

# Virtualization and Cloud Computing

Sorav Bansal

# Administrivia

- Instructors: Sorav Bansal, Huzur Saran, Gautam Shroff (Tata Consultancy Services)
- Webpage:  
<http://www.cse.iitd.ernet.in/~sbansal/csl862>
- Syllabus: Lectures + Papers on webpage + Gautam Shroff's book (see webpage)
- Lecture Hours:
- Assignments: Will be posted on Webpage
- Teaching Assistant:

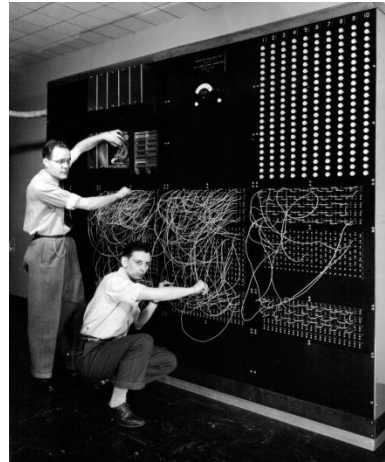
# What to Expect

- Deep Understanding of Virtualization Internals
- **Heavy Programming**
  - Pre-requisite: Must have done significant programming in OS course
  - Many new concepts and a new environment
- Understanding of Cloud Computing and Related Technologies
- Systems Research Papers
  - Expect you to read the paper before attending lecture
  - Often, the lecture will be organized as a discussion based on your understanding of the paper

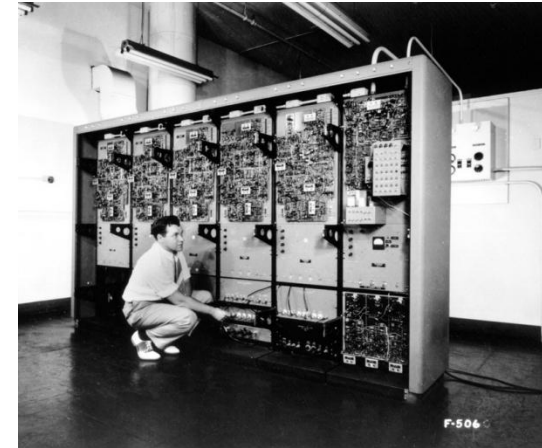
# History of Computing



Babbage Difference Engine  
1879



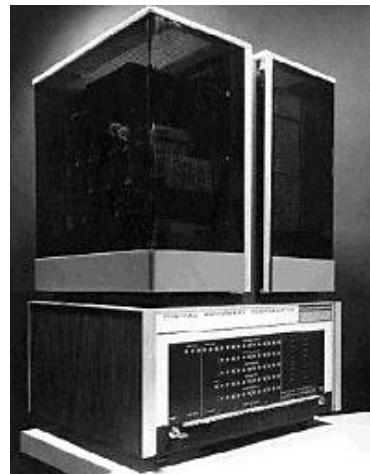
Harvard Mark-I, 1944



MIT Whirlwind, 1951



IBM 360 Mainframe  
1964, \$2.5-3 million



DEC PDP-8 **mini**computer  
1965, \$18,000



ATARI **micro**computers  
Gaming + home computing  
1979.

# History of Computing - 2



Apple Macintosh, 1984



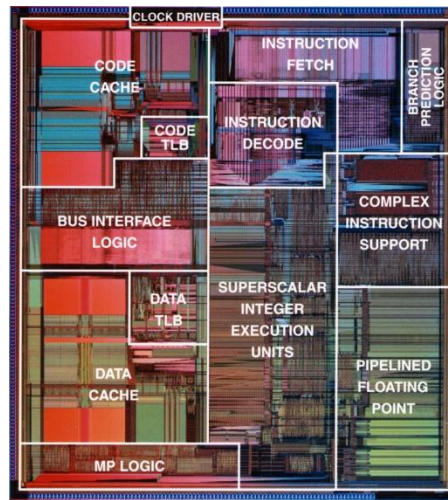
IBM PC/AT, 1984  
Intel 80286 Microprocessor



