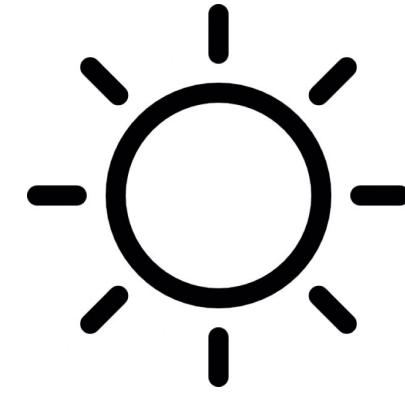


# CS 644: Introduction to Big Data

Daqing Yun  
New Jersey Institute of Technology

# Outline

- Introduction
  - Motivation
  - Where are the big data sets from?
  - Key enablers
- Objectives and Challenges
- Computing Technologies
  - High-performance Computing
  - Grid Computing
  - Cloud Computing
  - Mobile Computing
- Ongoing Research
  - High-performance Networking
  - Data-intensive Computing



# Motivation

- Solving the *big problems*

Then what are  
the big problems?



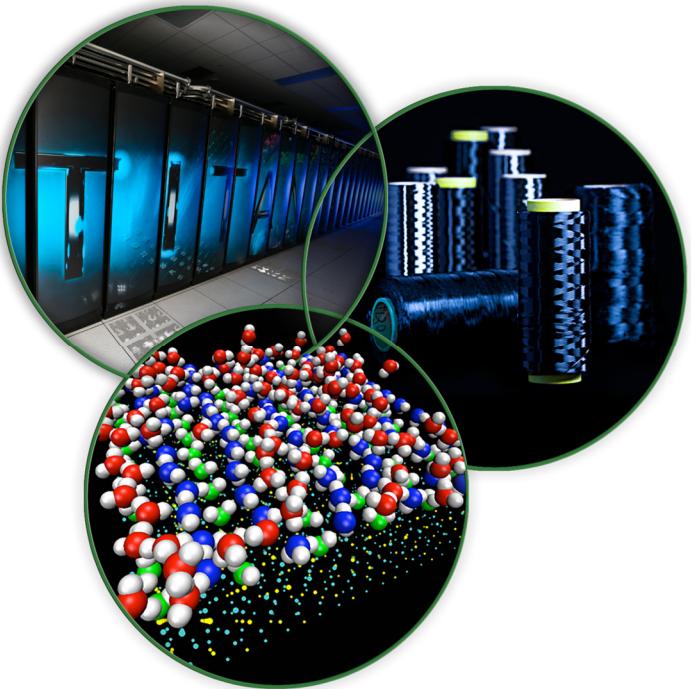
# What is big data?

- “Big” data is similar to small data, but only bigger?
- Having data bigger consequently requires different approaches
  - Techniques, tools, and architectures,
  - With an aim to solve new problems,
  - And old problems in a better way



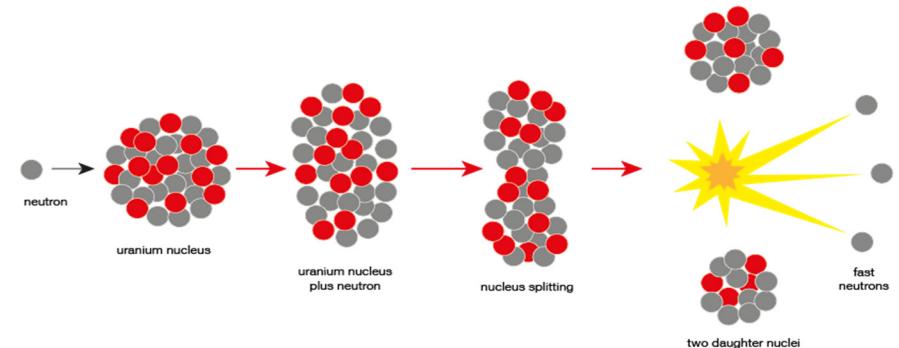
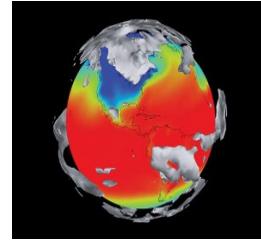
# Where are the big data sets from?

- Large-scale scientific applications



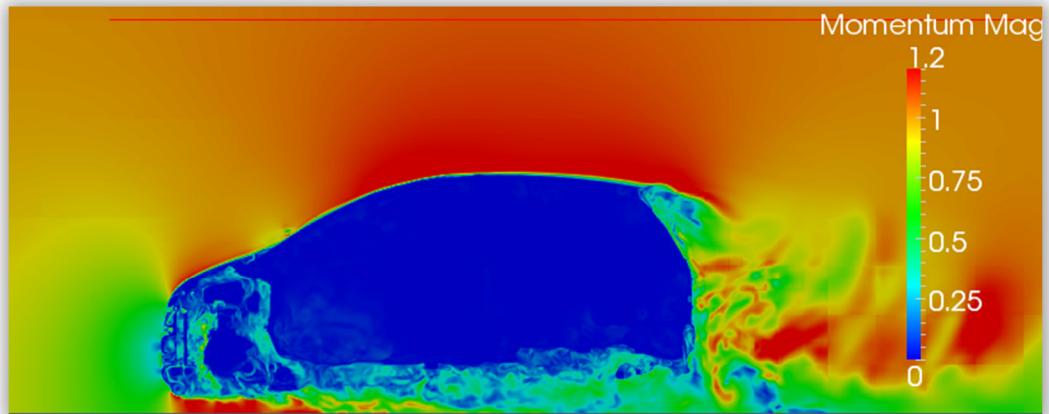
# Where are the big data sets from?

- Large-scale scientific applications
  - Simulation
    - Astrophysics, climate modeling, combustion research, etc.
  - Experimental
    - Spallation neutron source, large hadron collider, microarray, genome sequencing, protein folding, etc.
  - Observational
    - Large-scale sensor networks, astronomical imaging/radio devices (Dark Energy Camera, Five-hundred-meter Aperture Spherical Telescope – FAST, etc.)



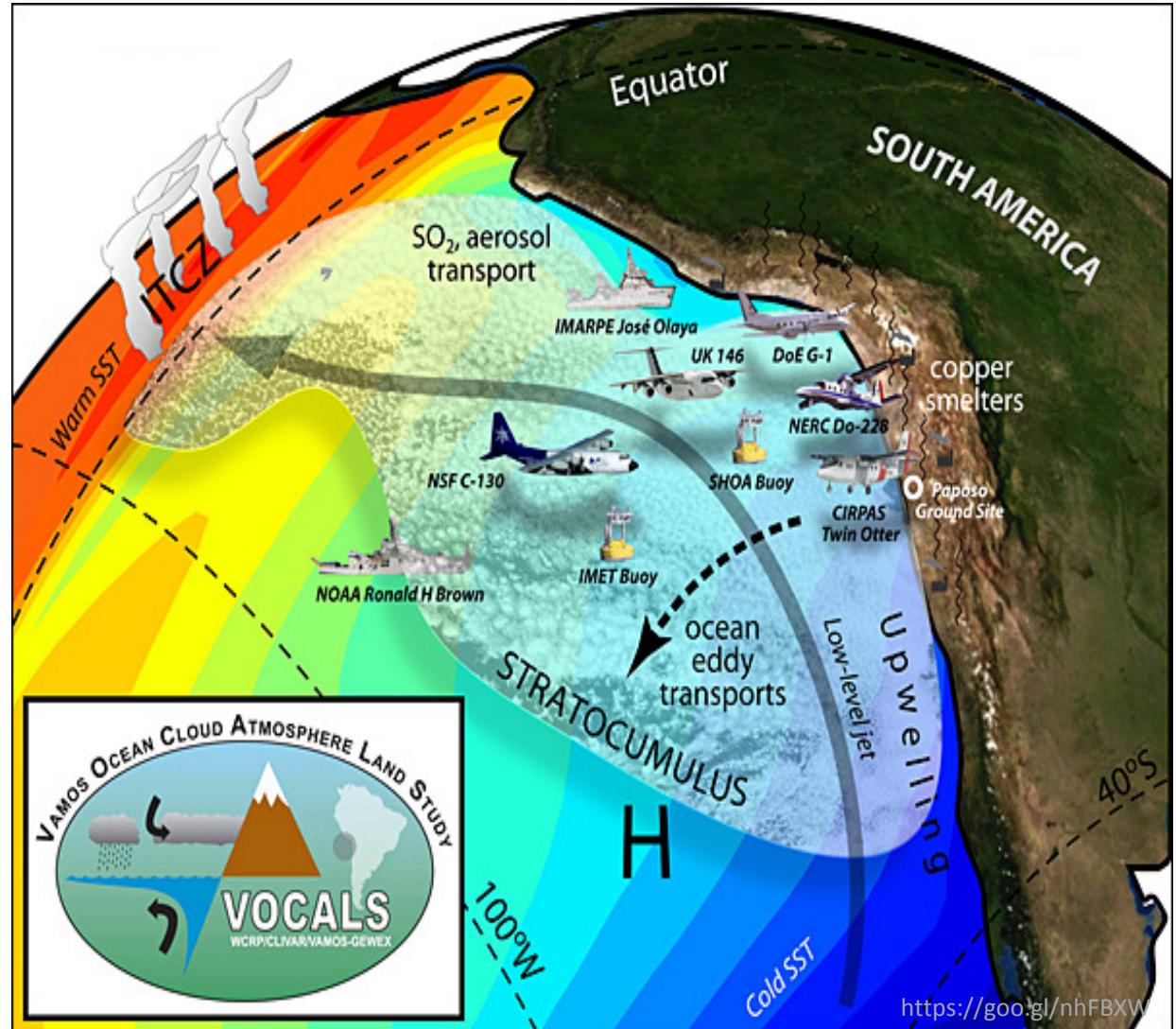
# Science is changing...

- Thousand years ago, science was empirically describing natural phenomena
- Last few hundred years, theoretical branch using models, generalizations
- Last few decades, a computational branch simulating complex phenomena
- Today, data-intensive science, synthesizing theory, experiment and computation with statistics, i.e., new way of thinking required



# Science is becoming data-driven

- Convergence of physical and life sciences through Big Data (statistics and computing) and a scientific revolution in how discovery takes place, which is a rare and unique opportunity



Albert Einstein  
Old Grove Rd.  
Nassau Point  
Peconic, Long Island  
August 2nd, 1939

F.D. Roosevelt,  
President of the United States,  
White House  
Washington, D.C.

Sir:

Some recent work by E. Fermi and L. Szilard, which has been communicated to me in manuscript, leads me to expect that the element uranium may be turned into a new and important source of energy in the immediate future. Certain aspects of the situation which has arisen seem to call for watchfulness and, if necessary, quick action on the part of the Administration. I believe therefore that it is my duty to bring to your attention the following facts and recommendations:

In the course of the last four months it has been made probable - through the work of Joliot in France as well as Fermi and Szilard in America - that [it may become possible to set up a nuclear chain reaction in a large mass of uranium, by which vast amounts of power and large quantities of new radium-like elements would be generated. Now it appears almost certain that this could be achieved in the immediate future.

This new phenomenon would also lead to the construction of bombs, and it is conceivable - though much less certain - that extremely powerful bombs of a new type may thus be constructed. A single bomb of this type, carried by boat and exploded in a port, might very well destroy the whole port together with some of the surrounding territory. However, such bombs might very well prove to be too heavy for transportation by air.

-2-

The United States has only very poor ores of uranium in moderate quantities. There is some good ore in Canada and the former Czechoslovakia, while the most important source of uranium is Belgian Congo.

In view of this situation you may think it desirable to have some permanent contact maintained between the Administration and the group of physicists working on chain reactions in America. One possible way of achieving this might be for you to entrust with this task a person who has your confidence and who could perhaps serve in an unofficial capacity. His task might comprise the following:

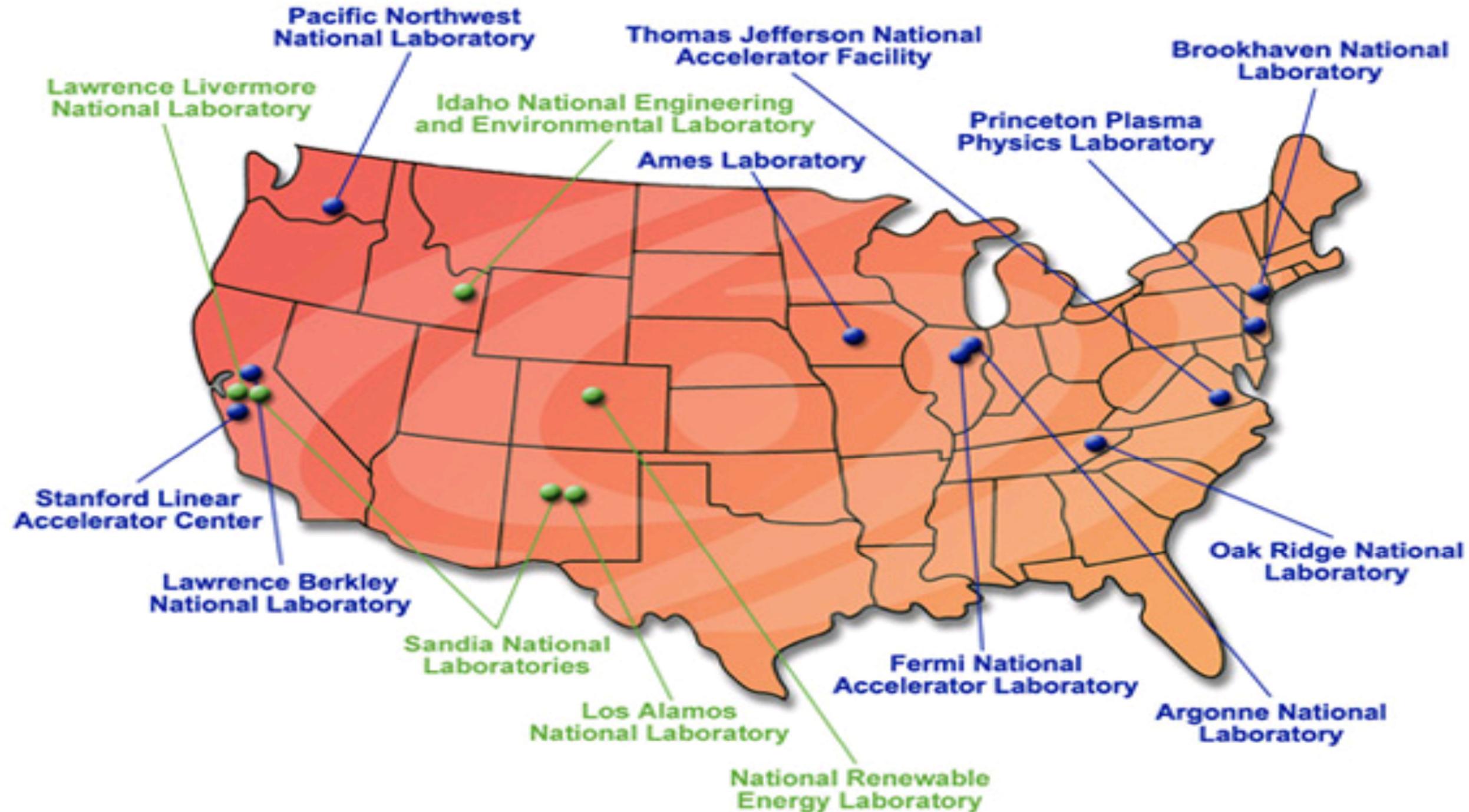
a) to approach Government Departments, keep them informed of the further development, and put forward recommendations for Government action, giving particular attention to the problem of securing a supply of uranium ore for the United States;

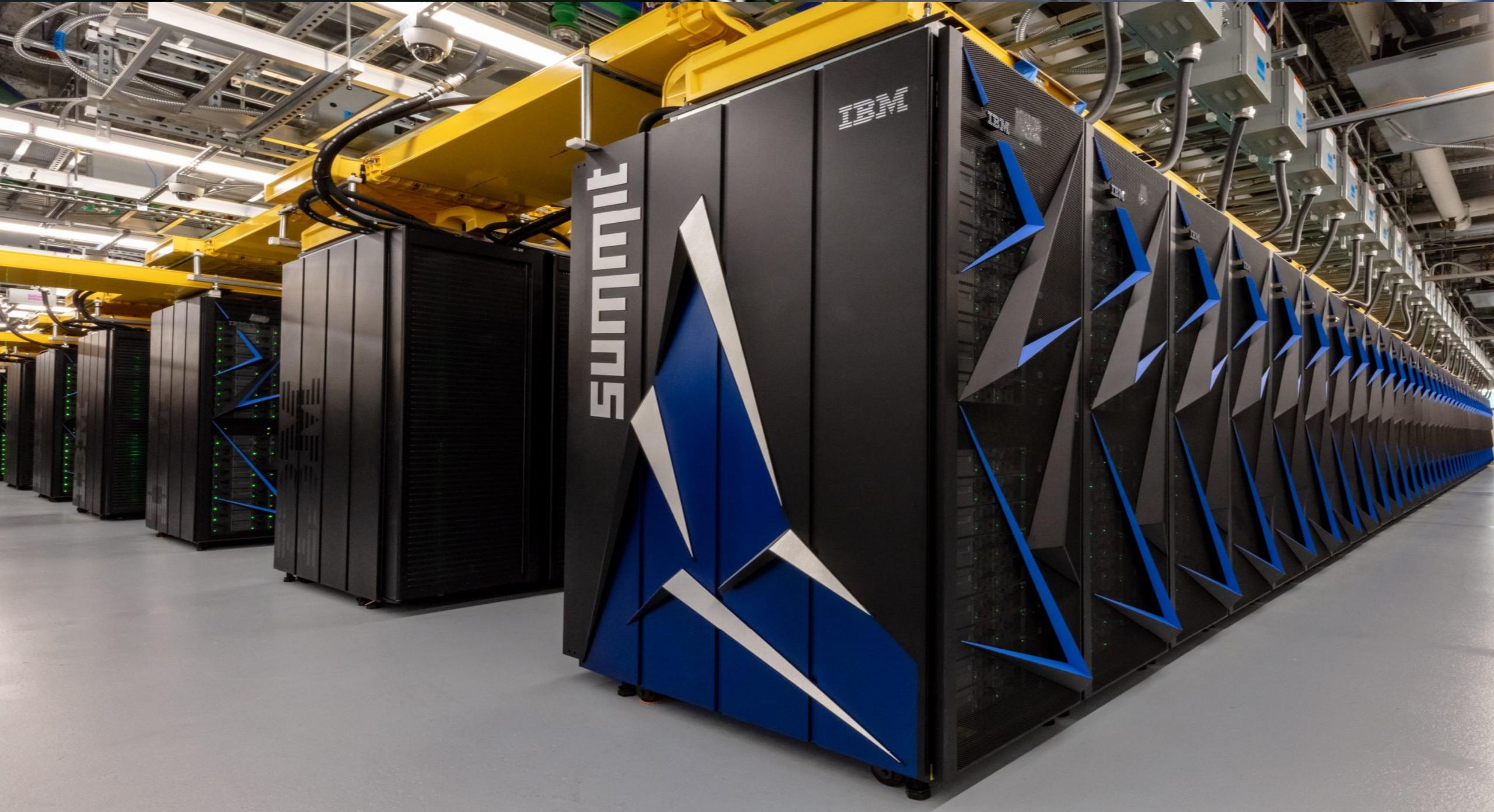
b) to speed up the experimental work, which is at present being carried on within the limits of the budgets of University laboratories, by providing funds, if such funds be required, through his contacts with private persons who are willing to make contributions for this cause, and perhaps also by obtaining the co-operation of industrial laboratories which have the necessary equipment.

I understand that Germany has actually stopped the sale of uranium from the Czechoslovakian mines which she has taken over. That she should have taken such early action might perhaps be understood on the ground that the son of the German Under-Secretary of State, von Weizsäcker, is attached to the Kaiser-Wilhelm-Institut in Berlin where some of the American work on uranium is now being repeated.

Yours very truly,

A. Einstein  
(Albert Einstein)





# solving the **BIG PROBLEMS**

---

O A K   R I D G E   N A T I O N A L   L A B O R A T O R Y  
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY



Supercomputer

# Six Scientific Themes

*Born of necessity. Inspired by our quest to know.  
We have always been called upon to address  
America's greatest scientific challenges.*

*"Men love to wonder, and that is the seed of science."*

*...Ralph Waldo Emerson*

## BIOLOGICAL SYSTEMS

Developing New Options

Whether converting biomass to fuel or understanding the impacts of climate change, biological research at ORNL is helping develop new options for energy, environmental protection, and human health. [mannre@ornl.gov](mailto:mannre@ornl.gov)



## NEUTRON SCIENCE

Leading the World

The Spallation Neutron Source and the High Flux Isotope Reactor together make Oak Ridge the world's foremost center for neutron science.

[andersonian@ornl.gov](mailto:andersonian@ornl.gov)

## ADVANCED MATERIALS

Strengthening American Industry

With DOE's first Nanoscience Center, the world's most powerful electron microscope, and the High Temperature Materials Laboratory, Oak Ridge plays a critical role in American industrial competitiveness.

[buchananmv@ornl.gov](mailto:buchananmv@ornl.gov)



## NATIONAL SECURITY

Guarding the Gates

From biochemical sensors to stopping the proliferation of nuclear weapons, technologies that make America safer are among the laboratory's top research priorities.

