractions, restaurants, gas stations, weather
- Healthcare: personalized, diagnosis, decision support, monitoring, public health policy/management
- Green causes: air quality, environmental monitoring
- Science and research: physics, astronomy, climate, life sciences, geology
- Social computing
- Commerce: shopping, inventory/supply chain management, advertising
- Security: physical and IT
- Power/water management
- Government services

Cloud Computing Patterns



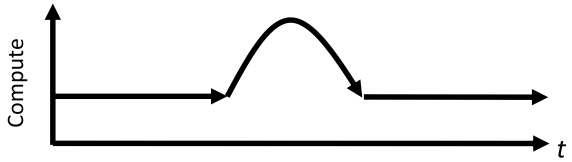
On and Off

On & off workloads (e.g. batch job)
Over provisioned capacity is wasted
Time to market can be cumbersome



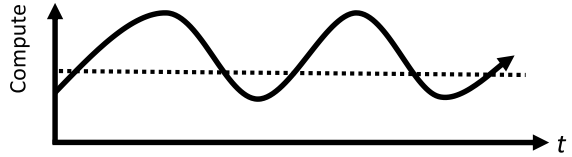
Growing Fast

Successful services needs to grow/scale
Keeping up w/ growth is big IT challenge
Cannot provision hardware fast enough



Unpredictable Bursting

Unexpected/unplanned peak in demand
Sudden spike impacts performance
Can't over provision for extreme cases



Predictable Bursting

Services with micro seasonality trends
Peaks due to periodic increased demand
IT complexity and wasted capacity

Industry is building out massive Infrastructure

MS Data Center

Virtual machines

Cloud services

Web sites

Pay only for what is used

- There is no upfront cost
- There is no need to buy any server licenses
- If you use a SQL database, no need to buy a SQL Server license
- For Virtual Machines and Web Sites, pay by the hour
- Scale up and scale down your solutions – or even turn them on and off as necessary – cost is scaled appropriately, automatically
- Evaluation period is no cost at all, of course – including this training

eScience and the Fourth Paradigm

Thousand years ago – **Experimental Science**

- Description of natural phenomena

Last few hundred years – **Theoretical Science**

- Newton's Laws, Maxwell's Equations…

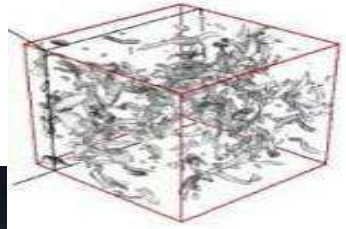
Last few decades – **Computational Science**

- Simulation of complex phenomena

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - \mathrm{K}\frac{c^2}{a^2}$$

Today – **Data-Intensive Science**

- Scientists overwhelmed with data sets from many different sources
  - Data captured by instruments

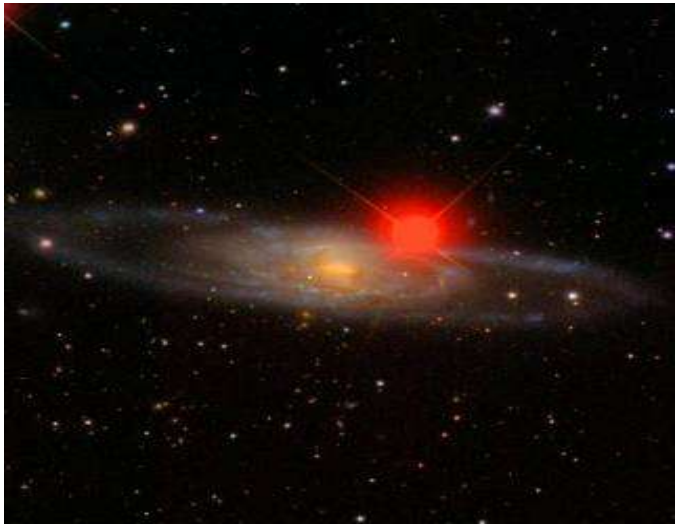


eScience is the set of tools and technologies to support data federation and collaboration

