

### Windows Azure for Research

#### Guobin Wu, Ph. D.

Research Program Manager Microsoft Research Asia

guobinwu@Microsoft.com

Messearch Connections 激软亚洲研究院·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)





#### 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

#### Simulations **Archives** Literature **Experiments** Instruments Cornell University Library arXiv.org Enable Discovery. The "meshes" created on top of Deliver the capability to mine, the data by various apps and search and analyze this data further "interconnect" in near real time 1PB=1,125,899,906,842,624(1024^5)B and form "intelligence inside" for future apps and **Enhance** our Lives Petabytes services Participate in our own heath **Doubling every** care. Augment experience 2 years Data becomes a strategic asset with deeper understanding

## **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

Microsoft Research Connections

# 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
- 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**









## 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**

These are the million galaxies in the Sloan Digital Sky Survey

A trip from here to our Solar System is 600 megaparsecs or about 1.6 billion light years

## 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**



The second matter mark local both story who examines									Printing State (1. Annual and State									
		Table Follow Bar	10	E a			1 52 55	1	100 H	. 1		-	1000	1	-	10.00		
44		1 = + 4	are four-term															
	· 4			1811	- E.		14.7	10.1	4	4	1	4	- M					
	Taxy Number	Care Block		1938 Descy	NO	Location Description	See.7	244447	61-171	in in T	EQ.7	war i	anned the	Service	unphole .			
E	0-7120346	1110001.0.0 1000 e 4004		2811 /005 1	ANNALL SHART OF LCCC.	UTIMAN	-0480	10	10	14	18.	清助	1/4/2003 0:29	1.011110.0	-61/10/0011	6.		
١.	-1180N	1152304-5-41 Pacin 2 aM1 8	100	ALL WELL I		HELERAL TOPOLE HALEAST			. 16.			3011	04/1811-0-04	(c).70884/3				
	10004	3/1/2012 010 01404 W 1980			ANALY NUMBER OF A	TO DE VEALA	114	e	12.1	. 61		300.	14001039	<1.71.21106	47,69810	κ.		
C	+TUBLEY	Artifold Line Blaters ON/A	DAMOR (HART	301 AND 1		2000 and all	186	. 10	44	141		3844.	24/001816	12001100	<ul> <li>47 (m) (i) (ii)</li> </ul>			
0	-1088144	MUTTER COLUMN 4 MOT	2081-01/6	- 2011 0000 C	00494	PARTICICULATION ACCORDING INCOME	460	14	14		14.	361	14/10111-014	41,29407	41 6418101	611		
c	+1100323	WEPPERS PLAN GERRY IN LANSE	Artik Anile		MAAR'S SOMETHING	471487		- 18	10	181	18.	3844	37473023.0.84	41.5221765	87.750000	61		
c	+1100.005	1/1/1981 J.31 (0000-1 (0403)	07.04	and yone o	DOANE.	AND TRADING THE	414	1.1	-46	38.	18.	1011.	1/4/0011036	11,489,004	415338413	65		
6	011001044	INTERESTING AND ADDRESS OF ADDR	4018.8009	301 Philo 1	10.0.04	1111111	189	- 11	100	23	18.	1011	10,000 100	41,8683915	41,763000	67		
ĸ.	17145346	WEREA AND REDOCTORY	100	3417 0030 0	Math	UDFglAst.	444	1.14	4	48.	- 68	3041	434/1891-0-94	417171668	47.019441	87		
۴.	ATOMETY	WINDOLL & WINDOWS & DOWL	HERITER AND	JELL POTO C	AMAGE STRATE OF LTD.	THEFT.	311			42	18	1100	149/2011 0:38	ALM PROP	-07.041188	67		