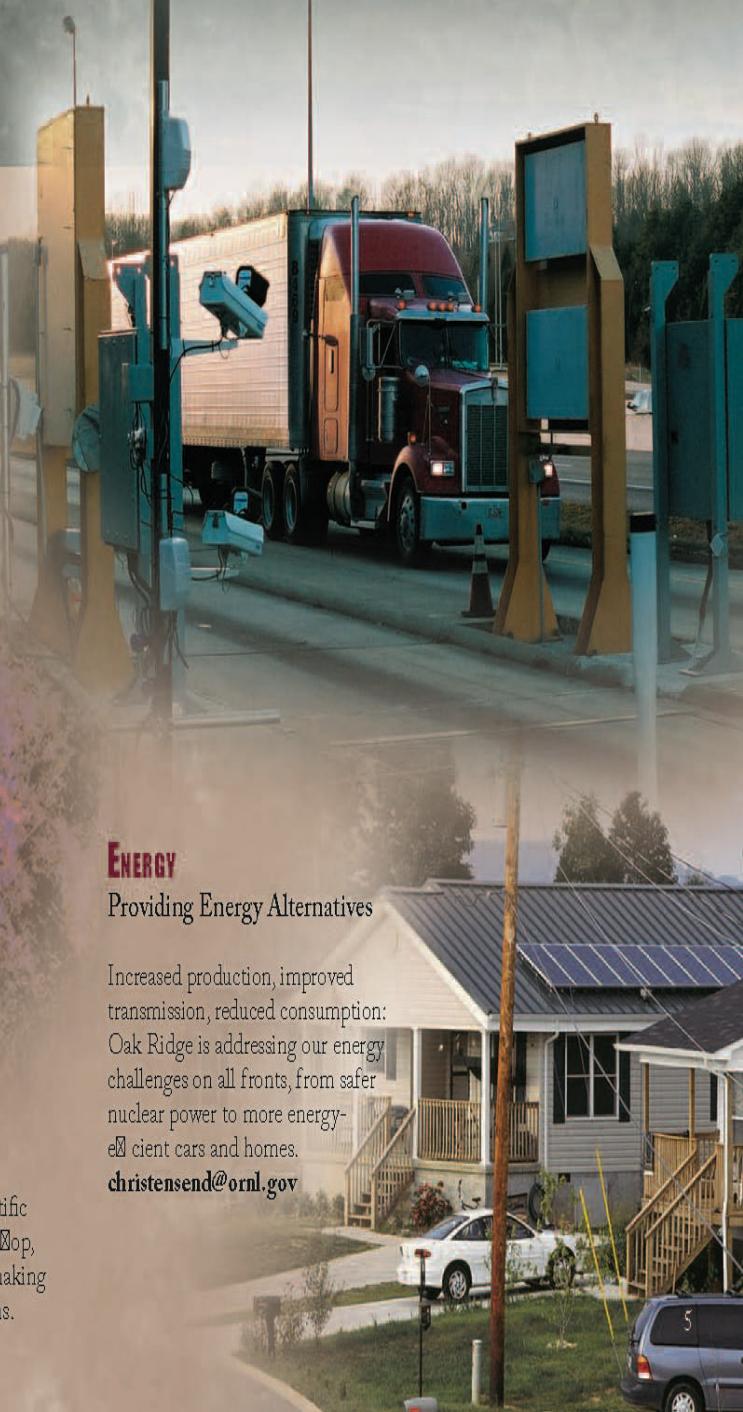
[akershjr@ornl.gov](mailto:akershjr@ornl.gov)

## ENERGY

Providing Energy Alternatives

Increased production, improved transmission, reduced consumption: Oak Ridge is addressing our energy challenges on all fronts, from safer nuclear power to more energy-efficient cars and homes.

[christensen@ornl.gov](mailto:christensen@ornl.gov)



## HIGH PERFORMANCE COMPUTING

Tackling the Big Problems

With unmatched computational capacity for open scientific research, Oak Ridge is on a path by 2009 to reach a petaOp, or 1 quadrillion mathematical calculations per second, making it possible to model the most complex scientific problems.

[zachariat@ornl.gov](mailto:zachariat@ornl.gov)

# Where are the big data sets from?

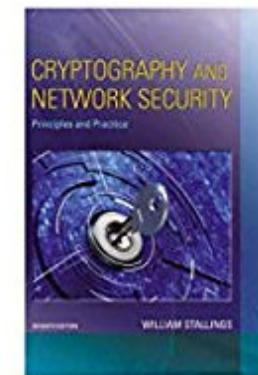
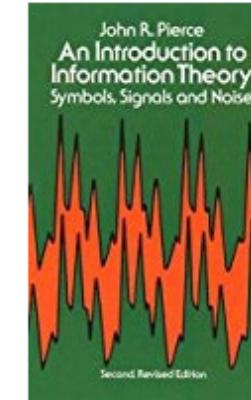
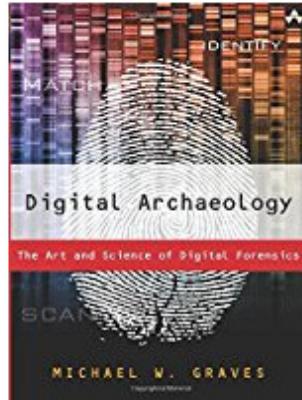
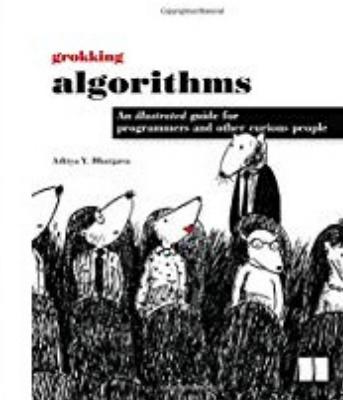
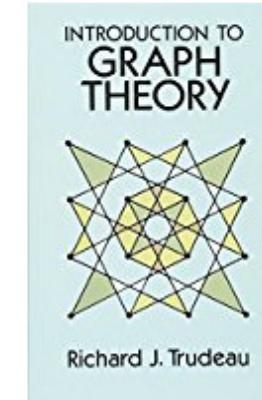
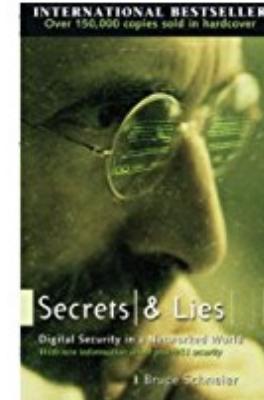
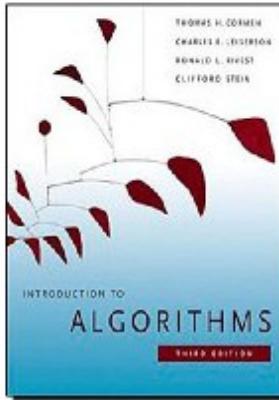
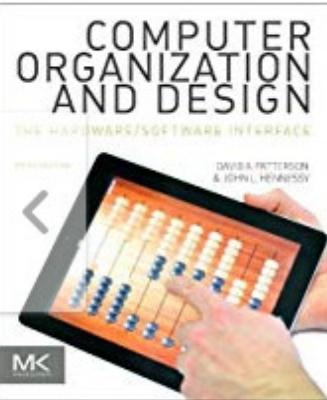
- Business
  - Financial transactions: walmart headquarter, stock traders, etc.



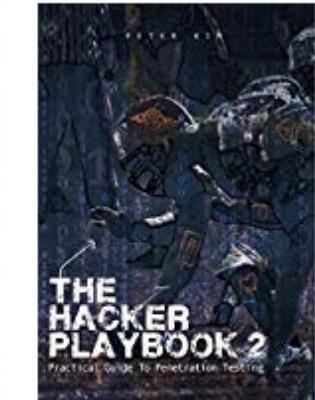
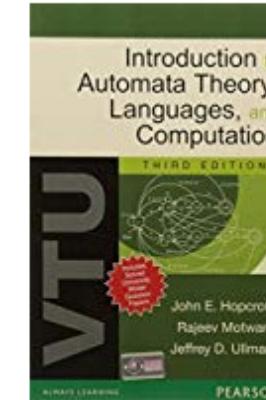
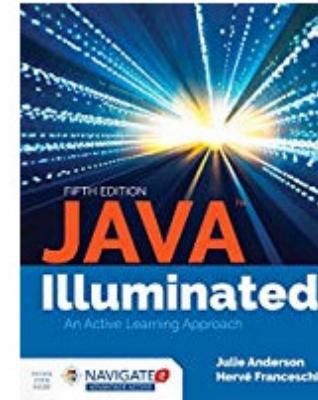
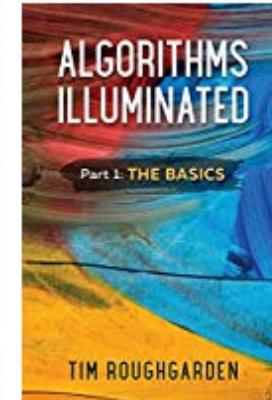
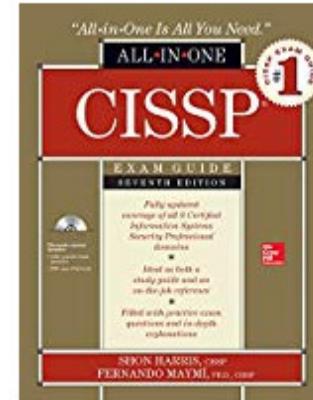
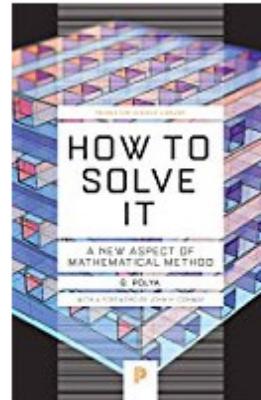
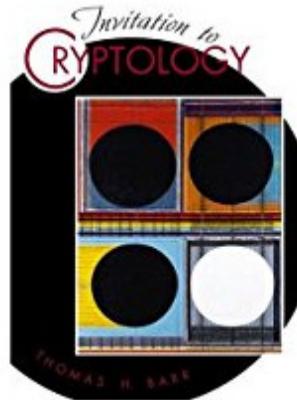
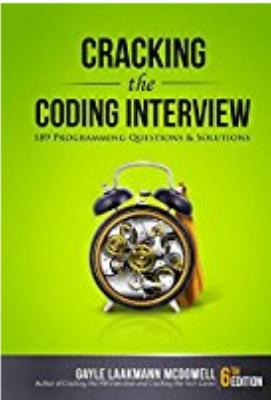
- Social media
  - Weblogs, twitter feeds, click streams, WeChat, etc.



## Recommendations for you in Books

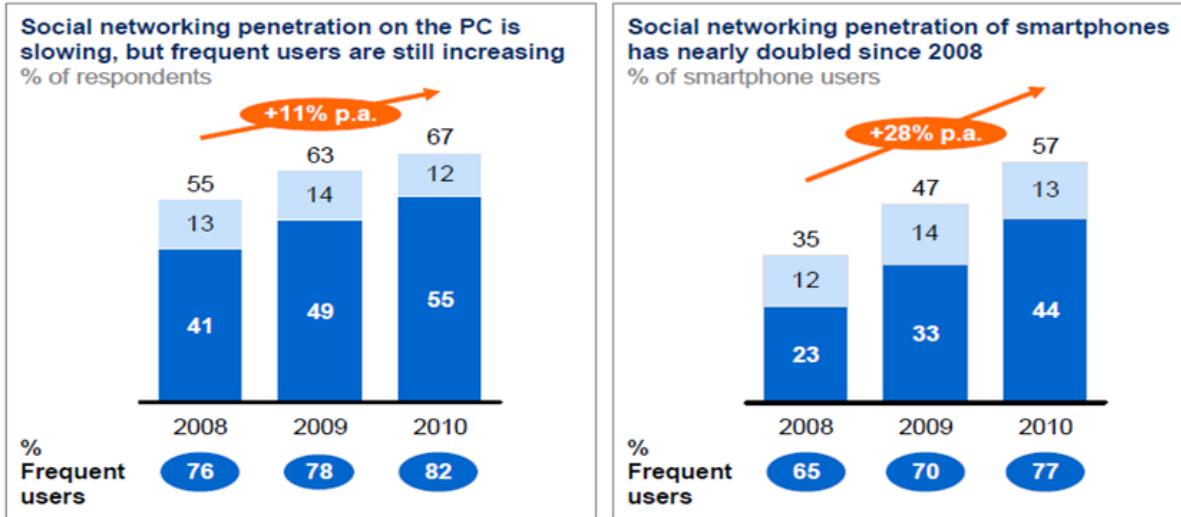


## Recommendations for you in Books



# Example - Recommendation

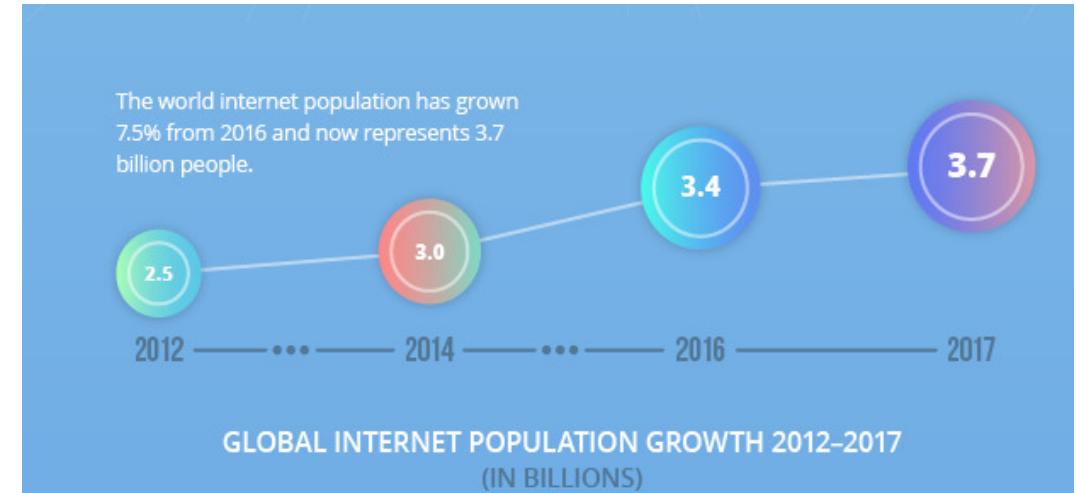
# Social networks and mobile devices



## DATA NEVER SLEEPS 6.0

How much data is generated *every minute*?

There's no way around it: big data just keeps getting bigger. The numbers are staggering, but they're not slowing down. By 2020, it's estimated that for every person on earth, 1.7 MB of data will be created every second. In our 6th edition of Data Never Sleeps, we once again take a look at how much data is being created all around us every single minute of the day—and we have a feeling things are just getting started.



## IBM CIO monitoring categories

## Monitoring filter

Select CIO Catetory(-ies): EXECDB BLADE HRTEANNT IBM SecurityAnalysis SWG WATSON

or Word: Egypt

GO STOP RESUME

language: Arabic

Total Tweets: 231

Positive: 35 15%

Negative: 31 13%

EGYPT wearing @RawyaRageh beauty brutality Mor  
e || Am Egypt's 12 police hijab Egypt's 12 police  
ozen sponge allege Port Egypt than Cairo  
you my Egyptian Said egypt lady call

**Saloom Butilla** @SaloomButilla  
إعفاء الصفيرين الغرفة في البحرين على المرافق العامة ورجال الأمن #Bahrain #Egypt #Syria #KSA #UAE #News h ....

**Translation:** RT \*@Lion\_King\_Bhr\*: The traitors in Bahrain Safavid attack on public utilities and security men, 2/19/2013 \*LBahrain\* #Egypt \*LSyria\* \*LKSA\* \*LUAE\* \*LNews\* h \*...\* --Wed Feb 20 17:57:58 2013

**Zenza Raggi fan-club** @Zenzaclub  
Private Gold 64: Cleopatra 2 // A sect that worships ancient Egypt is attempting to bring Cleopatra back to life... http://t.co/TcvMDiwb --Wed Feb 20 17:57:53 2013

**@SH\_QalamSara** منقرضة هاتم RT @HebaFarooq: An #Egypt-ian beauty :) ▶ http://t.co/S9BZb5f3 --Wed Feb 20 17:57:53 2013

**Mona Metwally** @monametwally  
مرتضى محتاج متبر عن دم AB+ بمستوى الجامعه بالاسعاليه فضيله دم آب موجود 01024705247 #Egypt # مصر http://t.co/5oO6mtZ5.  
**Translation:** RT \*@EgyBloodBank\*: A

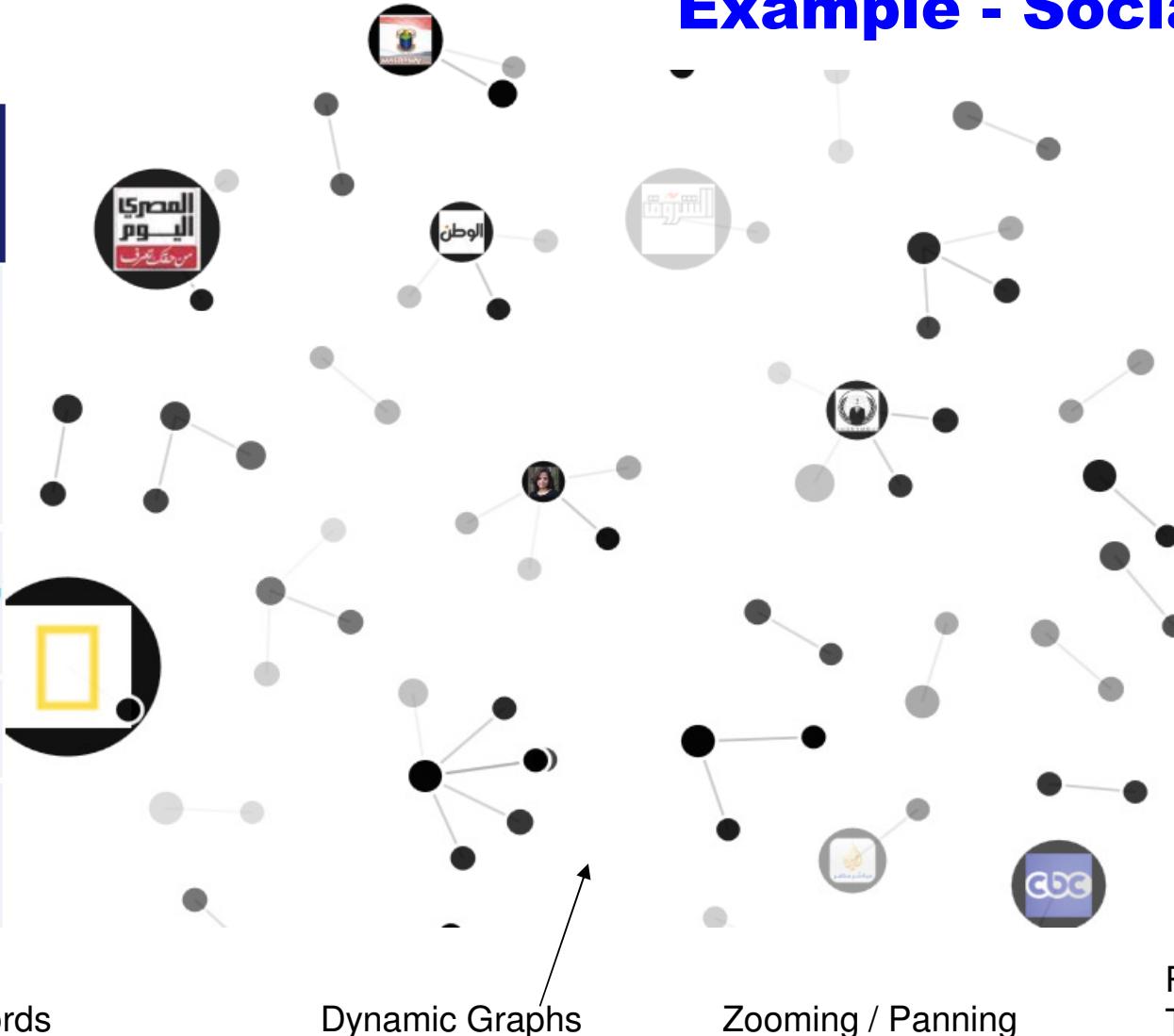
Live Tweets, Sentiment, Keywords

Dynamic Graphs

Zooming / Panning

Real-Time Translation, Locations, Top Retweets

## Example - Social Media Monitoring



**@1Derland** @1Derland 48,230 --> @1DRana 157  
And One Way Or Another is also number 1 in Guatemala, Peru, Israel, Brazil, Egypt and Panama! OMG  
@Lion\_King\_Bhr 44,12025 --> @SaloomButilla 1351