- For analysis and data mining
- For data visualization and exploration
- For scholarly communication and dissemination

(With thanks to Jim Gray)

Astronomical Astronomical Data

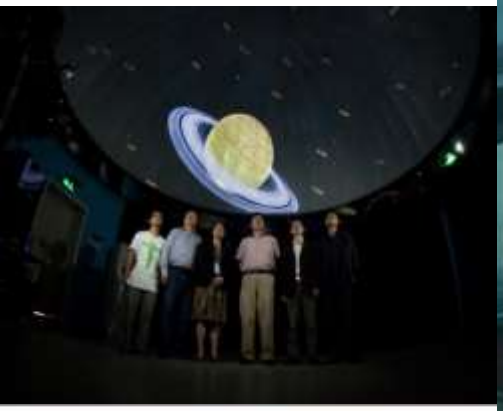


Courtesy of Jeannette Wing

WorldWide Telescope

- Over ten million users around the globe.
- Used by the largest planetariums in the US – SF, NY, Chicago

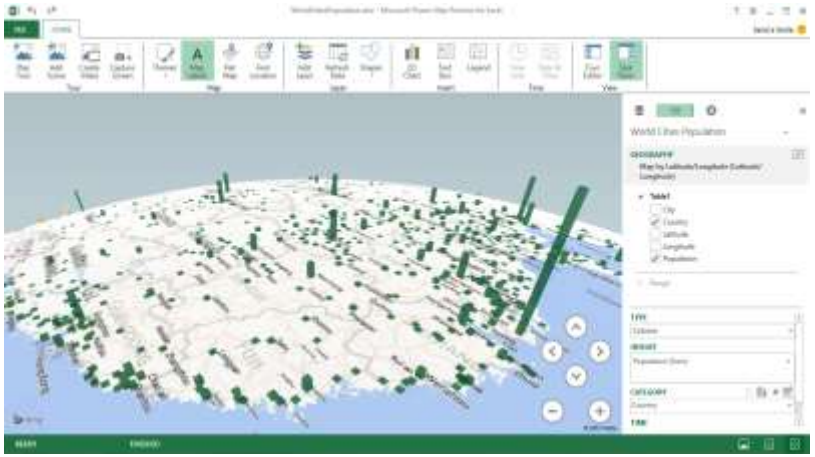
WorldWide Telescope Use in China



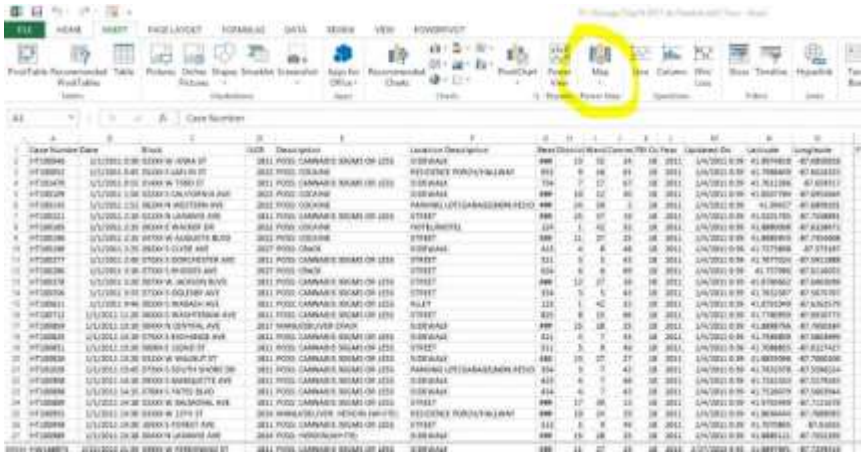
Tour Through the Universe

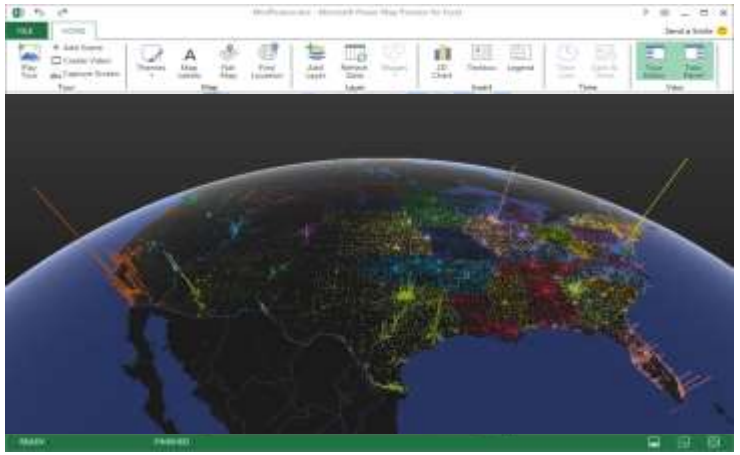
From the Sky to Earth

- Based on the WorldWide Telescope, we built a powerful generic visualization tool for daily use



Power Map in Excel 2013



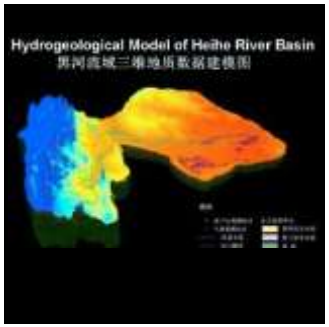


2008 China Olympic Games

