

**CS670/ECE670:  
Special topics in Architecture/Systems:  
Distributed Systems**

Shrideep Pallickara  
Computer Science  
Colorado State University

# Some interesting quotes - I

Computers in the future may weigh no more than 1.5 tons.

- *Popular Mechanics*, 1949

I think there is a world market for maybe five computers.

- *Thomas Watson*, chairman of IBM, 1943

<http://ifaq.wap.org/computers/famousquotes.html>

# Some interesting quotes - II

640K of memory should be enough for anybody.

– Bill Gates, Microsoft CEO and Chairman, 1983

There is no reason anyone would want a computer in their home.

Ken Olson, president, chairman and founder of DEC

<http://ifaq.wap.org/computers/famousquotes.html>

# I like to build things... learned more from why things fail

- NaradaBrokering: Content Distribution Network for data streams
  - High energy physics, Geosciences, Environmental modeling, Defense Applications and Commercial conferencing systems
- Granules: Lightweight cloud runtime
  - Earthquake science, epidemiological modeling

# We will be reviewing several papers in this course.

- Look at the forces that drive evolution of distributed systems
- Computational Economics
  - Challenges, pitfalls and opportunities
- Virtualization of physical machines (Xen)
- Map-Reduce
  - Granules and Hadoop
- Data Structures for Large-scale computing

# What is expected of you?

- Reports and Reviews
  - Some of these will be peer-reviewed.
- Assignments
  - Small projects, and a Final Project
- Apply concepts you learn to YOUR area
- NO Auditing/Sitting-in PLEASE
  - Discussions are based on workloads, unfair to other students

# At the end of each class ...

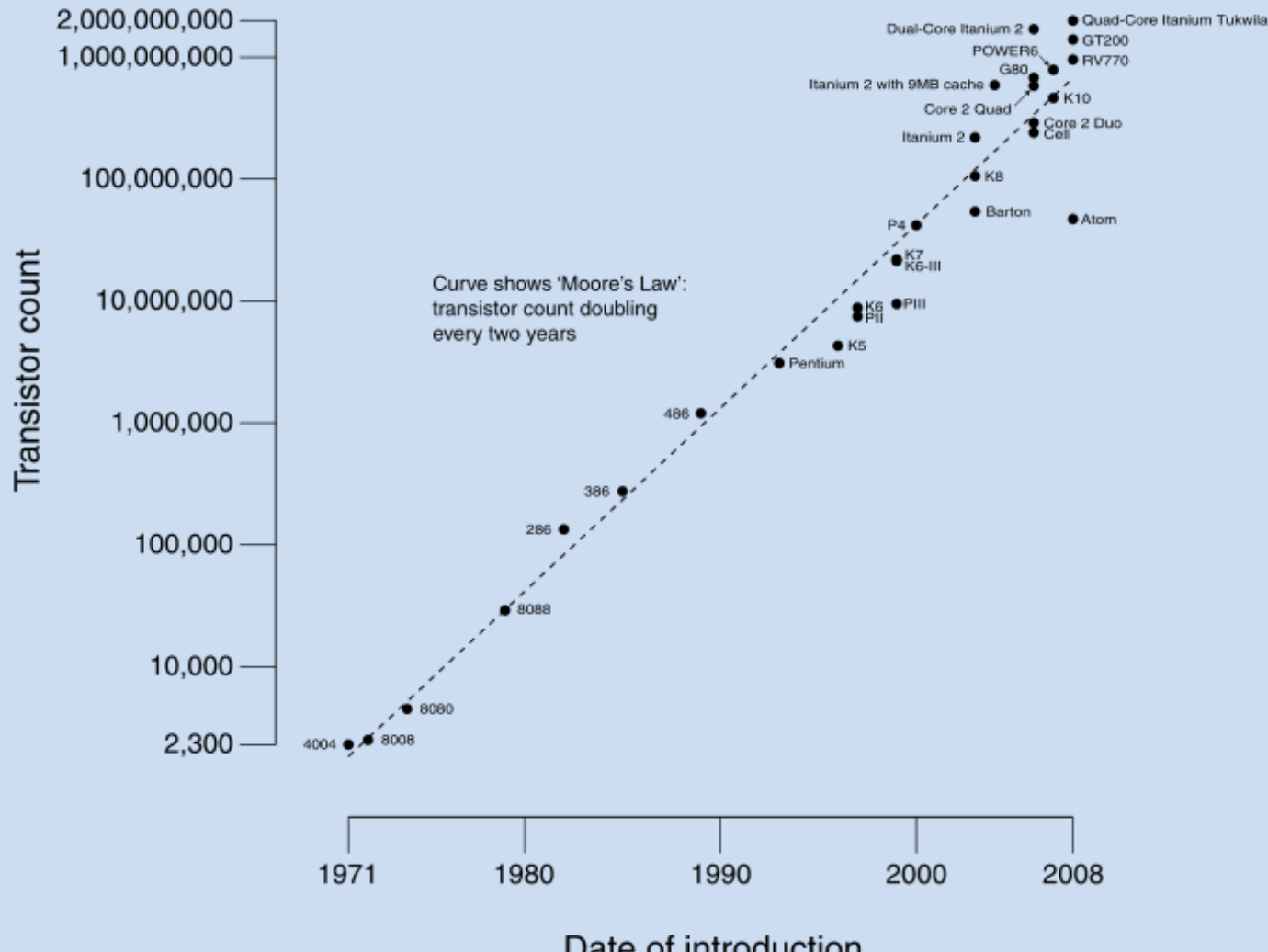
- We will take a survey
  - 3 Things that were very clear in the class
  - 3 Things that were difficult to understand
- These surveys will take 5 minutes
  - Will improve as the class goes on.