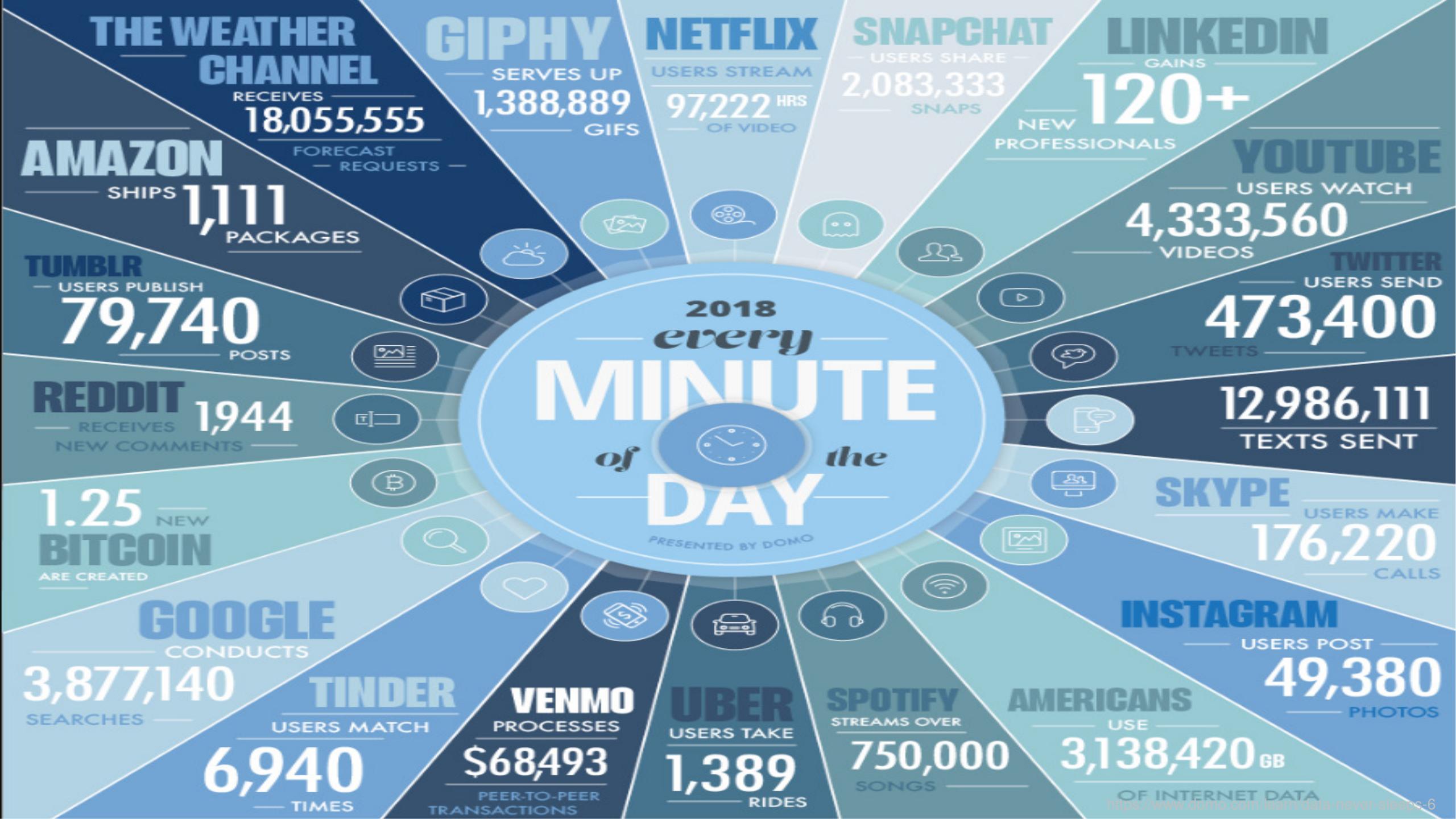
إعفاء الصفيرين الغرفة في البحرين على المرافق العامة ورجال الأمن #Bahrain #Egypt #Syria #KSA #UAE #News http://t.co/M18TdDE4.  
**Translation:**

**Vote4Squash** @Vote4Squash 42,4123 --> @JamesOxbury 22  
Big thanks to all who #vote4squash! There were over 5k tweets sent worldwide reaching over 1.3mil ppl trending in M'sia, Aus, Egypt & the UK

**NatGeo** @NatGeo 38,3039548 --> @abeenueve 216  
Now under a state of emergency, Egypt's Port Said flourished in the '20s http://t.co/N5mcFM6m

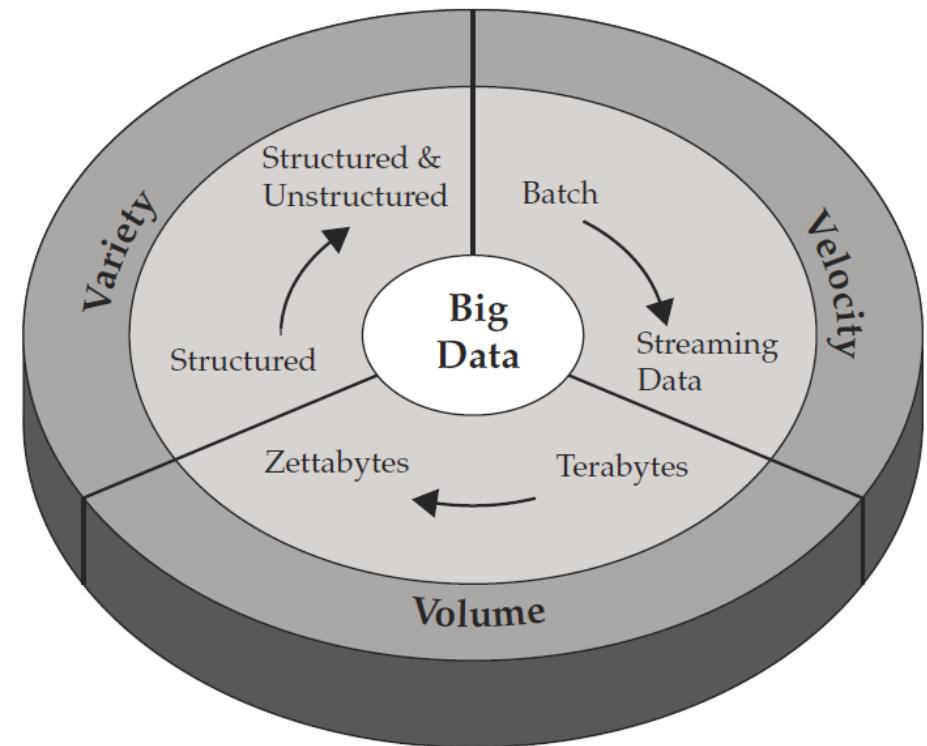
**EgyptBloodBank** @EgyBloodBank 29,5003 --> @monametwally 846  
مرتضى محتاج متبر عن دم AB+ بمستوى الجامعه بالاسعاليه فضيله دم آب موجود 01024705247 #Egypt # مصر http://t.co/5oO6mtZ5.  
**Translation:**

**@ABplus** @ABplus 29,4470 --> @monametwally 846



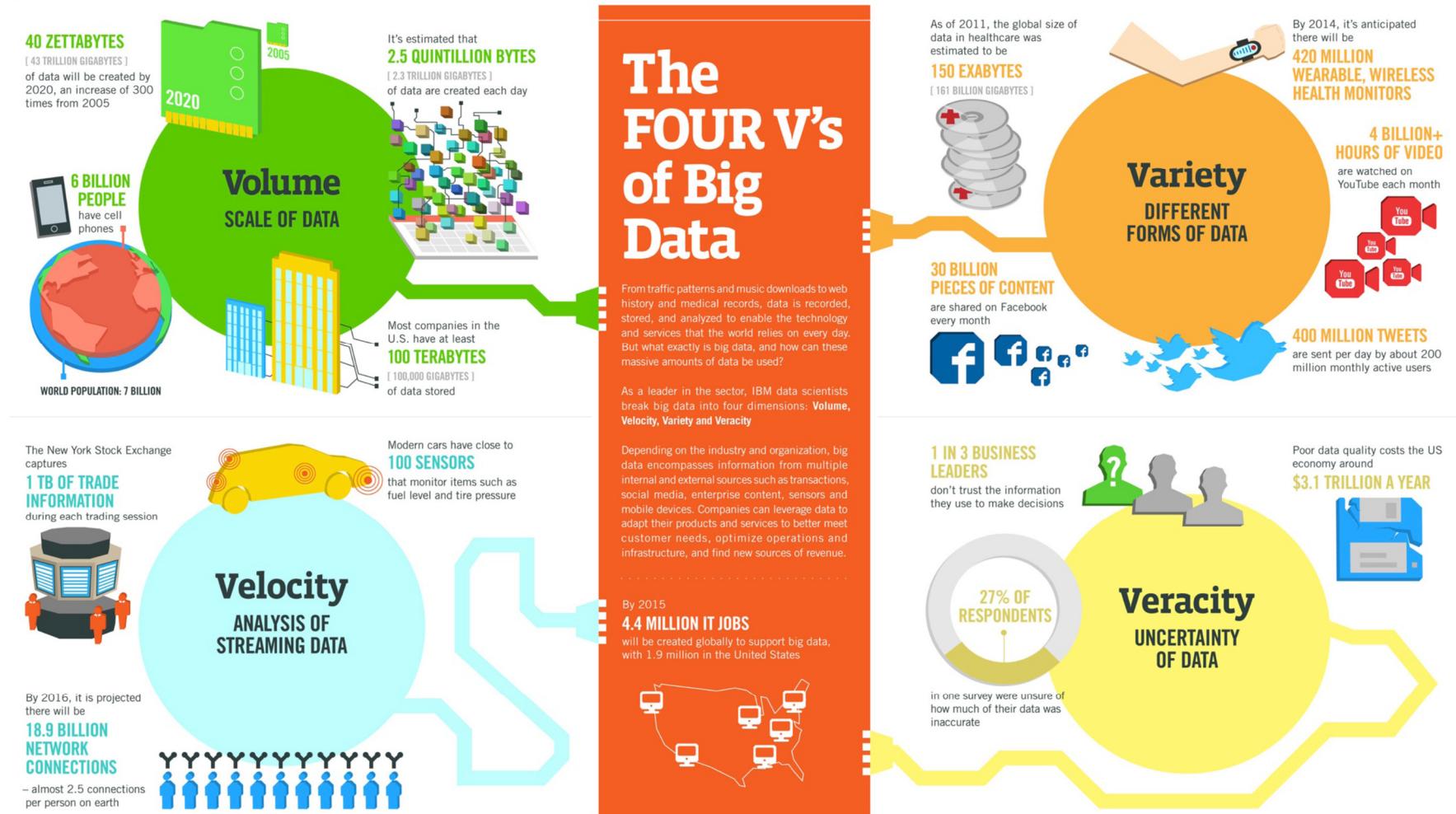
# Characterization of big data

- Types, quality, relationships, and processing of data



# Characterization of big data

- Four V's



# Typical operations on data sets

- Smart sampling of data
  - Reducing the original data while not losing the statistical properties of data
- Finding similar items
  - Efficient multi-dimensional indexing
- Incremental updating of the models
  - (v.s. building models from scratch)
  - Crucial for streaming data
- Distributed linear algebra
  - Dealing with large sparse matrices
- Supervised learning
  - Classification, regression, etc.
- Non-supervised learning
  - Clustering, different types of decompositions, etc.

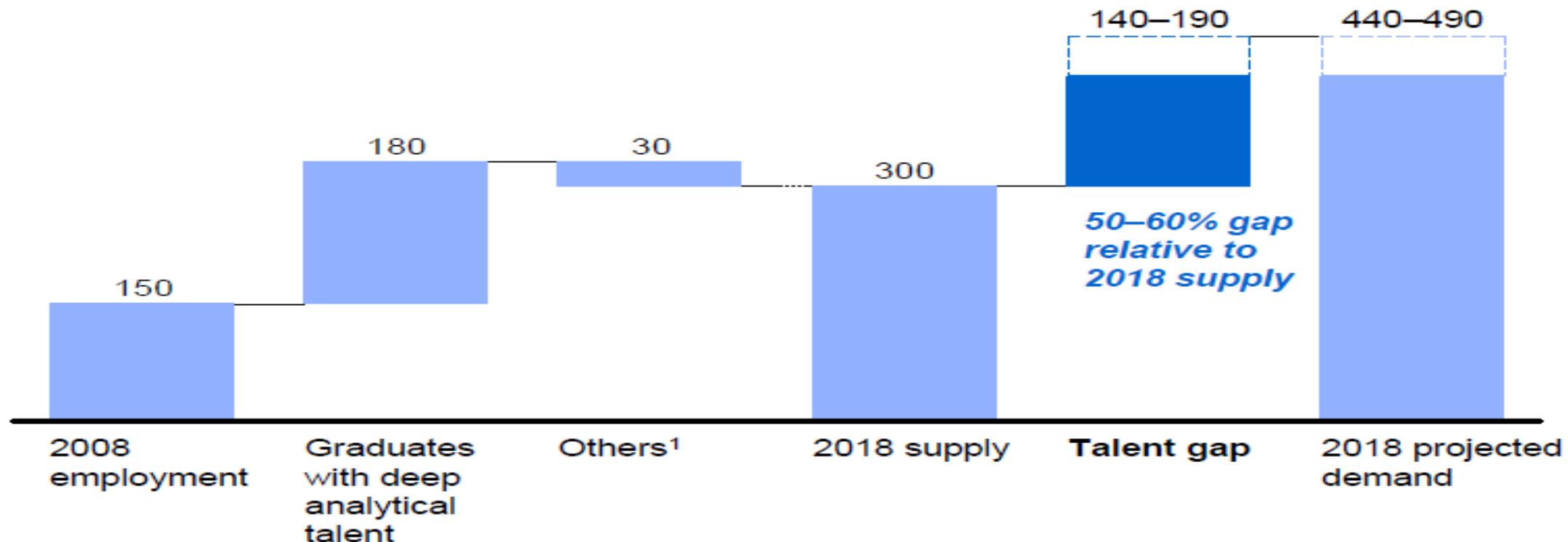


# Predicted lack of talent for big data

Demand for deep analytical talent in the United States could be 50 to 60 percent greater than its projected supply by 2018

Supply and demand of deep analytical talent by 2018

Thousand people

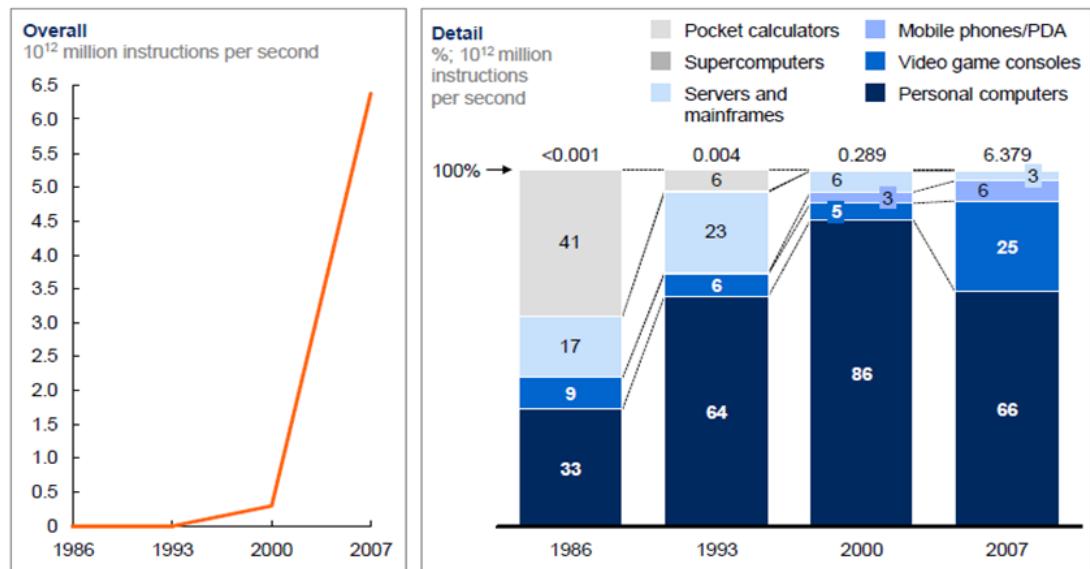
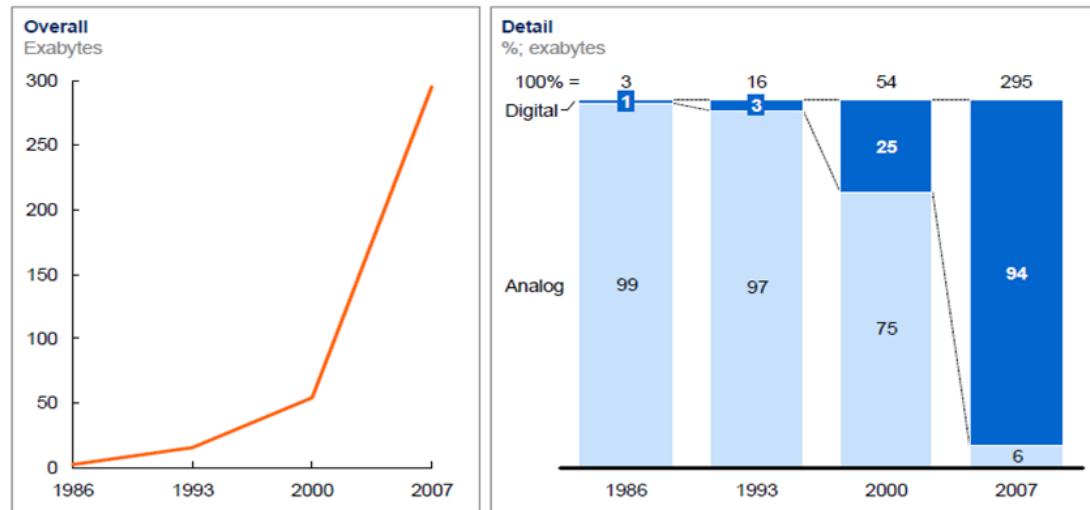


<sup>1</sup> Other supply drivers include attrition (-), immigration (+), and reemploying previously unemployed deep analytical talent (+).

SOURCE: US Bureau of Labor Statistics; US Census; Dun & Bradstreet; company interviews; McKinsey Global Institute analysis

# Key enablers

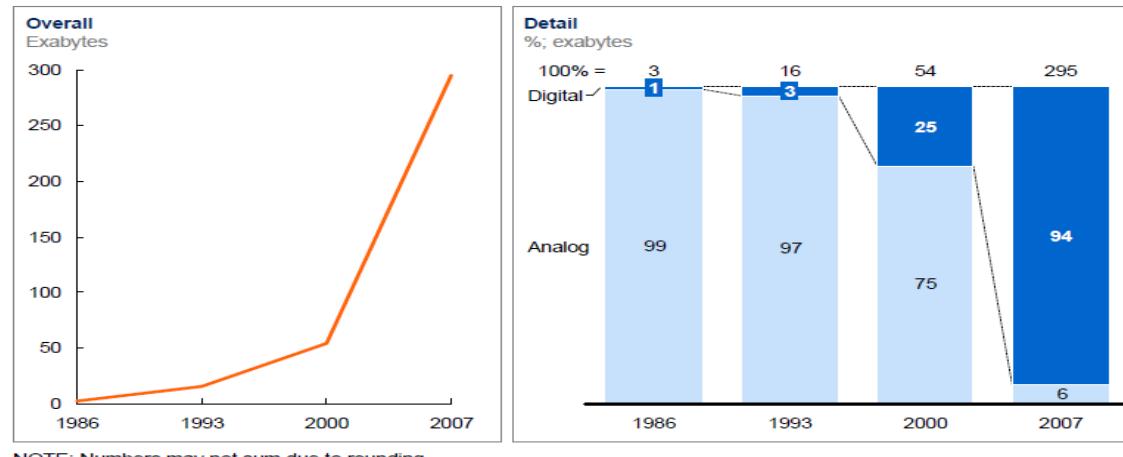
- Increase of storage capabilities
- Increase of computing power
- Increase of network capacities
- Availability of data sets



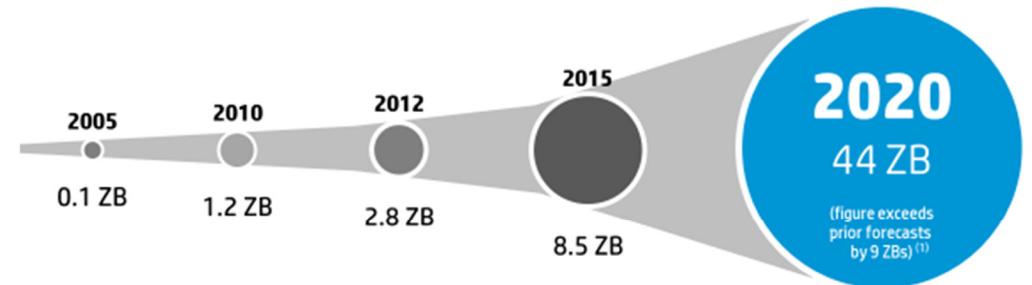
# Increase of storage capabilities

Data storage has grown significantly, shifting markedly from analog to digital after 2000

Global installed, optimally compressed, storage



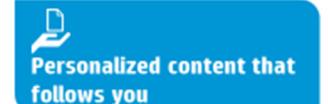
Data explosion outpacing technology



Next-generation competitive advantage delivered through:



Business insight at real-life speeds

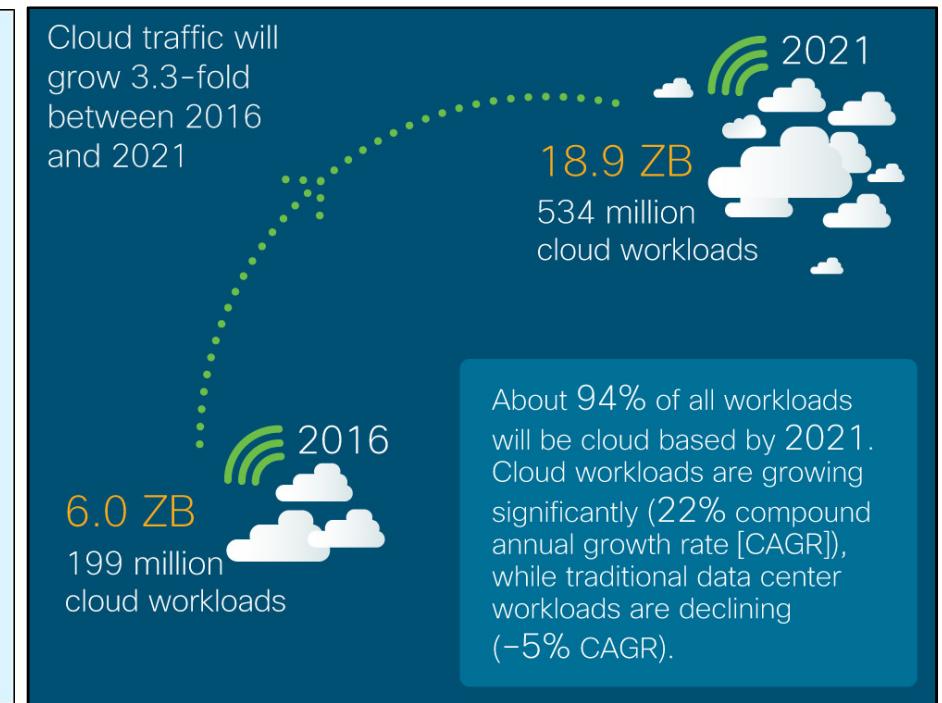
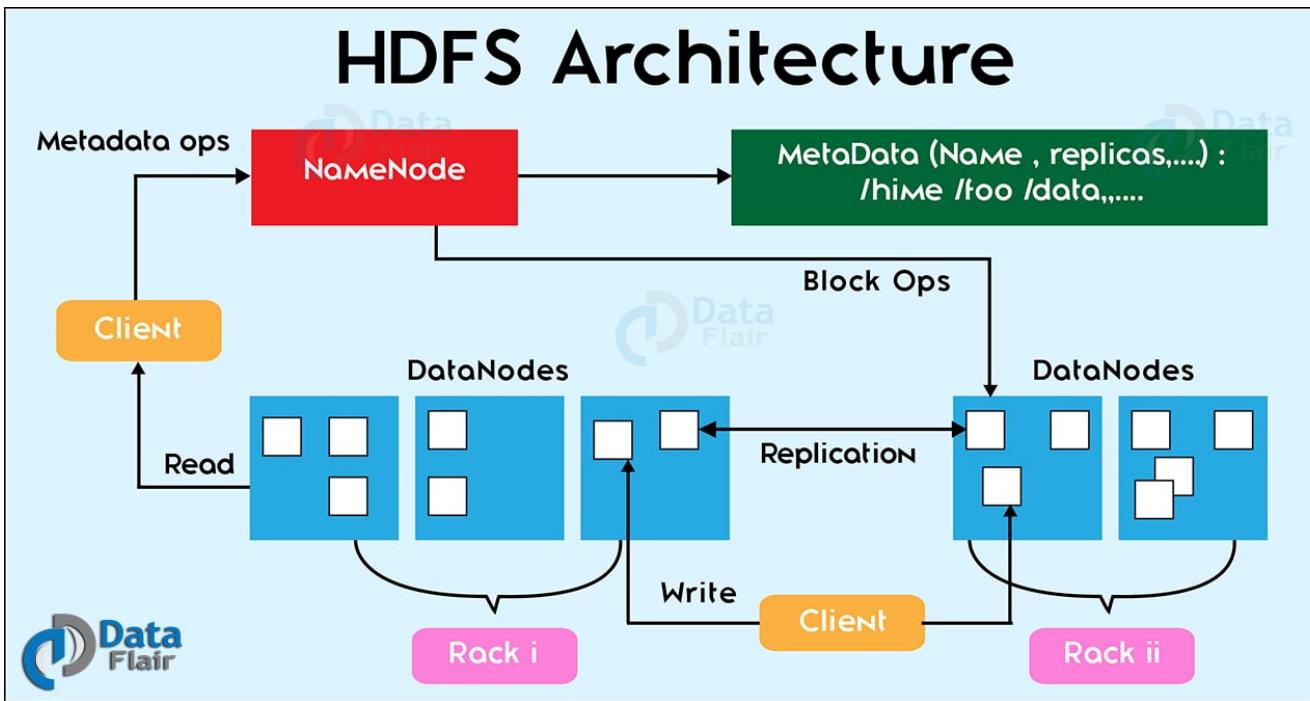


Personalized content that follows you



Questions that arise automatically from data

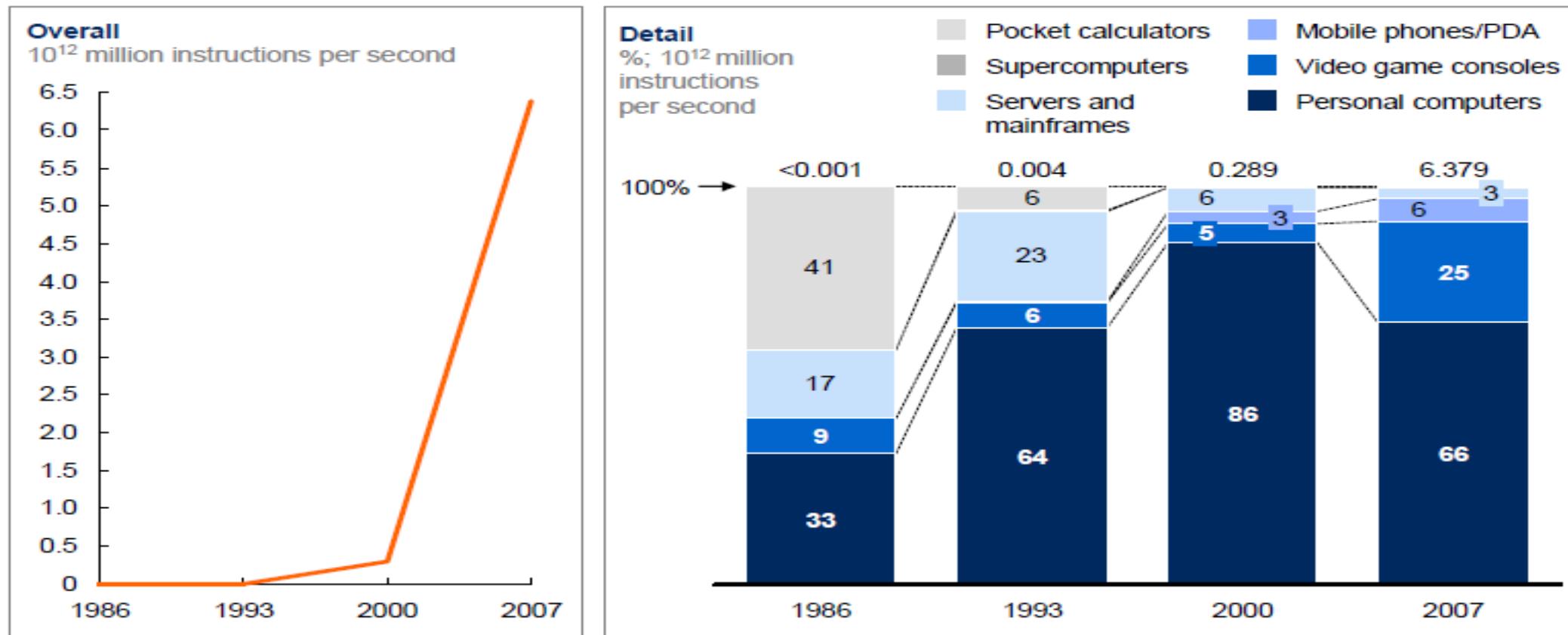
# Increase of storage capabilities



# Increase of computing power

Computation capacity has also risen sharply

Global installed computation to handle information



NOTE: Numbers may not sum due to rounding.

SOURCE: Hilbert and López, "The world's technological capacity to store, communicate, and compute information," *Science*, 2011

# More platforms, techniques, and tools



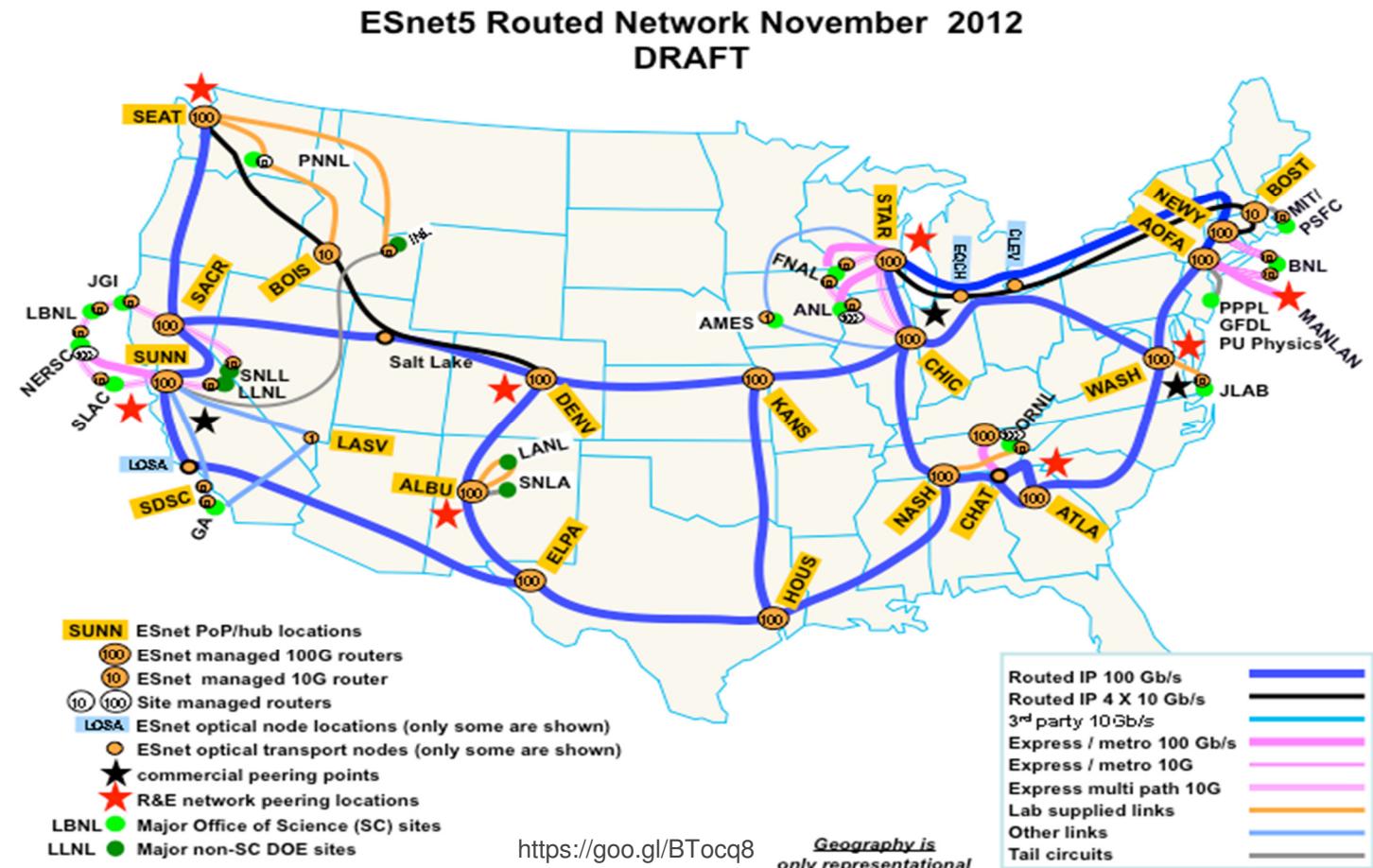
# Increase of network capacities

- HPNs: High-performance Networks
- e.g., Internet2



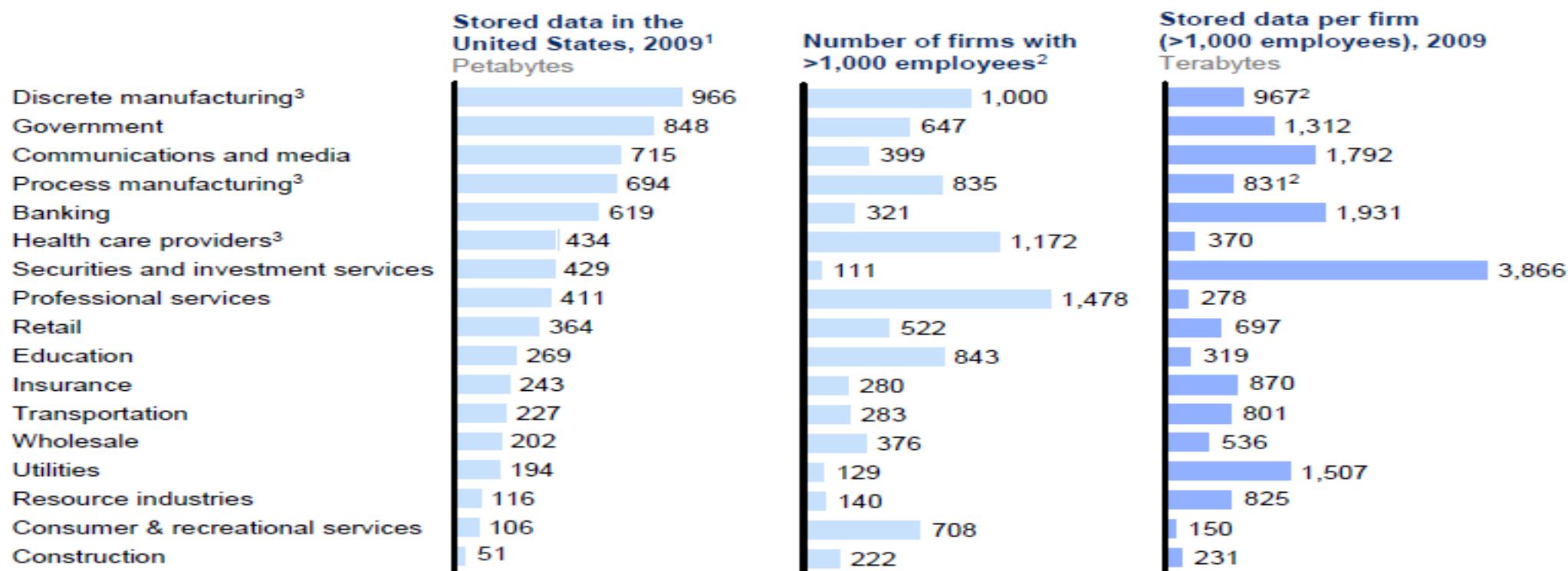
# Increase of network capacities

- HPNs: High-performance Networks
- e.g., ESnet



# Availability of data sets

**Companies in all sectors have at least 100 terabytes of stored data in the United States; many have more than 1 petabyte**



1 Storage data by sector derived from IDC.

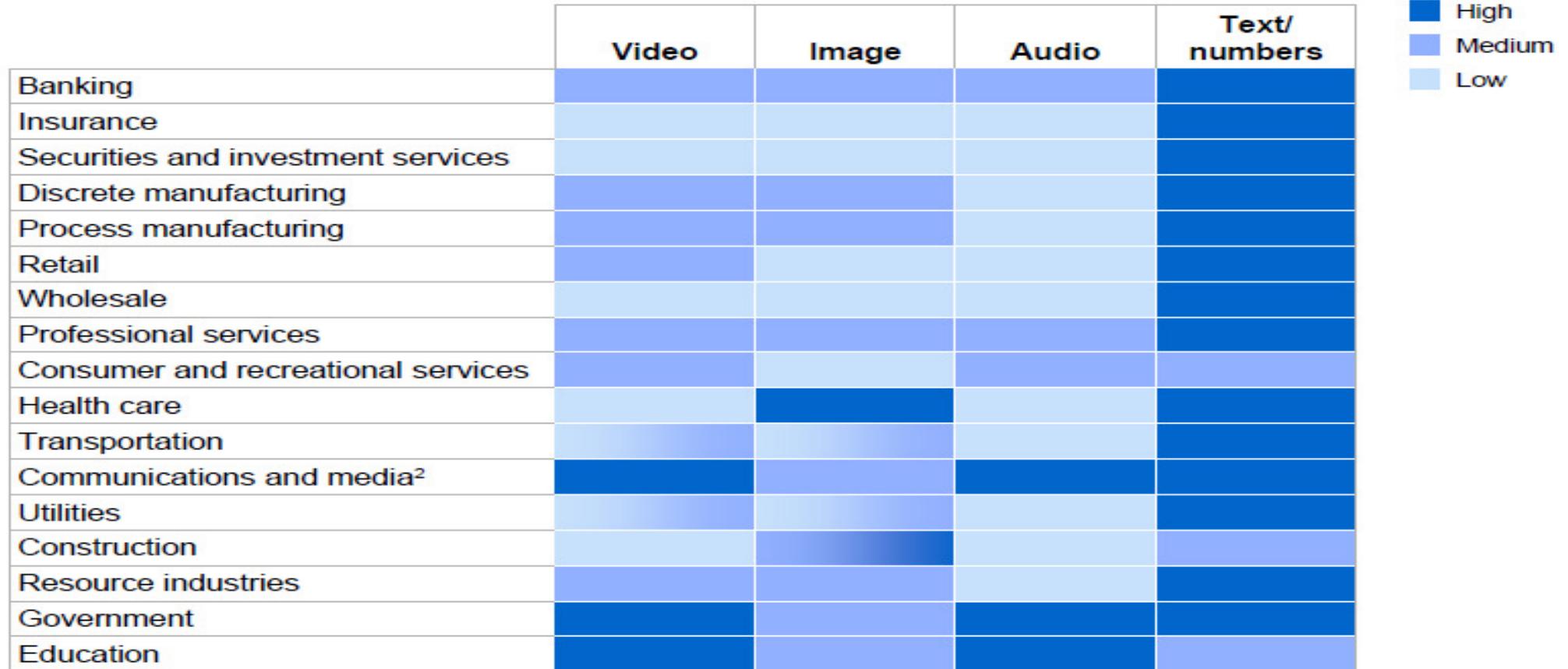
2 Firm data split into sectors, when needed, using employment

3 The particularly large number of firms in manufacturing and health care provider sectors make the available storage per company much smaller.

SOURCE: IDC; US Bureau of Labor Statistics; McKinsey Global Institute analysis

# Types of available data sets

The type of data generated and stored varies by sector<sup>1</sup>



1 We compiled this heat map using units of data (in files or minutes of video) rather than bytes.

2 Video and audio are high in some subsectors.

SOURCE: McKinsey Global Institute analysis

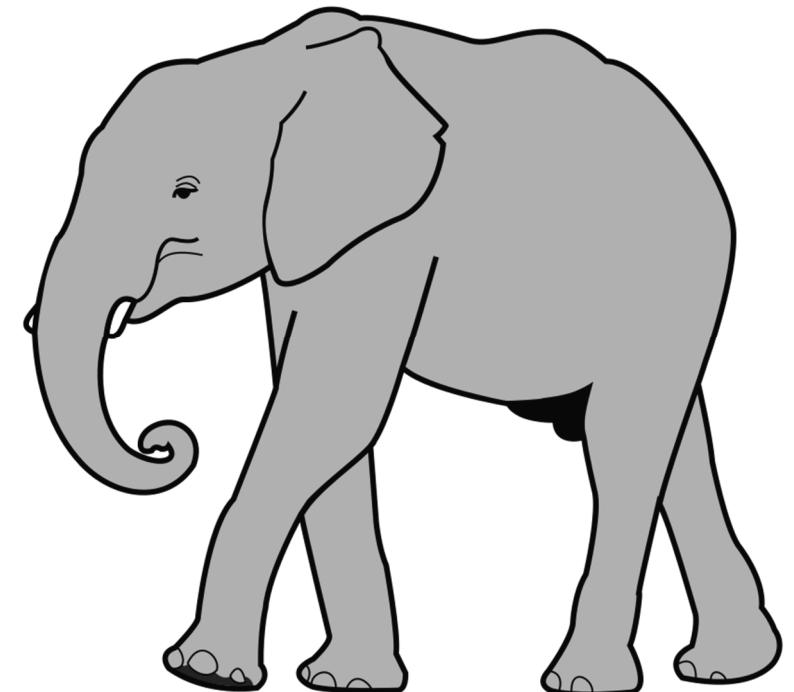
# Big-data Applications

- From **T**(erabyte) to **P**(etabyte), to **E**(xabyte), to **Z**(ettabyte), to **Y**(ottabyte), and beyond...
- Sciences: Simulation, Experimental, Observational
- Business: Financial Transactions
- Social Media: Weblogs, Twitter feeds, etc.

No matter which type of data is considered, we need  
a high-performance end-to-end computing solution  
to support data generation, storage, transfer,  
processing, and analysis!



V.S.



# Objective

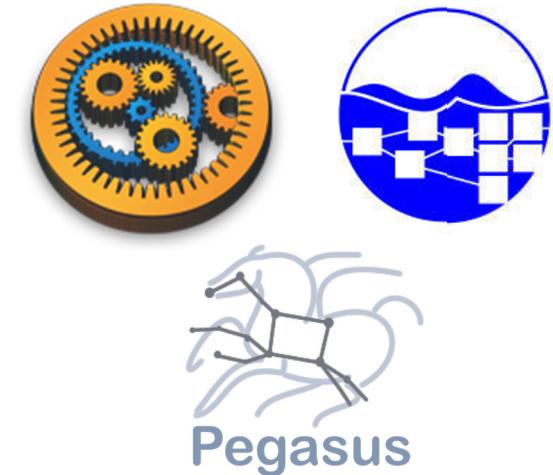
- A high-performance end-to-end computing solution to support
  - Data generation
  - Data storage
  - Data transfer
  - Data processing
  - Data analysis
  - Data visualization
  - ...