c	whiteday.	UL7011 / W.ET00-1-M-818	19 AV	HIT NOG -	NV.	279827	454				10	MO.	1/4/1811-0-04	41.717996	-en act 44001	ĸ		
ŧ.	withman .	WHICH IN MONY AND	adiny they be	3811 7050 4	MANUE REMAINS OF LTREE	utest?	144	-10	10.	. 64	18.	attr.	144/00110-00	41.010462	#7.5465046	67		
۰.	+1181106	1/1/2012 IV-00 17/00 IV-0/04/8	BLANK	1811 POSE 4	MANUE REALTON 1994	179417	394	1.1	- 6	44.0	18	ant.	3/4/2011/039	41,7612987	42.5670391	87		
6	-CORNELL-	AUTORIT PHE MOOP'S MARKA	decession -	HELL POIDC 0	AMPARES REAR OF LESS	8447	118.	1.	-	41	10.	3811	14/1011038	41.0705249	#PANESH	60		
÷	-rootu	L/L/ICC2 LL JE MCCO/L WALK	PERMIT AVE:	HALL PERSON OF	ANSAGE INCALLING 1810	A79467	1426		10.		18	3000	1/4/2012 8:88	1017100900	40 00 MILTIN	80		
×.	+toney	1/1/2011 to at large holds in Ohm	14, 14	2011 10040.0	DRUVER DAMA	ALCOMAL <sup>®</sup>	1.000	15	18.	125	18	3811.	140001031	41,8886154	47.760244	61		
÷	-TINKS	UNLIDED IN BE DREAK & BROWN	\$103 AVR.	0411 0000.2	MANAGE SECURITIES AND	BENEAU	121	- 4	100	44	14	1110	1/4/2012 0-04	41,7540010	47.3828499	67		
÷	withhere a	UNDER THE REPORT OF	194	1411 0039 0	AMMAGED STOLATION (22)	119651	. 111	5		44	18	344.0	144/1003 0-04	433086805	-41.5323423	67		
Π.	+7189634	UTARD IF BUSY was	8.大时	1811 POND 0	SMRARE STOLET	ADDATAS	440	38	- 27.	27.	18	3811	1/4/10110.08	01-003594	47,706201	£.		
ĸ.	+114044	GURLINE PRODUCT	10000	19811 PONIA I	MAAKS NOME OF 1555	Ashing critishashhim Alter	184		18	40	18.	anti-	14/08110-08	4130009	40.594634	£Ľ,		
£	oruses.	STUDIES NAME INVOLTANCE	BUTTE OVE	BELL PRESS C	AMPARES NO. WEAKS OF LESS	T-DF-D-Aut.	410			44	18.	100	1/4/1811.0.19	<b>#1.7382.00</b> 8	67.5176.81	63		
ĸ.	YTIMM	SPECIES SAME AND A VERY SAME	840	1011 POSS -	AMMARIE WORKS ON 1930	millionature .	414	- 4	1.0	41	18	ant.	244/08118-98	43,7126079	47.500744	£		
×.	1710040	UUTILI IN BELLEVIN DAM	ACTING, MUR.	IIII /000 1	ANNALS INCAS OF LESS	299427	100	10	- 66	- 14	18	anti,	4/4/MELIUM	AL BROOMS	87,7121010	67		
ĸ.		1/1/2010 10:00 00:00 w 10**	94	Dis mand.	NUME NEW WITH	RELEASE FOR DATE AND		10	- 14	11	140	1041	AAA/18(1)(0.89	e) behalo	47,7488181	67		
	+1100448	1010202 DRIN INVESTIGATION	11.6/8	10111 10020-1	MANAGER INFUMD-OF 1228	177347	141	1.0			18.	3011	244/080218189	as trimain	-87.5 alat	67		
й.	+1100489	while a most mean at the	and Anti-	2634 (1055) 1	0000000100	(5-30) shifted	1.85	15-	18	10.	18	1146	4/4/00010-00	VI. BARFLIS	41701110	67		
	POPULARITY.	ATTRACTOR OF AN ADDRESS OF ADDRESS	stranger by	SHALL PROVE 1	APPARTO DECARL OF LESS	and the second s	100	- 14	-	10	10.0	-	a (art/mean # ##	in hereiten.	87.729844	27		



## 2008 China Olympic Games

#### China's 2008 Beijing Olympics

In the 2008 Seijing Olympics. China won more medals than it had at any previous Olympic Games. Chinese athletes took home a total of 183 medals, and won more gold medals than any other country by a wide margin.

## How do we connect?



## **Collaboration on Scientific Data**



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





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



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

### **Windows Azure for Research**



Microsoft Research Connections

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