# Systems stand on the shoulders of those that have come before it

- Client-Server
- 3-tier (or N tiers)
- Distributed Objects (DCOM, CORBA, RMI)
- Message Passing based
  - Queuing, Publish/Subscribe, Peer-2-Peer
- Grid Computing
- Service Oriented Architectures (XML )
- Cloud Computing

# Moore's law: Number of transistors on an Integrated Chip doubles every 2 years

CPU Transistor Counts 1971-2008 & Moore's Law



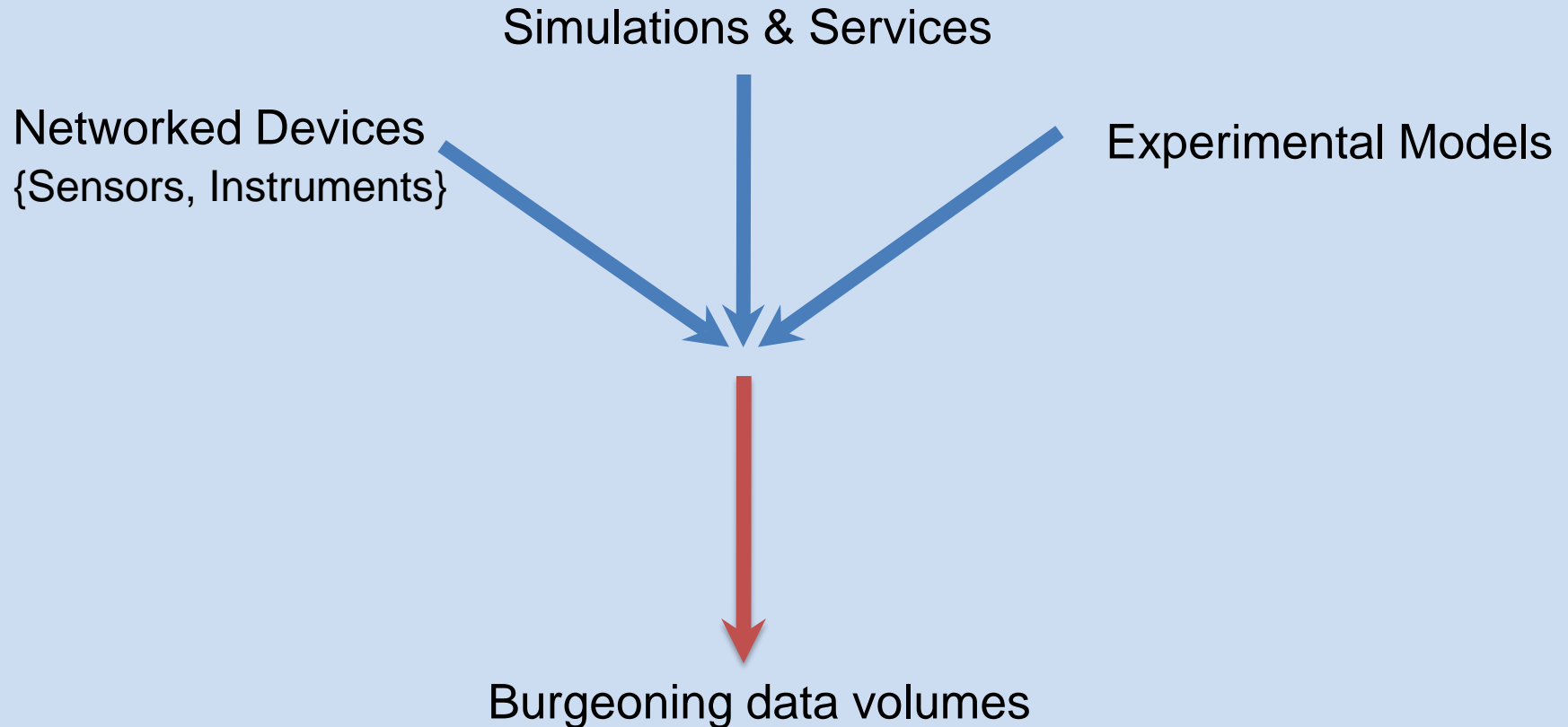
# Multi-core chips have multiple processor cores on a single chip

- ***Raison d'être***
  - Cope with heat dissipations
  - Improve Thread level parallelism
- Number of cores **doubling** every year
  - Each core also get more execution pipelines
  - Gartner Projection: 1024 cores in 2013!
- **Re-engineering** programs daunting.

# Bandwidth & Latency improvements NOT in lock-step

- Over past 20-25 years approximately
  - Bandwidth improvements: 220-1200 times
  - Latency improvements: 4-20 times.
- Ethernet 802.3 (1978)
  - 10 Mbps, Latency 3 millisecond
- Ethernet 802.3ae (2003)
  - 10,000 Mbps (1000 times), Latency 0.19 millisecond (15 times)

# Causative forces that contribute to data-volumes



# A unique confluence of factors have driven the need for cloud computing

- **DEMAND PULLS:** Process and store large data volumes
  - **Y02** 22-EB : **Y06** 161-EB : **Y10** 988-EB ~ | ZB
- **TECHNOLOGY PUSHES:** Falling hardware costs & better networks
- **RESULT:** Aggregation to scale

# Since the cloud is not monolithic it is easier to cope with flux and evolve

- Replacement and upgrades are two sides of the same coin
- Desktop: 4 GB RAM, 400 GB Disk, 50 GFLOPS & 4 cores
- 250 x Desktop = 1 TB RAM, 100 TB Disk, 6.25TFLOP and 1000 cores

# Cloud and traditional HPC systems have some fundamental differences

- One job at a time = Underutilization
  - Execution pipelines
  - IO Bound activities
- An application is the **sum of its parts**
- Cloud strategy is to **interleave** 1000s of tasks on the same resource
- Topology agnostic

# With great power comes great responsibility ...

- Literal: Data centers projected to use 4% of world's power consumption.
- Second Life avatar = 1752 kWh/year.  
Average human = 2436kWh/year
- Incredible ability to profile usage patterns

# Issues in Cloud Computing

- Scaling
- Data Management
- Privacy
- Economic Issues
- Political Issues
- Environmental issues
  - Power Consumption, disposing old computers