# Big-data Workflows



- Require massively distributed resources
  - **Hardware:** computing facilities, storage systems, special rendering engines, display devices (tiled display, powerwall, etc.), network infrastructures, etc.
  - **Software:** Domain-specific data analytics/processing tools, programs, etc.
  - **Data:** Real-time, archival
- Feature different complexities
  - **Simple case:** linear pipeline (a special case of DAG)
  - **Complex case:** DAG-structured graph
- Different application types have different performance requirements
  - **Interactive:** minimize total end-to-end delay for fast response
  - **Streaming:** maximize frame rate to achieve smooth data flow

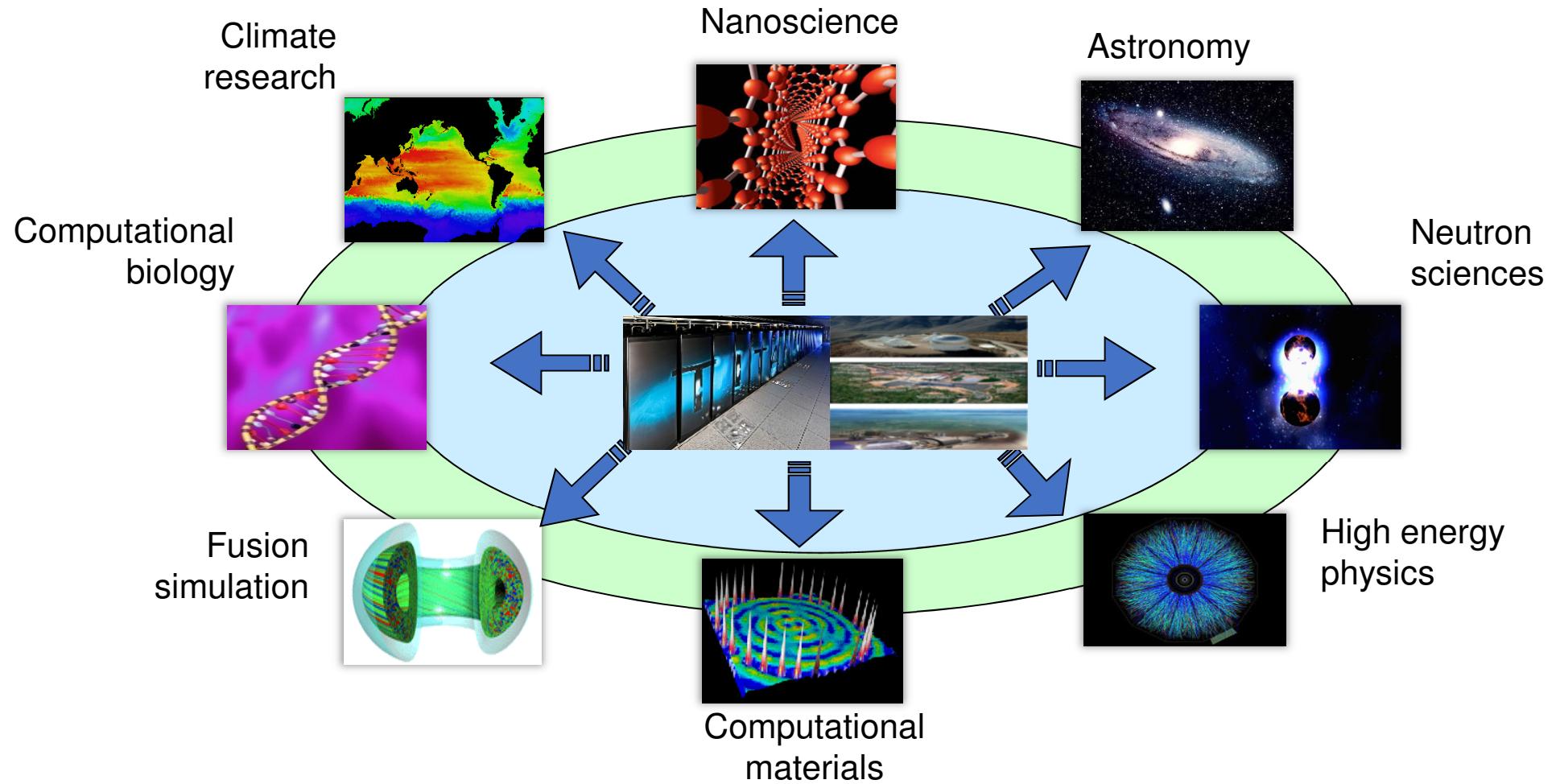


# Computing Paradigms: an Overview

- **Client-Server Model:** client-server computing refers broadly to any distributed application that distinguishes between service providers (servers) and service requesters (clients)
- **High-performance Computing (Supercomputing, Cluster Computing)**
  - Powerful computers: supercomputer, PC cluster
  - Used mainly by large organizations for critical applications, typically bulk data processing such as scientific computing, enterprise resource planning, and financial transaction processing
  - Programming models: MPI, OpenMP, CUDA, MapReduce/Hadoop, Spark
- **Grid Computing:** a form of distributed computing and parallel computing, whereby a “super and virtual computer” is composed of a cluster of networked, loosely coupled computers acting in concert to perform very large tasks
- **Cloud (Utility) Computing:** the packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility, such as electricity
- **Service-Oriented Computing:** cloud computing provides services related to computing while, in a reciprocal manner, service-oriented computing consists of the computing techniques that operate on software-as-a-service
- **Peer-to-peer (P2P) Computing:** distributed architecture without the need for central coordination, with participants being at the same time both suppliers and consumers of resources (in contrast to the traditional client–server model)
- **Mobile Computing:** computing on the go!

# **High-performance Computing (Supercomputing, Cluster Computing)**

# Supercomputing for scientific applications



# Why do we care about computing power?

## Computer Security: Exhaustive Key Search

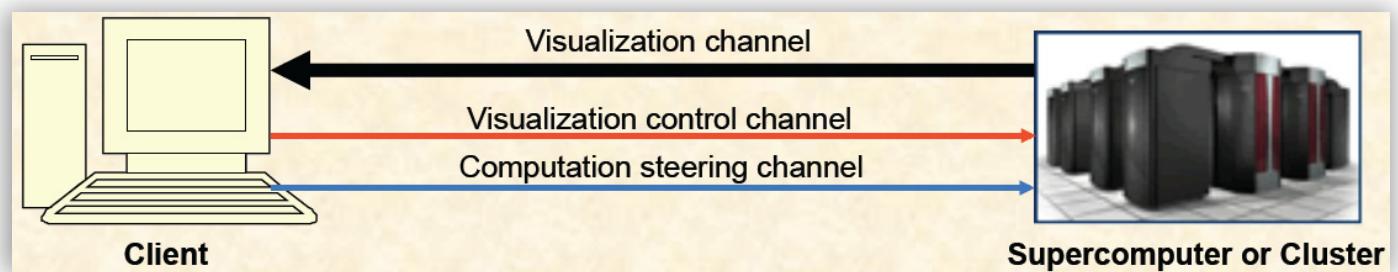
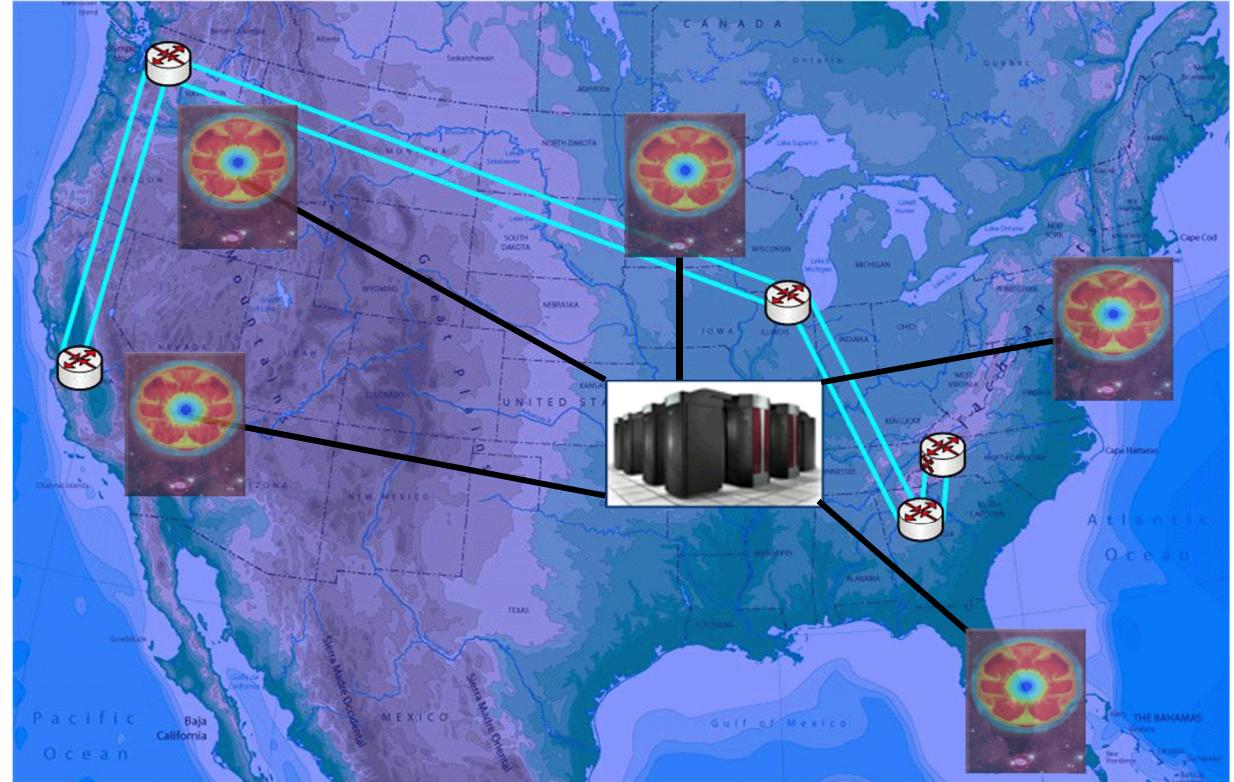
- Two types of security
  - Computational security
  - Unconditional security
- Attack on computational security
  - Always possible to simply try every key
  - Most basic attack, proportional to key size
  - Assume either know / recognize plaintext

Key Size (bits)	Number of Alternative Keys	Time required at 1 decryption/ $\mu$ s	Time required at $10^6$ decryptions/ $\mu$ s
32	$2^{32} = 4.3 \times 10^9$		
56	$2^{56} = 7.2 \times 10^{16}$		
128	$2^{128} = 3.4 \times 10^{38}$		
168	$2^{168} = 3.7 \times 10^{50}$		
26 letters (permutation)	$26! = 4 \times 10^{26}$		

# Science is becoming data-driven

## TSI: Terascale Supernova Initiative

- Collaborative project
  - Supernova explosion
- TSI simulation
  - 1 terabyte a day with a small portion of parameters
  - From TSI to PSI to ESI
- Transfer to remote sites
  - Interactive distributed visualization
  - Collaborative data analysis
  - Computation monitoring
  - Computation steering



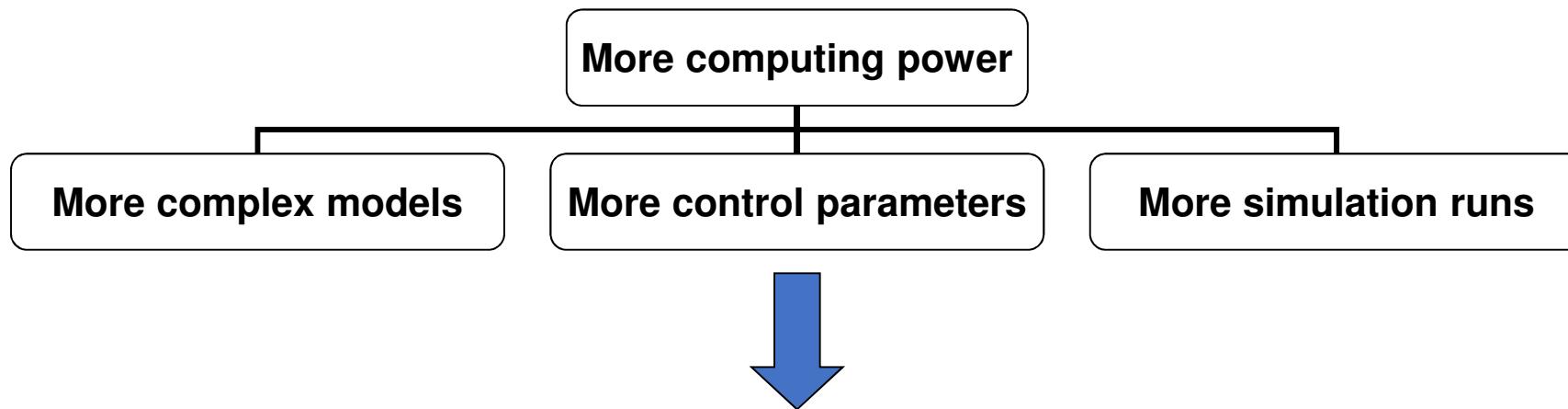
# **Computing power continues to increase over time**

DOE NLCF, Cray XK, IBM BlueGene, SGI Altix

- Two next-generation supercomputers:

- ORNL's **Summit**
- LLNL's **Sierra**

- Use NVIDIA® Tesla® GPU accelerators and NVIDIA NVLink™ high-speed interconnect technology on next-generation IBM POWER servers
- Each will be significantly faster than today's fastest supercomputer, achieving between 100 to 300 petaflops of peak performance



**Colossal amounts of scientific datasets!**

# Increase of computing power

## Computer Cluster



<http://www.mmds.org>



<http://aggregate.org/press001122.html>

# Increase of computing power

- World-class facilities



## US once again boasts the world's fastest supercomputer

The US Department of Energy on Friday unveiled Summit, which can perform 200 quadrillion calculations per second.

 By Stephanie Condon for Between the Lines | June 8, 2018 -- 1610 GMT (1110 PDT) | Topic: Artificial Intelligence

## China Launches Exascale Supercomputer Prototype

Aug 07, 2018

Email

Print

Text Size A A

Share



## Hyperion Provides Update to Exascale Efforts in US, China, Japan, and Europe

Michael Feldman | September 6, 2018 21:10 CEST

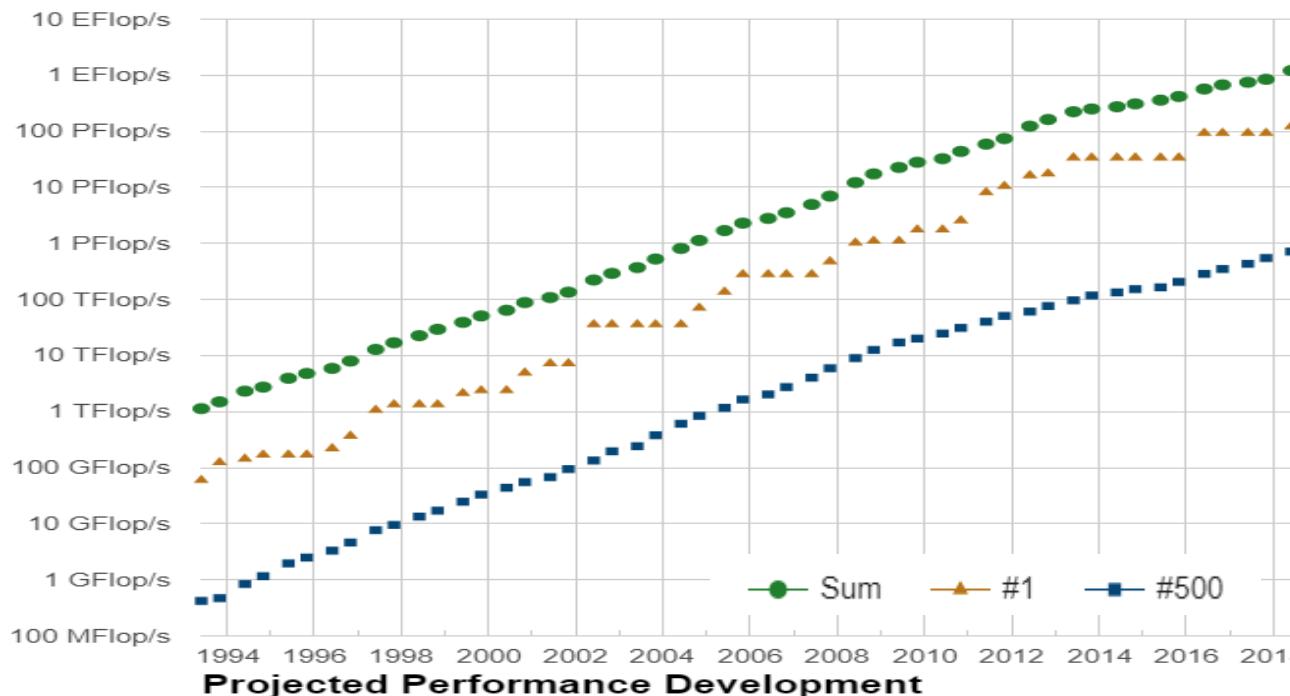
According to the latest analysis from Hyperion Research, the various global efforts to reach exascale supercomputing are making good headway. But in some cases, the decision to develop domestically-produced processors for these systems and the inclusion of new application use cases appears to be stretching out the timelines.

[Read more](#)

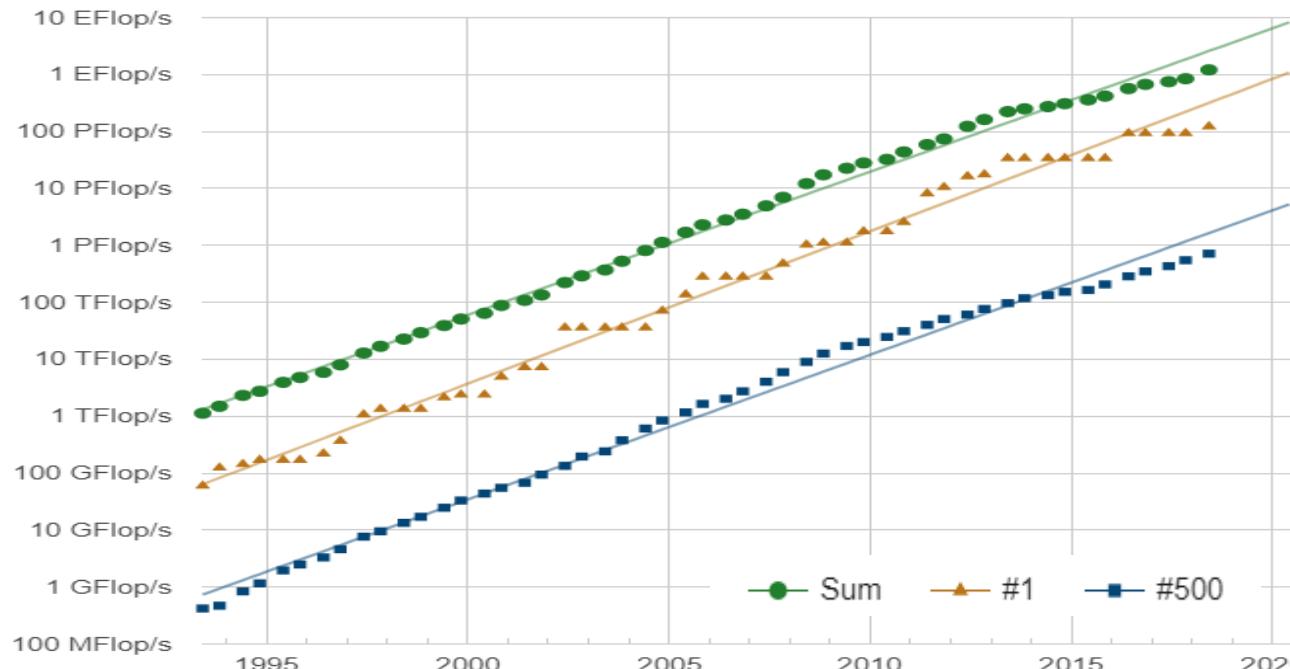
U.S.	EU
Sustained ES*: 2022-2023	
Peak ES: 2021	PEAK ES: 2023-2024
Vendors: U.S.	Pre-ES: 2021-2022
Processors: U.S. (some ARM?)	Vendors: Likely European
Initiatives: NSC/VECP	Processors: Likely ARM or RISC-V
Cost: \$600M per system, plus heavy R&D investments	Initiatives: EuroHPC
China	Cost: Over \$350M per system, plus heavy R&D investments
Sustained ES*: 2021-2022	
Peak ES: 2020	PEAK ES: Likely as a AI/ML/DL system
Vendors: Chinese (multiple sites)	Vendors: Japanese
Processors: Chinese (plus U.S.?)	Processors: Japanese
13th 5-Year Plan	Cost: \$80M-\$1B, this includes both system and the R&D costs
Cost: \$350-\$500M per system, plus heavy R&D	They will also do many smaller size systems