Microsoft Windows, 1985



Linus Torvalds, 1991



Intel Pentium Processor  
Diagram, 1993



Internet and Search, 1998

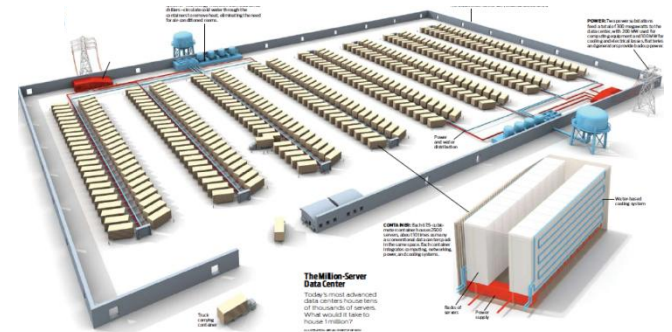
# History of Computing - 3



IPAD, 2010



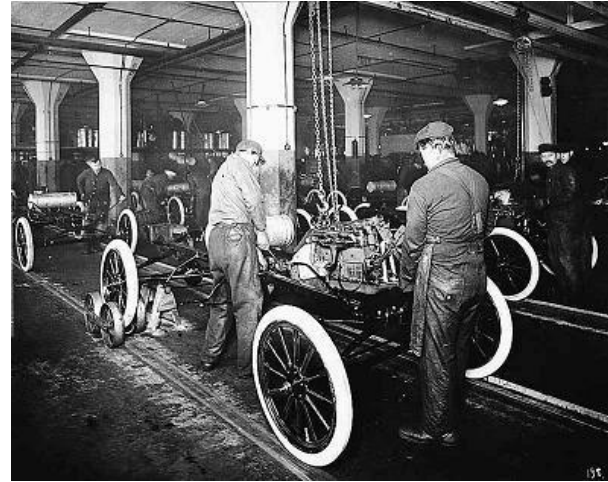
SaaS Computing



The Million Server Datacenter  
Providing IaaS



Source acknowledgement: Raghavan Subramaniam, Infosys



Source acknowledgement: Raghavan Subramaniam, Infosys



# More Buy-vs-Rent examples

- Each IIT student must buy a house in Delhi *versus* Rent a hostel room
- One workstation per student *versus* GCL lab
- Networks: Circuit-switched *versus* Packet-switched

Common Theme: Virtualization of Resource + Scheduling

# Virtualization inside an OS

- CPU → Process
- Memory → Virtual Memory
- Disk → Files
- Network Card → TCP/UDP Sockets
- Screen → Windows

NOW: Physical Machine (CPU+Mem+Disk+Net+...)  
→ Virtual Machine (VCPU+Vmem+Vdisk+Vnet+...)

# How to Virtualize

- Divide a resource in Time and/or Space
- Share
- Protect
- Schedule
- Make Pre-emptible
- Build a Cost Model