How do we connect?

Collaboration on Scientific Data

e-Science on Azure



Global Land Cover Map Visualization and Analysis with WWT (Tsinghua)

Numerical Modeling of Eco-hydrological Processes in the Heihe River Basin Using Windows Azure Cloud (PKU)

2013 RFP on **e-Sciences** (CAS & MS)

MS e-Science Workshop (10/23-25, Beijing)

Windows Azure for Research

Global RFP

RFP every two months

12/15 second submission deadline

Apply on line:
http://www.windowsazurepass.com/research

Award large Windows Azure storage and compute resources for one year

40K USD
200,000 hours
20Tbytes

Azure for Research Training

Windows Azure Training

CAS: 2/28/2013
SJTU: 4/7-8/2013
CUHK: 8/2/2013

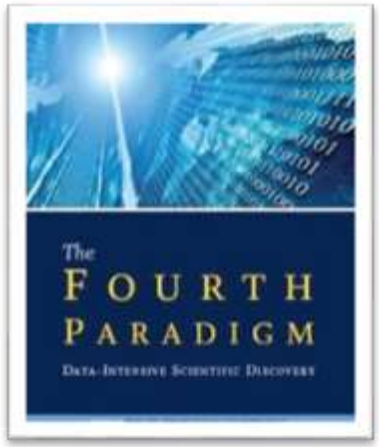
Offer two-day training events around the world (free of charge)