\* 1 exaflop on a 64-bit real application

## Performance Development



## Projected Performance Development

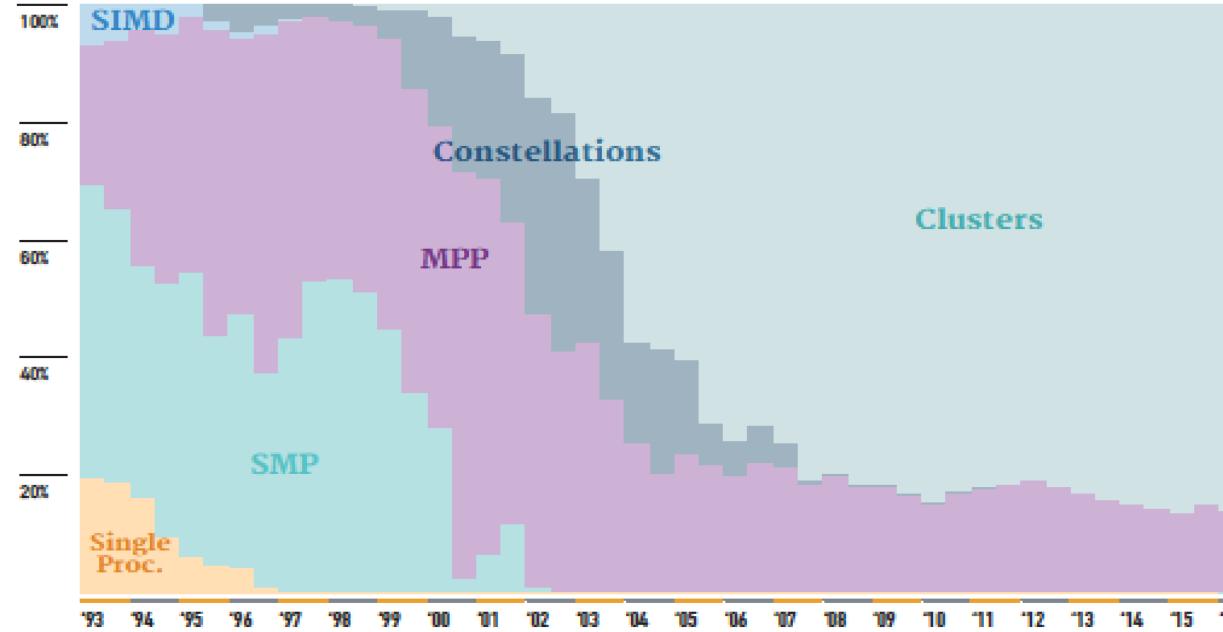


<https://www.top500.org/>

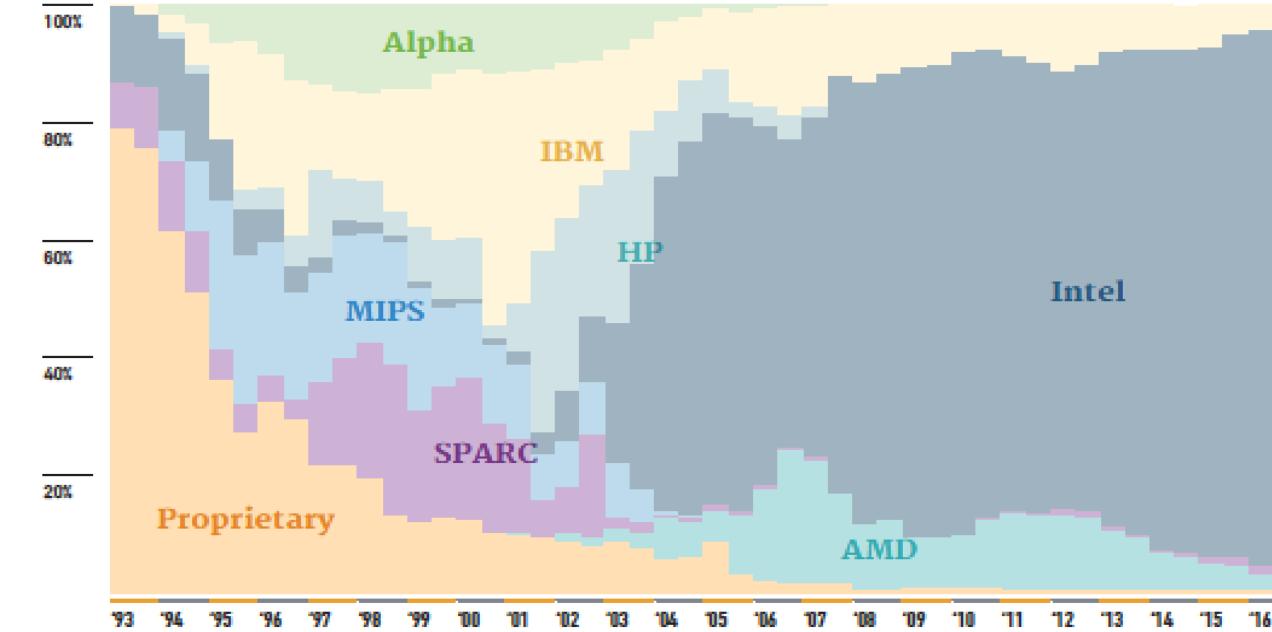


Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	<b>Summit</b> - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,282,544	122,300.0	187,659.3	8,806
2	<b>Sunway TaihuLight</b> - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
3	<b>Sierra</b> - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/NNSA/LLNL United States	1,572,480	71,610.0	119,193.6	
4	<b>Tianhe-2A</b> - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
5	<b>AI Bridging Cloud Infrastructure (ABCi)</b> - PRIMERGY CX2550 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	32,576.6	1,649
6	<b>Piz Daint</b> - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
7	<b>Titan</b> - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
8	<b>Sequoia</b> - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890

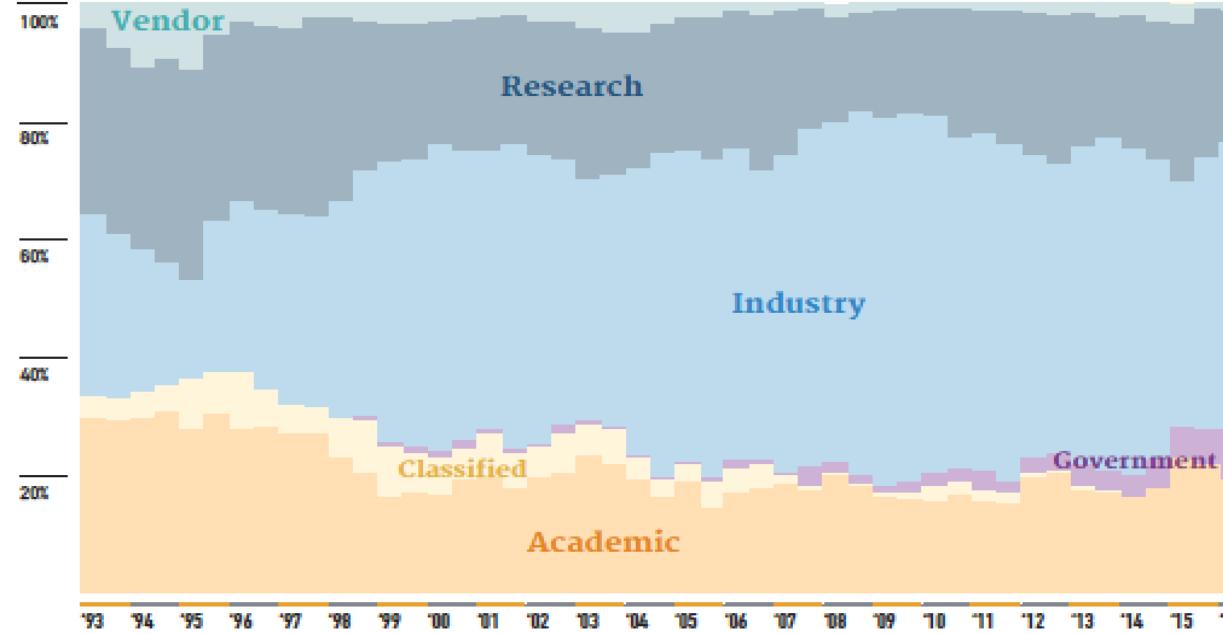
## ARCHITECTURES



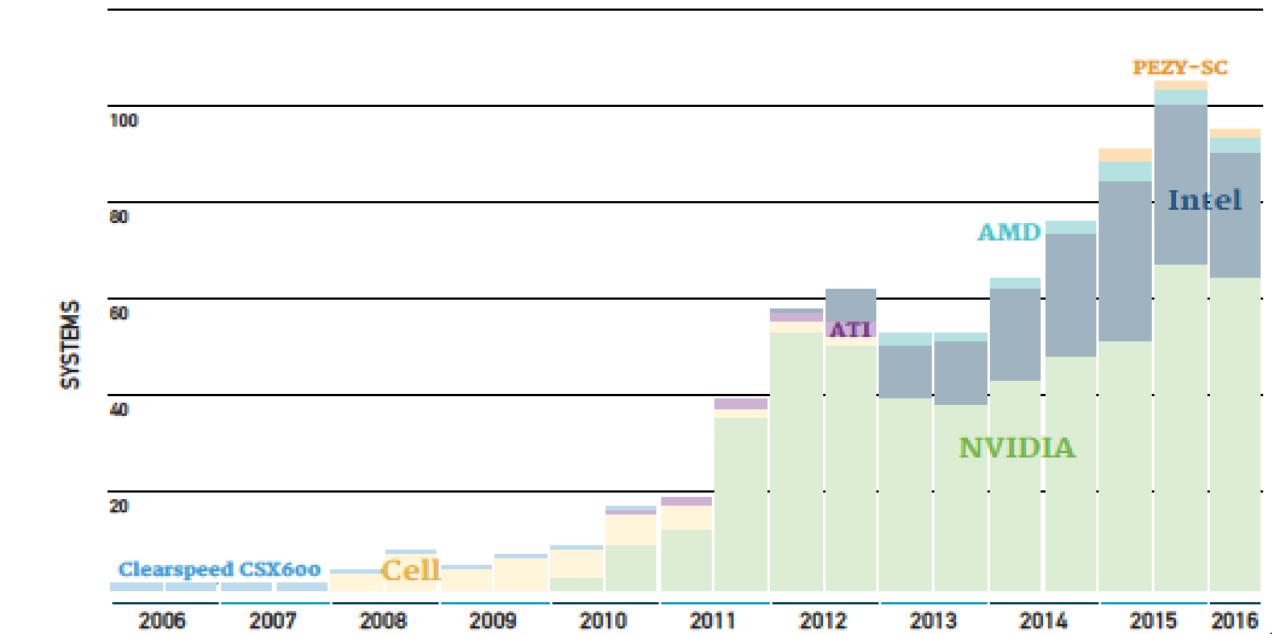
## CHIP TECHNOLOGY



## INSTALLATION TYPE



## ACCELERATORS/CO-PROCESSORS

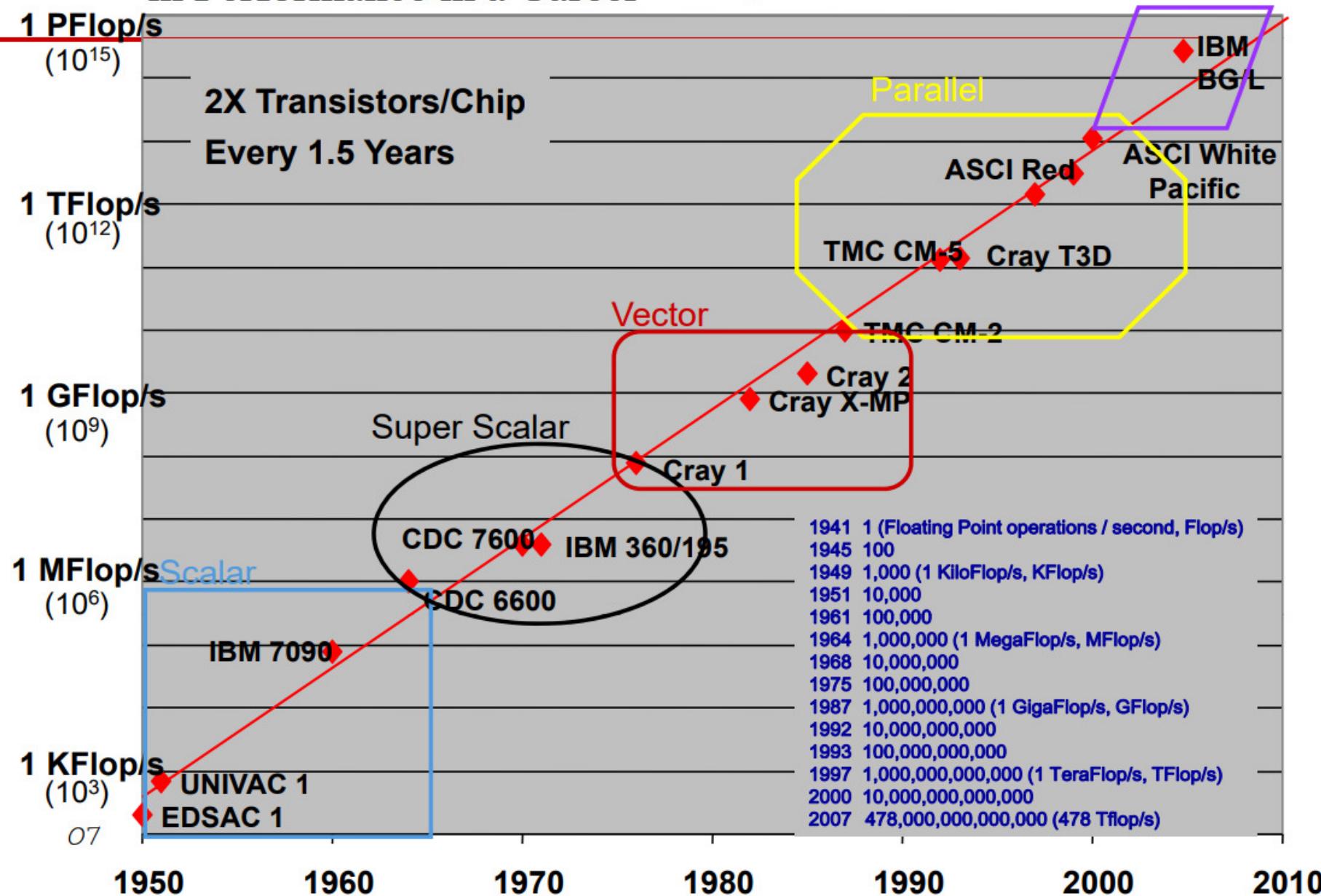


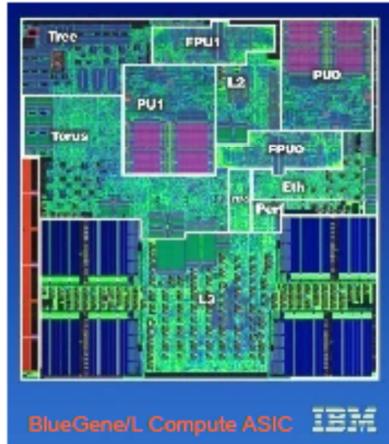
# ASCR Computing Upgrades At a Glance

System attributes	NERSC Now	OLCF Now	ALCF Now		NERSC Upgrade	OLCF Upgrade	ALCF Upgrades
Name Planned Installation	Edison	TITAN	MIRA	Theta 2016	Cori 2016	Summit 2017-2018	Aurora 2018-2019
System peak (PF)	2.6	27	10	>8.5	> 30	200	180
Peak Power (MW)	2	9	4.8	1.7	< 3.7	13.3	13
Total system memory	357 TB	710TB	768TB	>480 TB DDR4 + High Bandwidth Memory (HBM)	~1 PB DDR4 + High Bandwidth Memory (HBM)+1.5PB persistent memory	> 2.4 PB DDR4 + HBM + 3.7 PB persistent memory	> 7 PB High Bandwidth On- Package Memory Local Memory and Persistent Memory
Node performance (TF)	0.460	1.452	0.204	> 3	> 3	> 40	> 17 times Mira
Node processors	Intel Ivy Bridge	AMD Opteron Nvidia Kepler	64-bit PowerP C A2	Intel Knights Landing Xeon Phi many core CPUs	Intel Knights Landing many core CPUs Intel Haswell CPU in data partition	Multiple IBM Power9 CPUs & multiple Nvidia Volta GPUS	Knights Hill Xeon Phi many core CPUs
System size (nodes)	5,600 nodes	18,688 nodes	49,152	>2,500 nodes	9,300 nodes 1,900 nodes in data partition	~4,600 nodes	>50,000 nodes
System Interconnect	Aries	Gemini	5D Torus	Aries	Aries	Dual Rail EDR-IB	2 <sup>nd</sup> Generation Intel Omni-Path Architecture
File System	7.6 PB 168 GB/s, Lustre®	32 PB 1 TB/s, Lustre®	26 PB 300 GB/s GPFS™	10PB, 210 GB/s Lustre initial	28 PB 744 GB/s Lustre®	120 PB 1 TB/s GPFS™	150 PB 1 TB/s Lustre®

# A Growth-Factor of a Billion in Performance in a Career

Super Scalar/Vector/Parallel





## IBM BlueGene/L #1 212,992 Cores

**Total of 26 systems all in the Top176**

2.6 MWatts (2600 homes)

70,000 ops/s/person

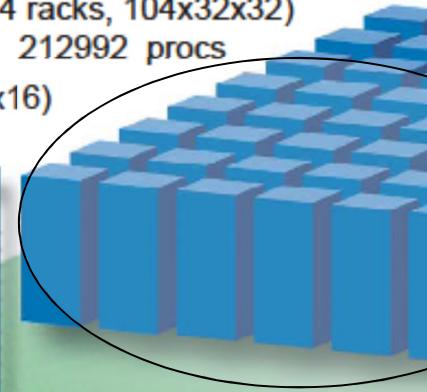
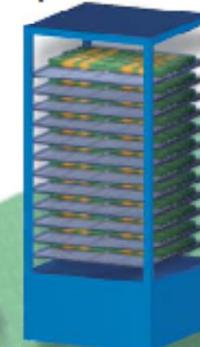
(104 racks, 104x32x32)

212992 procs

Rack

(32 Node boards, 8x8x16)

2048 processors

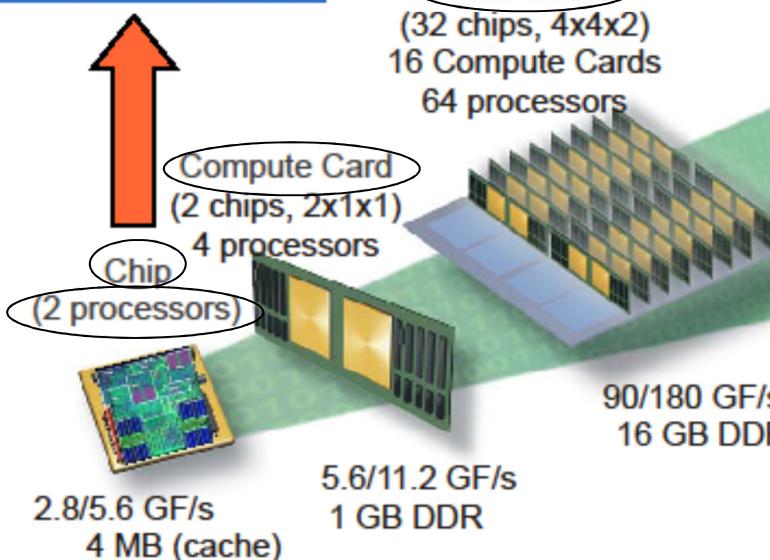


180/360 TF/s  
32 TB DDR

90/180 GF/s  
16 GB DDR

2.9/5.7 TF/s  
0.5 TB DDR

**Full system total of  
131,072 processors**

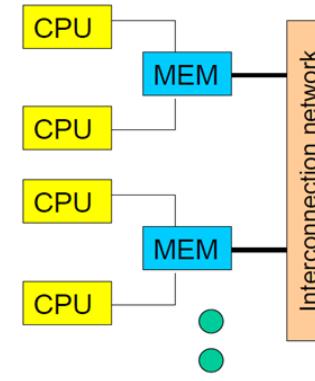
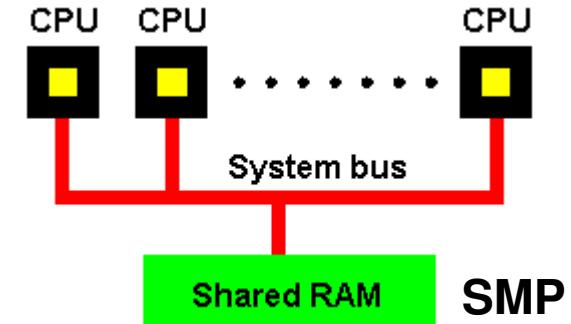


The compute node ASICs include all networking and processor functionality.  
Each compute ASIC includes two 32-bit superscalar PowerPC 440 embedded cores (note that L1 cache coherence is not maintained between these cores).  
(20.7K sec about 5.7hours; n=2.5M)

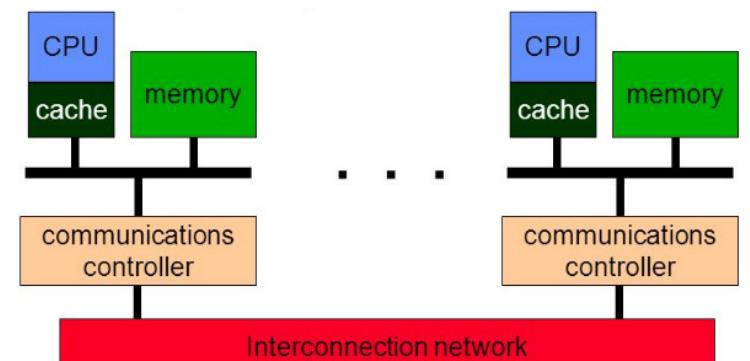
**“Fastest Computer”**  
**BG/L 700 MHz 213K proc**  
**104 racks**  
**Peak: 596 Tflop/s**  
**Linpack: 498 Tflop/s**  
**84% of peak**

# HPC Architecture

- SMP – symmetric multiprocessing (up to 64 CPUs)
  - All processors access common memory on the same rights
  - Used in desktops
- NUMA – nonuniform memory access
  - Global address space (as in SMP)
  - Faster access to local memory
  - Slower to remote
- Distributed memory multicompilers
  - Communication via messages
- Vector computers
  - Multiple functional units performing the same operation on vector registers (very long ones)
    - e.g., vector addition, dot product
  - Almost disappeared
- Distributed memory multiprocessors
  - MPP: Massively Parallel Processing
    - Tightly integrated
- Constellations
  - Clusters of supercomputers



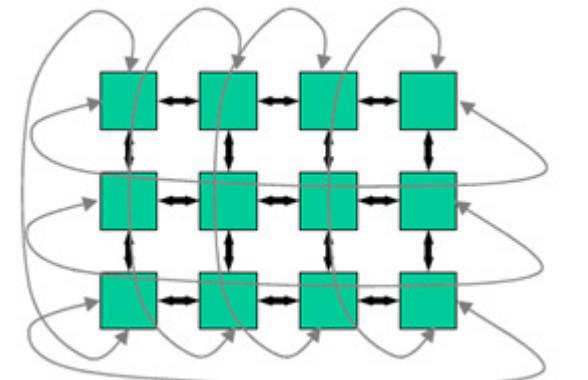
**NUMA**



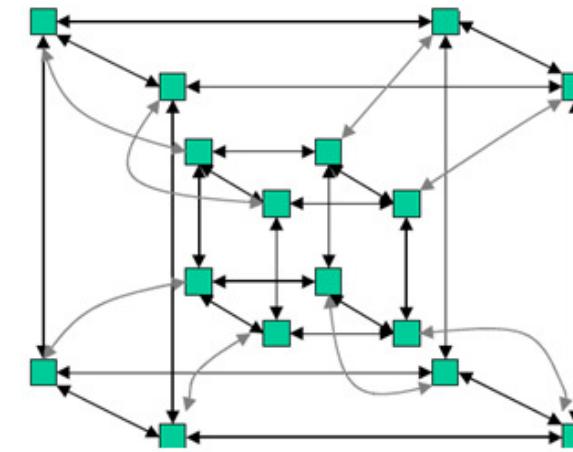
**Distributed memory multicompilers**

# Network Topologies

- Goal
  - Limited number of connections per node
  - Small width
  - Scalability
- Topologies
  - Mesh
  - Ring
  - Torus
  - Hypercube



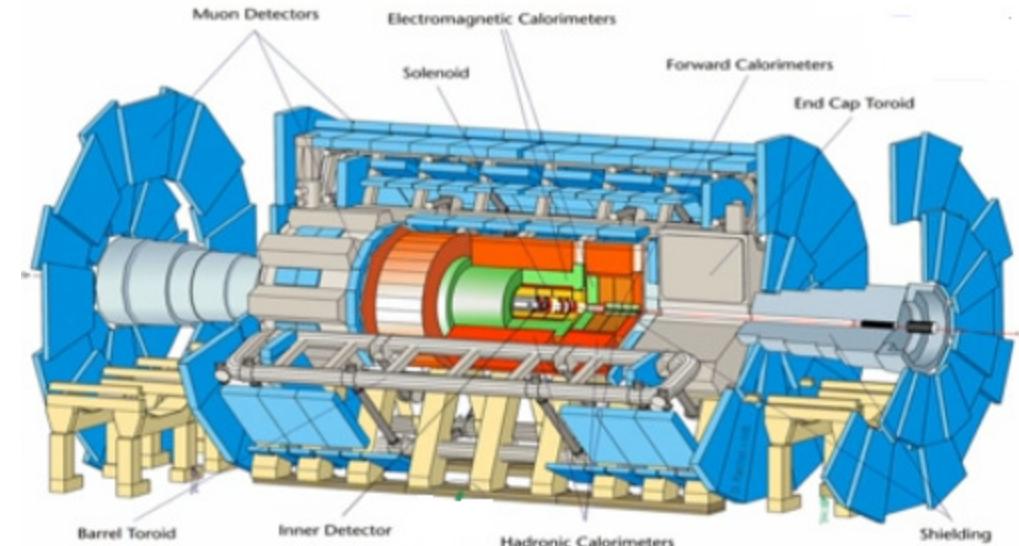
Torus



4D hypercube

# Parallel Application Example -- HEP

- Detector simulation
  - Simulate 10,000 events with Geant4
  - One event – (e.g.,) 1 minute
    - One week of computations!
- Trivial parallelization – give parts of all events to different CPUs
- Event analysis
  - Plot a quick histogram of 10,000 events
  - One event – (e.g.,) 0.1s
    - 20 minutes – this is not a quick histogram!



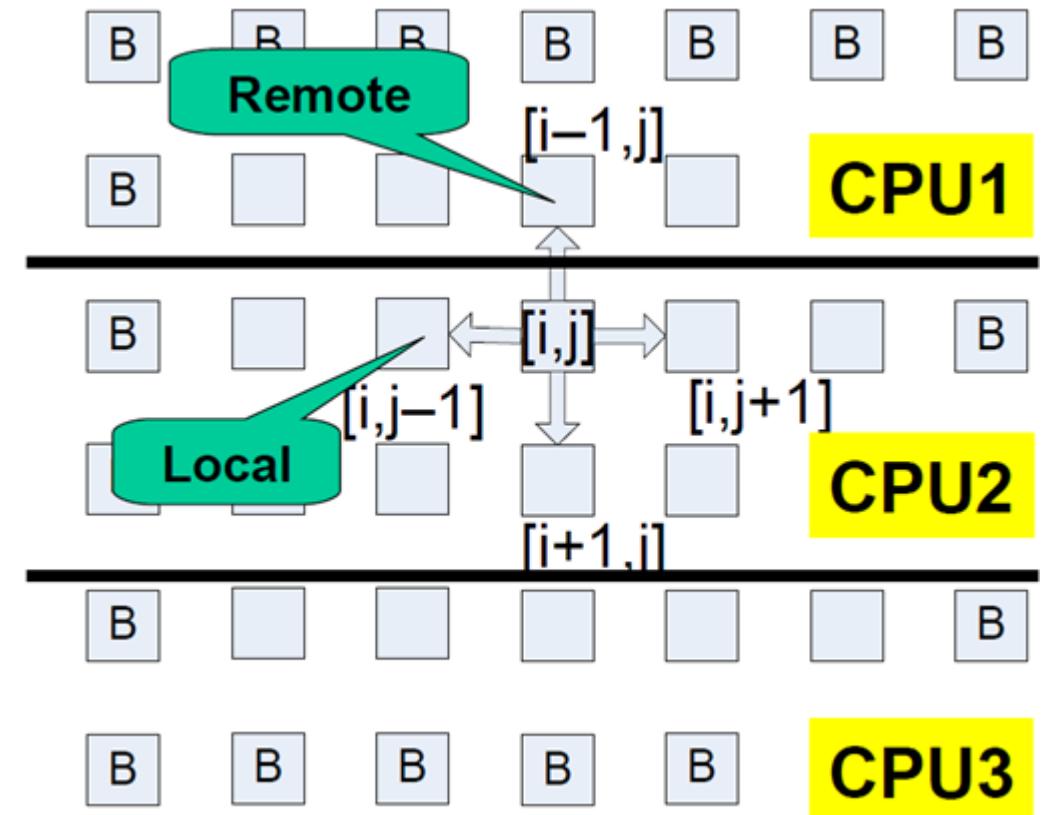
Trivial parallelization – each CPU analyzes a part of all the events.

At the end histograms are added. But making it interactive and transparent is a challenge.

# Nontrivial Example – SOR

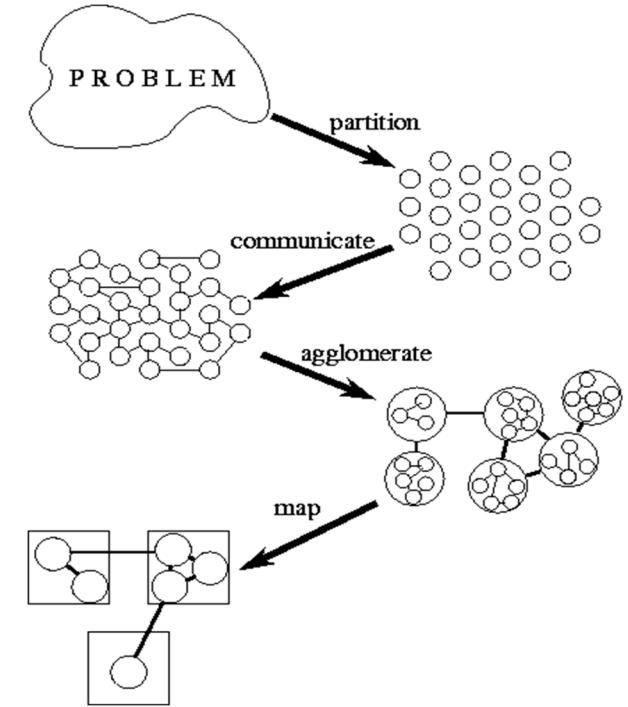
- SOR – Successive Overrelaxation
  - Solution to Laplace equation
- Boundary rows have to be transferred between CPUs

$$\begin{aligned} A_{new}[i, j] \\ = \frac{\alpha}{4} (A[i-1, j] + A[i+1, j] + \\ A[i, j-1] + A[i, j+1]) \\ +(1-\alpha)A[i, j] \end{aligned}$$



# Designing a Parallel Application

- **Partitioning:** decompose data and computations into small tasks
- **Communication:** analyze communication required to coordinate tasks
- **Agglomeration – reduce communication:** increase granularity, improve locality
- **Mapping – map processors to tasks**
  - Concurrent tasks on different CPUs
  - Frequently communicating tasks on the same CPU
- **Minimize communication overhead:** data locality is important
- **Load balancing:** make sure processors are never idle
  - Dynamic load balancing: divide work at runtime
- **Take system architecture into account!**
  - Fast local and slow remote memory for NUMA machines
  - Hardware for broadcasting, reduction
  - Faster access to neighboring nodes

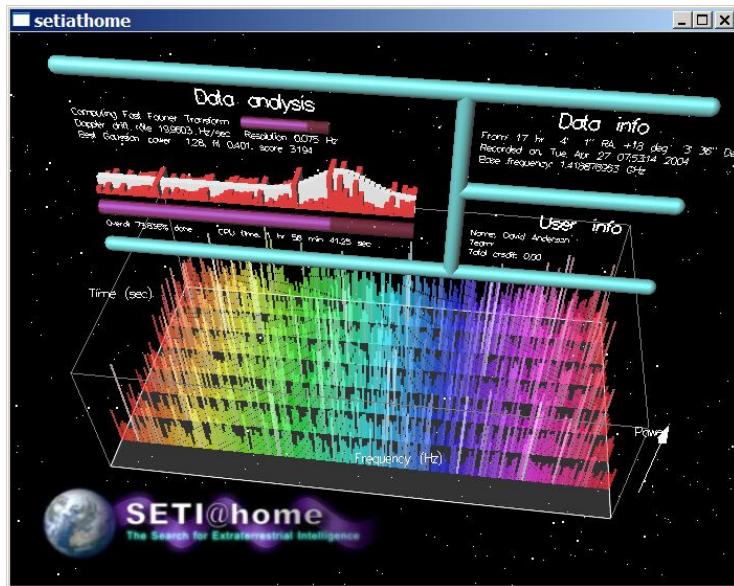


## Performance Metrics

- **Execution time**
  - Time when the last processor finishes its work
- **Speedup**
  - $(\text{time on 1CPU}) / (\text{time on P CPUs})$
- **Efficiency**
  - Speedup / P

# Seti@home-like Computing

- Idle computers can be used
  - Applications running as a **screen saver**
- Only computational-intensive application
  - No communication while computing
- Very successful
  - People are willing to share their computing power
- LHC@home
  - Testing stability of the beam (60 particles 100k loops)



# **Grid Computing**

# Grid Computing

- Who needs grid computing?
  - Particular software capabilities: modeling, simulation, etc.
  - High hardware/computing demands: processing, storage, etc.
  - Large network bandwidth: circuit provisioning to support large data transfer
- Problems, which are hard (or impossible) to solve at a single site, can be solved with the right kind of parallelization and distribution of the tasks involved
- There are two primary types of grids
  - Computational grids:
    - Open Science Grid (OSG)
    - Worldwide LHC Computing Grid (WLCG)
  - Data grids:
    - Earth System Grid (ESG)



# Requirements

- Computational grids
  - Manage a variety of computing resources
  - Select computing resources capable of running a user's job
  - Predict loads on grid resources
  - Decide about resource availability
  - Dynamic resource configuration and provisioning
- Data grids
  - Provide data virtualization service
  - Support flexible data access, filtering, and transfer mechanisms
  - Provide security and privacy mechanisms
- Grid computing environments are constructed upon three foundations
  - Coordinated resources
  - Open standard protocols and frameworks
  - Non-trivial QoS

# **Cloud Computing**

# Cloud Computing



- What is cloud computing?

- The phrase originated from the cloud symbol used to symbolize the Internet
- A model for enabling convenient, on-demand network access to a shared pool of configurable computing resources
  - e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction
- Provide computation, software, data access, and storage services
  - Do not require end-user knowledge of the physical location and configuration of the system that delivers the services

- Cloud architecture

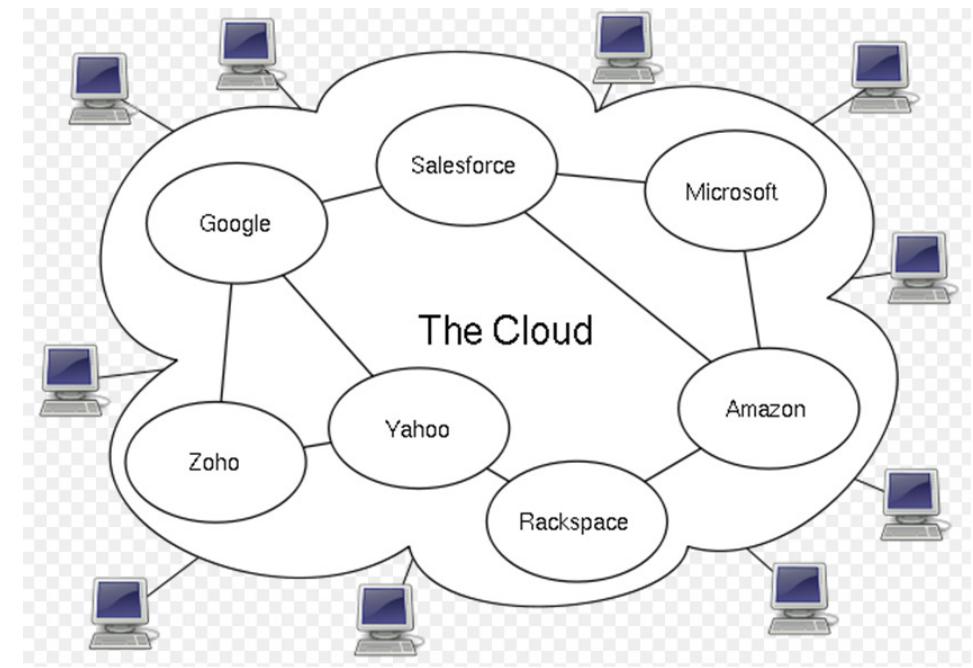
- Involve multiple cloud components communicating with each other over **application programming interfaces**, usually **web services** and **3-tier architecture**
  - Front end: seen by the user, such as a web browser
  - Back end: the cloud itself comprising computers, servers, data storage devices, etc.

# Five Layers in Cloud Computing

- **Client**: a cloud client consists of **computer hardware** and/or **computer software** that relies on cloud computing for application delivery, or that is specifically designed for delivery of cloud services
- **Application**: cloud application services or "**Software as a Service (SaaS)**" deliver **software** as a service over the Internet, eliminating the need to install and run the application on the customer's own computers and simplifying maintenance and support
- **Platform**: cloud platform services or "**Platform as a Service (PaaS)**" deliver a **computing platform** and/or **solution stack** as a service, often consuming cloud infrastructure and sustaining cloud applications
- **Infrastructure**: cloud infrastructure services, also known as "**Infrastructure as a Service (IaaS)**", delivers **computer infrastructure** – typically a **platform virtualization** environment – as a service
- **Server**: the servers layer consists of **computer hardware** and/or **computer software** products that are specifically designed for the delivery of cloud services, including multi-core processors, cloud-specific operating systems and combined offerings

# Real-life Cloud Computing Environment

- Microsoft Windows Azure
  - <https://azure.microsoft.com>
- Google Cloud
  - <https://cloud.google.com/>
- Amazon EC2
  - <https://aws.amazon.com/ec2/>
- Eucalyptus (first open-source platform for private clouds, 2008)
  - A software platform for the implementation of private cloud computing on computer clusters
  - <https://github.com/eucalyptus>
- Many others



# Managing Amazon EC2 Instances

- AWS management console
  - Web-based, most powerful
  - <http://aws.amazon.com/console/>
  - Command line tools
  - [http://aws.amazon.com/developertools/351? encoding=UTF8&jiveRedirect=1](http://aws.amazon.com/developertools/351?encoding=UTF8&jiveRedirect=1)
- Third-party UI tools
  - Example: ElasticFox browser add on

# Login AWS Management Console

S | Services ▾ | Edit ▾ jp238@njit.edu @ 934713874... ▾ Global ▾ | Help ▾

---

## Welcome

The AWS Management Console provides a graphical interface to Amazon Web Services. Learn more about how to use our services to meet your needs, or get started by selecting a service.

[Getting started guides](#)

[Reference architectures](#)

[Free Usage Tier](#)

## Set Start Page

Console Home ▾

---

 [AWS Marketplace](#)  
Find & buy software, launch with 1-Click and pay by the hour.

## Amazon Web Services

---

### Compute & Networking

-  [Direct Connect](#)  
Dedicated Network Connection to AWS
-  [EC2](#)  
Virtual Servers in the Cloud
-  [Elastic MapReduce](#)  
Managed Hadoop Framework
-  [Route 53](#)  
Scalable Domain Name System
-  [VPC](#)  
Isolated Cloud Resources

### Storage & Content Delivery

-  [CloudFront](#)  
Global Content Delivery Network
-  [Glacier](#)  
Archive Storage in the Cloud
-  [S3](#)  
Scalable Storage in the Cloud
-  [Storage Gateway](#)  
Integrates on-premises IT environments with Cloud storage

### Database

-  [DynamoDB](#)  
Predictable and Scalable NoSQL Data Store
-  [ElastiCache](#)  
In-Memory Cache
-  [RDS](#)  
Managed Relational Database Service

### Deployment & Management

-  [CloudFormation](#)  
Templated AWS Resource Creation
-  [CloudWatch](#)  
Resource & Application Monitoring
-  [Data Pipeline](#) NEW  
Orchestration for data-driven workflows
-  [Elastic Beanstalk](#)  
AWS Application Container
-  [IAM](#)  
Secure AWS Access Control

### App Services

-  [CloudSearch](#)  
Managed Search Service
-  [Elastic Transcoder](#) NEW  
Easy-to-use scalable media transcoding
-  [SES](#)  
Email Sending Service
-  [SNS](#)  
Push Notification Service
-  [SQS](#)  
Message Queue Service
-  [SWF](#)  
Workflow Service for Coordinating Application Components

## Announcements

---

- [Amazon Simple Workflow in Seven Additional Regions and Extends IAM Support](#)
- [Announcing Amazon Elastic Transcoder](#)
- [Announcing High Memory Cluster Instances for Amazon EC2](#)

[More...](#)

---

## Service Health

[Edit](#)

*Click [Edit](#) to add at least one service and at least one region to monitor.*

[Service Health Dashboard](#)

# Select AMI to create instance(s)

**Create a New Instance**

Select an option below:

**Classic Wizard**  
Launch an On-Demand or Spot instance using the classic wizard with fine-grained control over how it is launched.

**Quick Launch Wizard**  
Launch an On-Demand instance using an editable, default configuration so that you can get started in the cloud as quickly as possible.

**AWS Marketplace**  
AWS Marketplace is an online store where you can find and buy software that runs on AWS. Launch with 1-Click and pay by the hour.

**Name Your Instance:**  Pick a meaningful name, e.g. Web Server

**Choose a Key Pair:**  
Public/private key pairs allow you to securely connect to your instance after it launches.  
 **Select Existing**    **Create New**    **None**

**Choose a Launch Configuration:**

**More Amazon Machine Images NEW!**  
Search through public and AWS Marketplace AMIs or choose from your own custom AMIs.

<b>Amazon Linux AMI 2012.09</b> The Amazon Linux AMI 2012.09 is an EBS-backed, PV-GRUB image. <input checked="" type="radio"/> <b>64 bit</b> <input checked="" type="radio"/> <b>32 bit</b> <input type="radio"/> It includes Linux 3.2, AWS tools, and repository access to multiple versions of MySQL, PostgreSQL, Python, Ruby, and Tomcat. 
<b>Red Hat Enterprise Linux 6.3</b> Red Hat Enterprise Linux version 6.3, EBS-boot. <input checked="" type="radio"/> <b>64 bit</b> <input checked="" type="radio"/> <b>32 bit</b> <input type="radio"/>
<b>SUSE Linux Enterprise Server 11</b> SUSE Linux Enterprise Server 11 Service Pack 2 basic install, EBS boot with Amazon EC2 AMI Tools preinstalled; Apache 2.2, MySQL 5.0, PHP 5.3, and Ruby 1.8.7 available. <input checked="" type="radio"/> <b>64 bit</b> <input checked="" type="radio"/> <b>32 bit</b> <input type="radio"/>
<b>Ubuntu Server 12.04.1 LTS</b> Ubuntu Server 12.04.1 LTS with support available from Canonical ( <a href="http://www.ubuntu.com/cloud/services">http://www.ubuntu.com/cloud/services</a> ). <input checked="" type="radio"/> <b>64 bit</b> <input checked="" type="radio"/> <b>32 bit</b> <input type="radio"/> 

**Note:** You can customize your settings in the next step.

**Continue** 

[Submit Feedback](#) [Getting Started Guide](#)

Select instance type: <http://aws.amazon.com/ec2/instance-types/>

# Assign instance to security group(s)

- A **security group** defines firewall rules for instances
- At launch time, instance can be assigned to multiple groups
- Default group doesn't allow any network traffic
- Once an instance is running, it can't change to which security group(s) it belongs
- Can modify rules for a group at any time
- New rules automatically enforced for all running instances and instances launched in the future

# Illustration of security group

The screenshot shows the AWS EC2 Dashboard with the 'Security Groups' section selected. The left sidebar includes links for EC2 Dashboard, Events, Instances, Spot Requests, Reserved Instances, AMIs, Bundle Tasks, Volumes, Snapshots, and Network & Security. Under Network & Security, 'Security Groups' is highlighted.