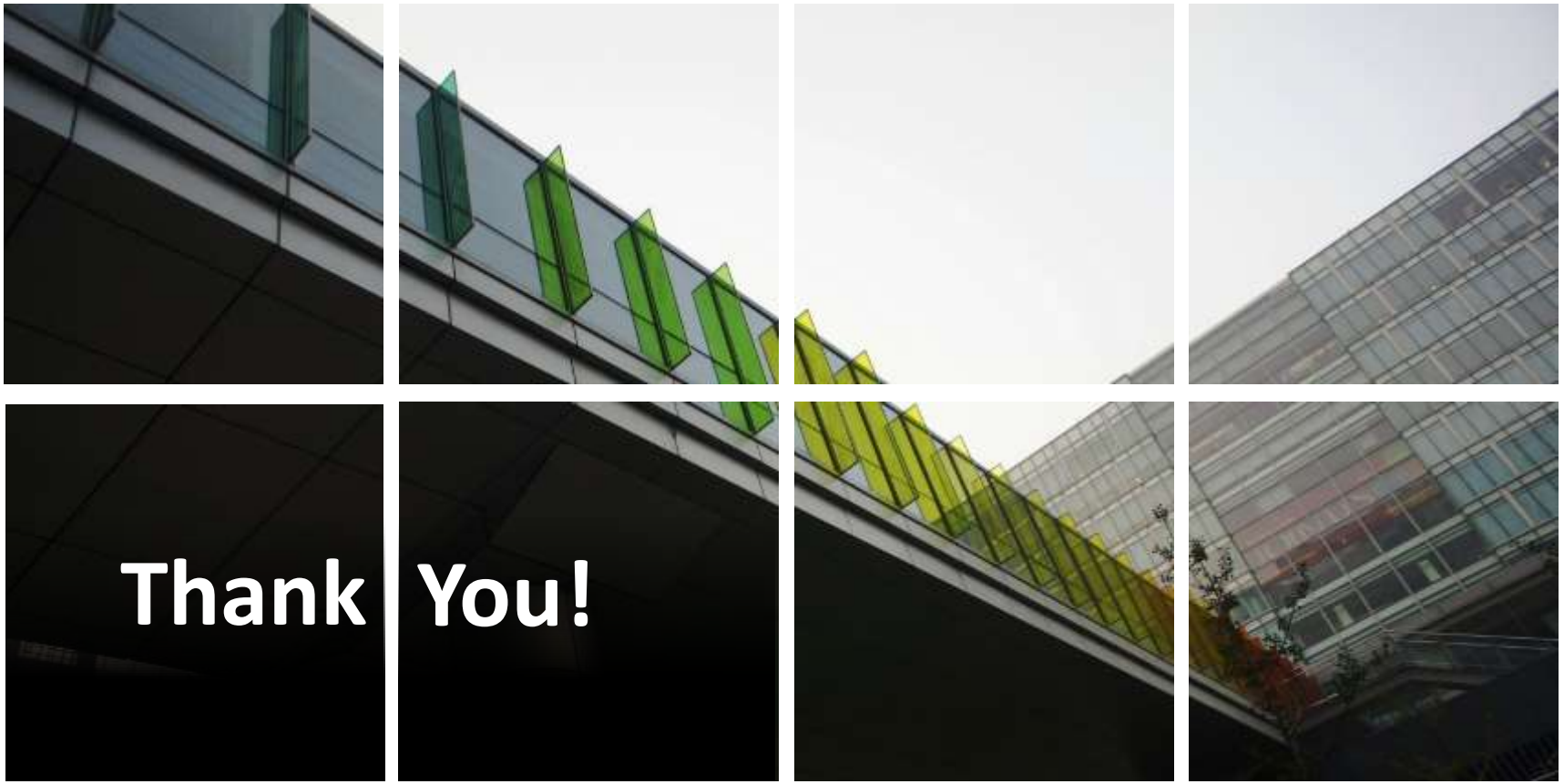
Beijing: 10/25-26, 2013
Guangzhou: 11/7-8, 2013
Beijing: 12/2-3, 2013
Nanjing: 12/14-15, 2013
Hefei: April, 2014

2014
Shanghai
Xi'an
Harbin
Hangzhou
Taipei

Some Resources

- Microsoft Research
  - http://research.microsoft.com
  - Microsoft Research downloads: http://research.microsoft.com/research/downloads
- Microsoft Research Connections
  - http://research.microsoft.com/en-us/collaboration/
- The Fourth Paradigm
  - http://research.microsoft.com/en-us/collaboration/fourthparadigm/
- Science at Microsoft
  - http://www.microsoft.com/science
- Azure for Research
  - http://azure4research.com




Thank You!