The main area displays a list of EC2 Security Groups:

	Name	VPC ID	Description
<input checked="" type="checkbox"/>	test-cluster		Group for Hadoop Slaves.
<input type="checkbox"/>	test-cluster-master		Group for Hadoop Master.
<input type="checkbox"/>	default		default group
<input type="checkbox"/>	mycluster1-master		Group for Hadoop Master.
<input type="checkbox"/>	mycluster1		Group for Hadoop Slaves.
<input type="checkbox"/>	quicklaunch-1		quicklaunch-1

A specific security group, 'test-cluster', is selected, and its details are shown in the bottom half of the interface. The 'Inbound' tab is active, showing the configuration for incoming traffic:

Create a new rule: Custom TCP rule

Port range: (e.g., 80 or 49152-65535)

Source: 0.0.0.0/0 (e.g., 192.168.2.0/24, sg-47ad482e, or 1234567890/default)

Add Rule

The inbound rules table shows two entries under ICMP and two under TCP:

ICMP	Port (Service)	Source	Action
ALL	sg-7ff73516 (test-cluster-master)	Delete	
ALL	sg-73f7351a (test-cluster)	Delete	

TCP	Port (Service)	Source	Action
0 - 65535	sg-7ff73516 (test-cluster-master)	Delete	
0 - 65535	sg-73f7351a (test-cluster)	Delete	

At the bottom of the page, there are links for © 2008 - 2013, Amazon Web Services, Inc. or its affiliates. All rights reserved., Privacy Policy, Terms of Use, and Feedback.

# Launch instance

Services ▾ | Edit ▾

jp238@njit.edu @ 934713874... ▾ | N. Virginia ▾ | Help ▾

EC2 Dashboard | Events

Instances

Spot Requests

Reserved Instances

Images

AMIs

Bundle Tasks

Elastic Block Store

Volumes

Snapshots

Network & Security

Security Groups

Elastic IPs

Placement Groups

Load Balancers

Key Pairs

Network Interfaces

Launch Instance Actions ▾

Viewing: All Instances ▾ | All Instance Types ▾ | Search | 1 to 2 of 2 Instances ▾

Name	Instance	AMI ID	Root Device	Type	State	Status Checks	Alarm Status	Monitoring	Security Groups	Ke
empty	i-9ccf74ec	ami-08ba6861	instance store	m1.small	terminated		none	basic	default	te
empty	i-984df6e8	ami-1624987f	ebs	t1.micro	pending	initializing...	none	basic	quicklaunch-1	ec

No EC2 Instances selected.

Select an instance above

Feedback

# Instance IP addresses

- Public IP and DNS
  - Public addresses are reachable from the Internet
- Private IP and DNS
  - Private addresses are only reachable from within the Amazon EC2 network
- Elastic IPs
  - Static IP addresses associated with account, not specific instances
  - If instance using elastic IP fails, this address can be quickly remapped to another instance
    - DNS propagation may take a long time if remapping the name to another IP address

# Checking instance IP/DNS

The screenshot shows the AWS EC2 Instances page. The left sidebar navigation includes EC2 Dashboard, Events, Instances (selected), Spot Requests, Reserved Instances, AMIs, Bundle Tasks, Volumes, Snapshots, Security Groups, Elastic IPs, Placement Groups, Load Balancers, Key Pairs, and Network Interfaces. The main content area displays two instances:

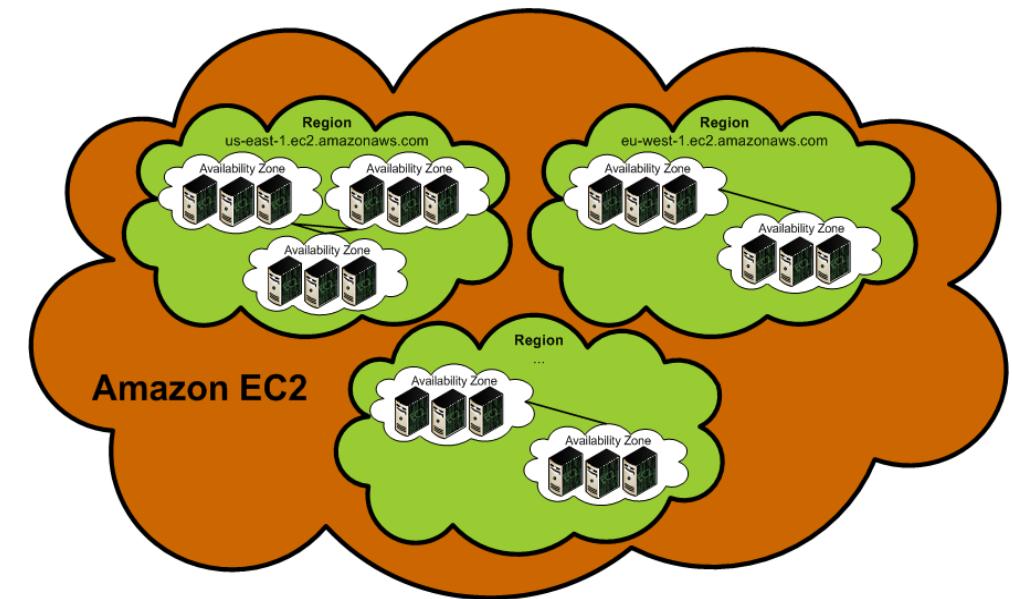
Name	Instance	AMI ID	Root Device	Type	State	Status Checks	Alarm Status	Monitoring	Security Groups	K...
empty	i-9ccf74ec	ami-08ba6861	instance store	m1.small	terminated		none	basic	default	te...
<input checked="" type="checkbox"/> empty	i-984df6e8	ami-1624987f	ebs	t1.micro	running	2/2 checks p...	none	basic	quicklaunch-1	ec...

Below the table, instance details are shown:

- Root Device Type: ebs
- Tenancy: default
- IAM Role: -
- Lifecycle: normal
- EBS Optimized: false
- Block Devices: sda1
- Network Interfaces:
  - Public DNS: ec2-23-20-83-139.compute-1.amazonaws.com
  - Private DNS: ip-10-196-102-134.ec2.internal
  - Private IPs: 10.196.102.134
- Product Codes: (empty)
- Secondary Private IPs: --
- Launch Time: 2013-01-31 13:58 EST (less than an hour)
- State Transition Reason: (empty)
- Termination Protection: Disabled

# Region and availability zones

- Regions are located in separate geographic areas (US: Virginia & California, Ireland, Singapore, etc.)
  - Each Region is completely isolated
  - Failure independence and stability
- Availability Zones are distinct locations within a Region
  - Isolated, but connected through low-latency links
  - Failure resilience
- Current pricing:  
<http://aws.amazon.com/ec2/pricing/>



# Connecting to instances

- Windows instance
  - Connect from your browser using the Java SSH Client
  - Connect using Remote Desktop
- Linux Instance
  - Connect from your browser using the Java SSH Client
  - Putty.exe/putty key generator
  - SSH command on linux machine
  - `ssh -i key.pem ec2-user@ec2-23-20-83-139.compute-1.amazonaws.com`

# Connecting to Linux instance

**Connect to an instance**

**Instance:** i-2ae8525a

► [Connect with a standalone SSH Client](#)

▼ [Connect from your browser using the Java SSH Client \(Java Required\)](#)

Enter the required information in the fields below to connect to your instance. AWS automatically detects the key pair name, and public DNS for your instance. You need to enter location and name of the .pem file containing your private key.

Public DNS ec2-23-20-250-138.compute-1.amazonaws.com

User name:

Key name: tester1

Private key path:   
Example: C:\Users\username\Downloads\tester1.pem

Save key location:  Stored in browser cache.

[Launch SSH Client](#)

[Close](#)

ec2-user@domU-12-31-39-0B-6D-96:~ [80x24]

File Edit Settings Plugins Tunnels Help

```
MindTerm home: C:\Users\Susan\Application Data\MindTerm\  
Initializing random generator, please wait...done  
Connected to server running SSH-2.0-OpenSSH_5.3  
  
Server's hostkey (ssh-rsa) fingerprint:  
openssh md5: fc:66:c6:97:30:04:04:a2:d8:c6:90:db:a0:e8:2e:15  
bubblebabble: xelofffatag-kyrak-gadol-fegek-sumez-fecol-beson-mokuv-vysis-fexax  
  
Host key not found in 'C:\Users\Susan\Application Data\MindTerm\hostkeys\key_22_ec2-23-20-250-138.compute-1.amazonaws.com.pub'  
  
_| _|_|_  
_| ( / Amazon Linux AMI  
_\_|_|_  
  
https://aws.amazon.com/amazon-linux-ami/2012.09-release-notes/  
There are 3 security update(s) out of 45 total update(s) available  
Run "sudo yum update" to apply all updates.  
[ec2-user@domU-12-31-39-0B-6D-96 ~]$ ls  
[ec2-user@domU-12-31-39-0B-6D-96 ~]$
```

# Connecting to Windows instance: get administrator password

**Console Connect - Remote Desktop Connection** Cancel 

**Instance:** i-34e55f44 **Public DNS:** ec2-54-242-92-134.compute-1.amazonaws.com

**Log in with your credentials**

Log in to your instance with your credentials:

**Public DNS:** ec2-54-242-92-134.compute-1.amazonaws.com  
**Username:** Administrator  
**Password:** [Retrieve Password](#) \*Click if you do not know your password.

You can download an RDP file for this instance which will launch Remote Desktop Connection and connect to your instance. You will need to note down your password because the Remote Desktop Connection software will open in a new window.

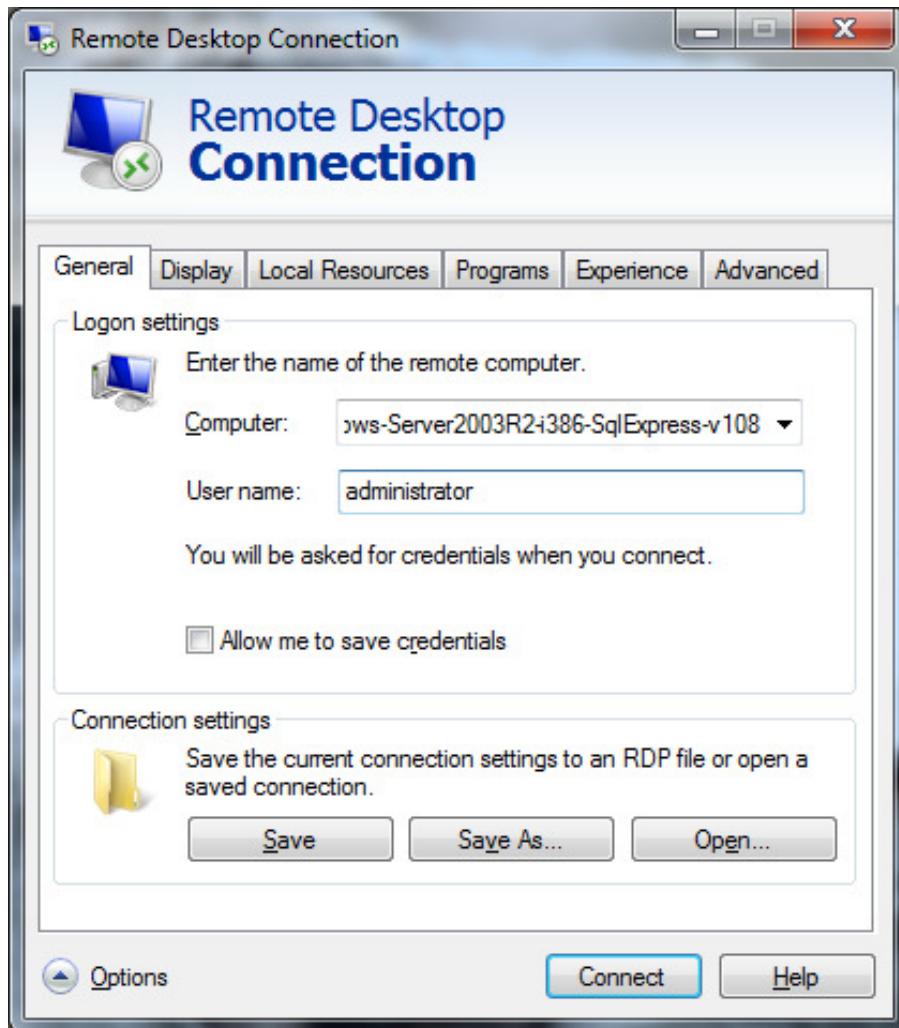
 [Download shortcut file](#)

If you need help configuring your remote desktop software, click [here](#).

► [Retrieve Windows Administrator password](#)  
► [Need help configuring your remote access software?](#)

[Close](#)

# Connecting to Windows instance



- Type in public DNS name of instance
- Use the retrieved administrator password

# IAM – Identity Access Management

- “AWS Identity and Access Management (IAM) helps to securely control access to AWS resources
- IAM is used to control who can use your AWS resources (authentication) and what resources they can use and in what ways (authorization).”

# **Mobile and Ubiquitous Computing**

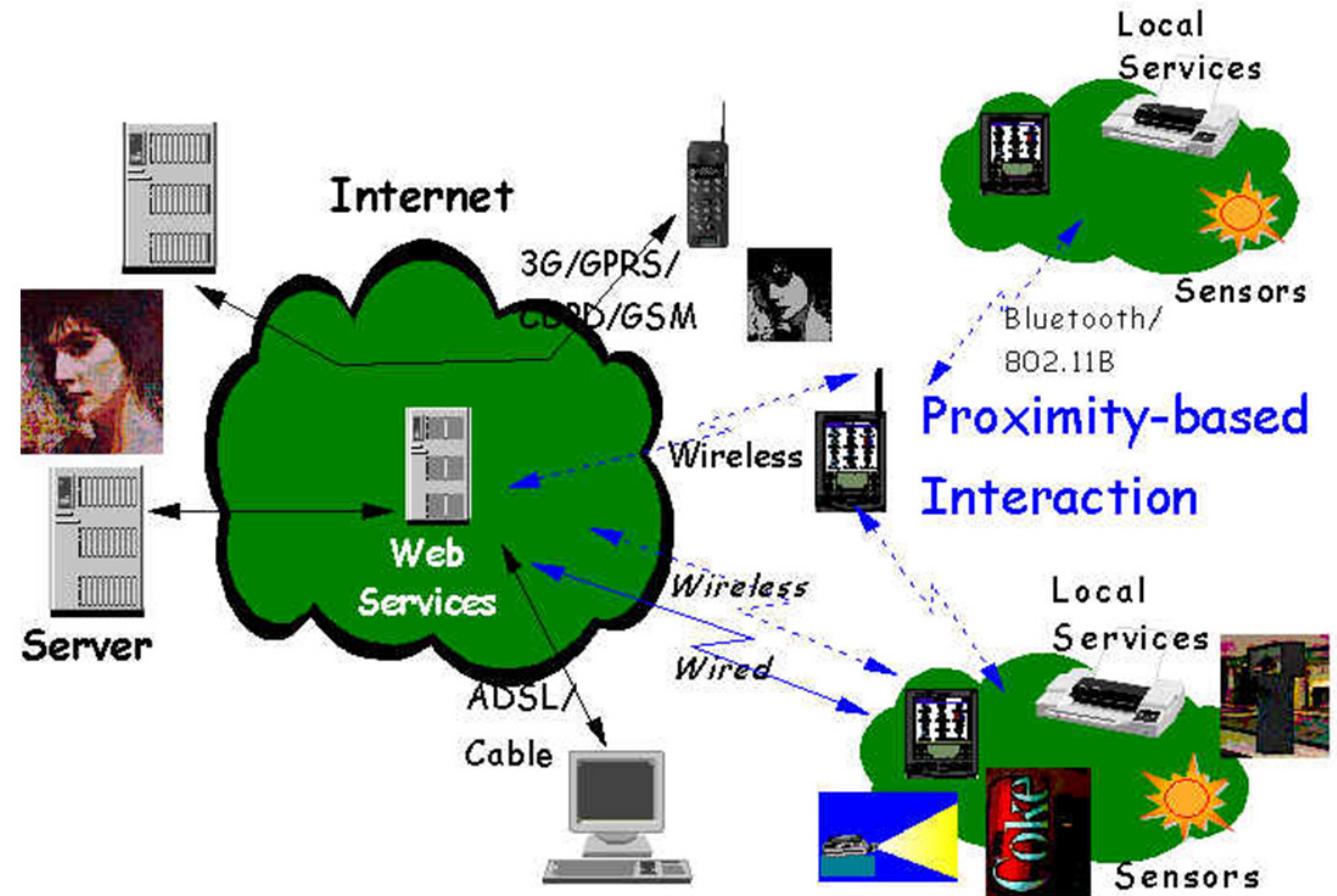


# Mobile and Ubiquitous Computing



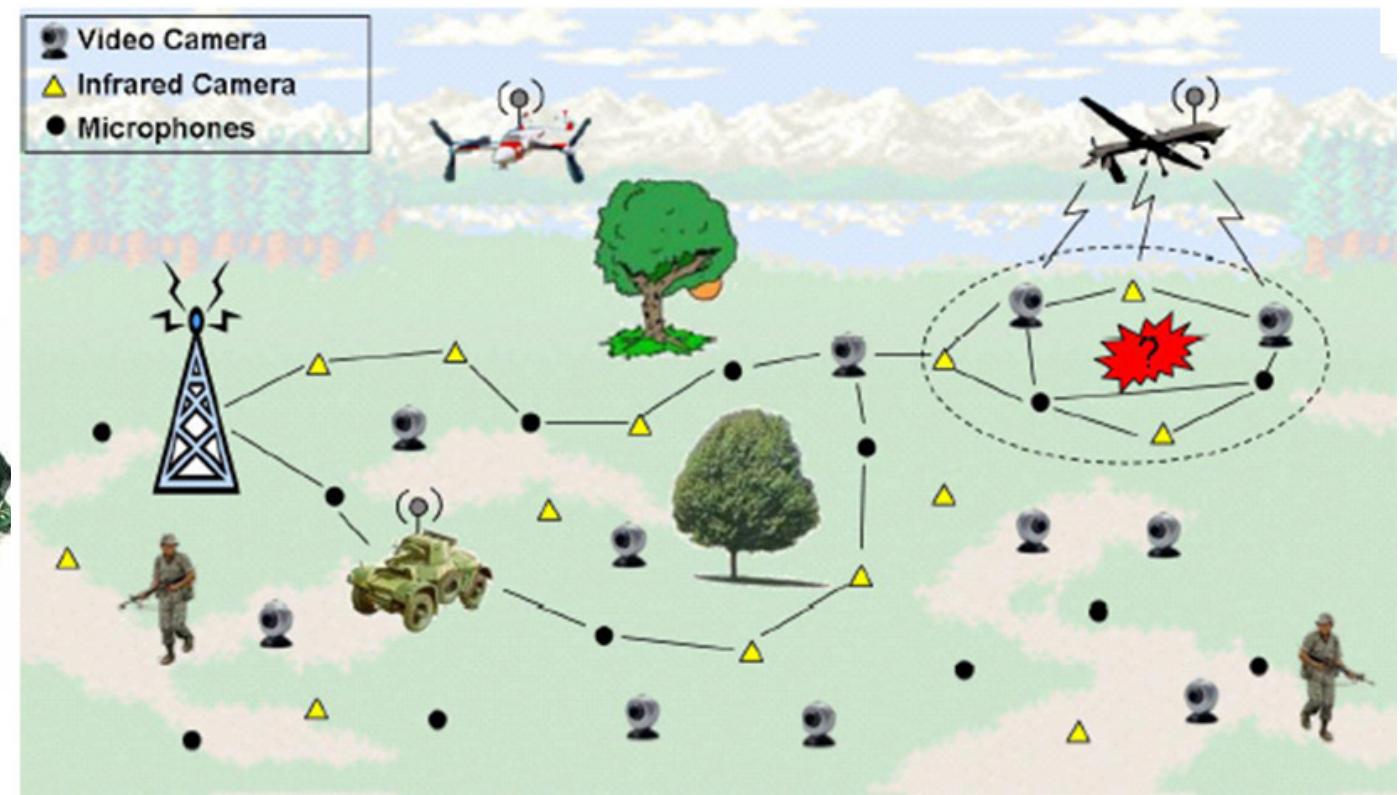
# Mobile and Ubiquitous Computing Internet of Things (IoT)

- GPRS: General Packet Radio Service
- WTP: Wireless Transport Protocol
- WAP: Wireless Application Protocol with 2 components: WML and WMLScript
- WML: Wireless Markup Language
- cHTML: Compact HTML
- HDML: Handheld Device Markup Language



# Wireless Sensor Networks

- Pervasive applications
  - Agricultural
  - Military
  - Industrial



# Ongoing Research in Our HPNC Group

## Data-intensive Computing

- Scientific Workflow Optimization: mapping, scheduling, modeling
- Various network environments HPN, grid, cloud

## High-performance Networking

- Bandwidth scheduling
- Transport control/profiling
- Control plane design

## Distributed Sensor Networks

- Deployment, routing, fusion

## Cyber Security

- Monitoring, game theory

## Visualization and Image Processing





*Thanks !* ☺

*Questions ?*