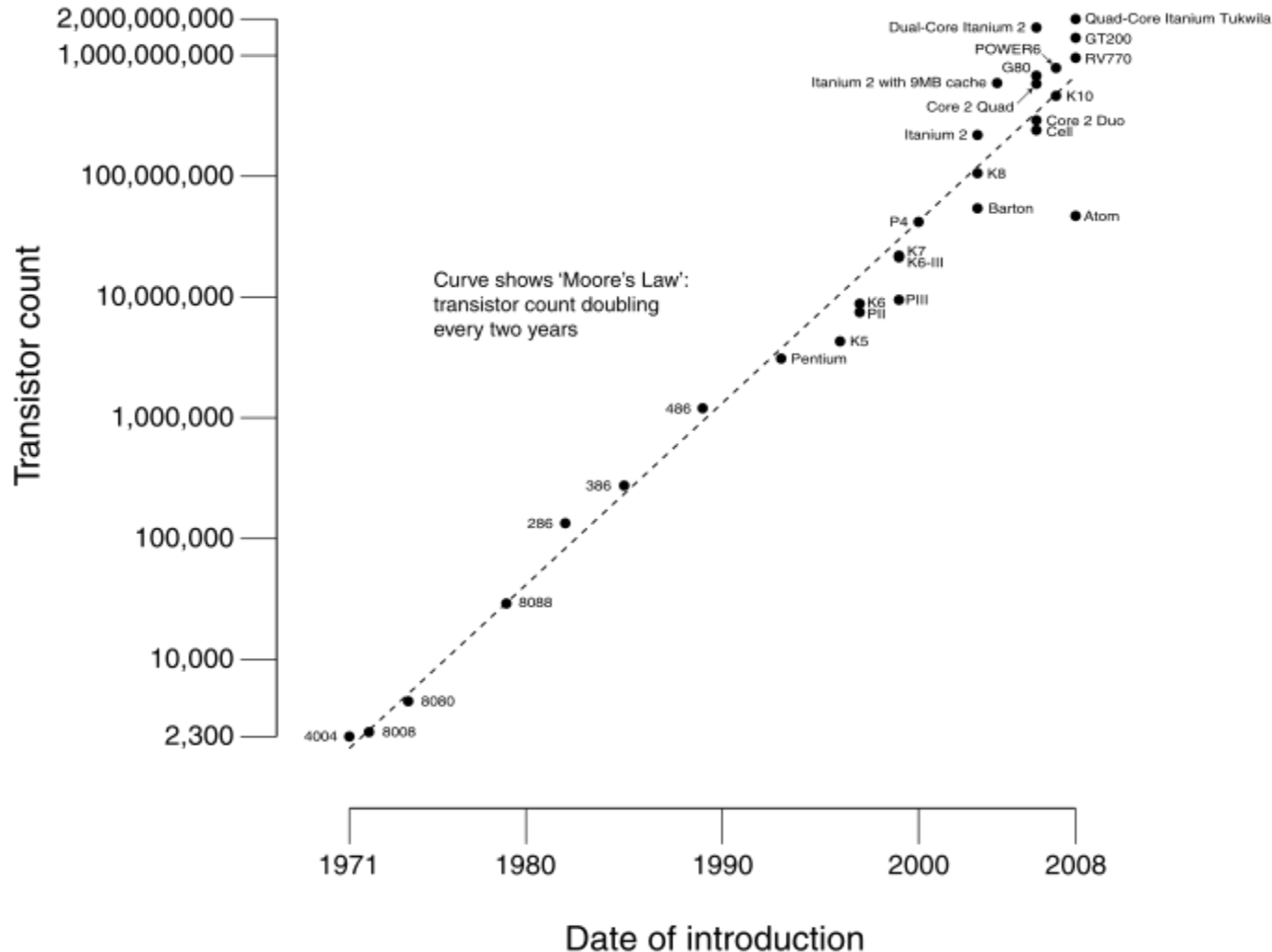
Consider examples: CPU, Memory, Disk, Cloud

# Why Virtualize?

- Usually cost
  - \$\$, power, space, maintenance personnel, ....
- Also, Hardware Trends
- Plus, much more
  - Mobility, Security, Encapsulation, ...

# Hardware Trends: Moore's Law

CPU Transistor Counts 1971-2008 & Moore's Law



Number of Transistors double every 2 years

Not expected to stop until 2015 or later!

# Hardware Trends: CPU Frequency

- Altair 8800 (1975), used an Intel 8080 CPU with a clock rate of **2 MHz**
- IBM PC (1981) : **4.77 MHz**
- Pentium (1995): **100 MHz**
- Pentium4 (2002): **3 GHz**

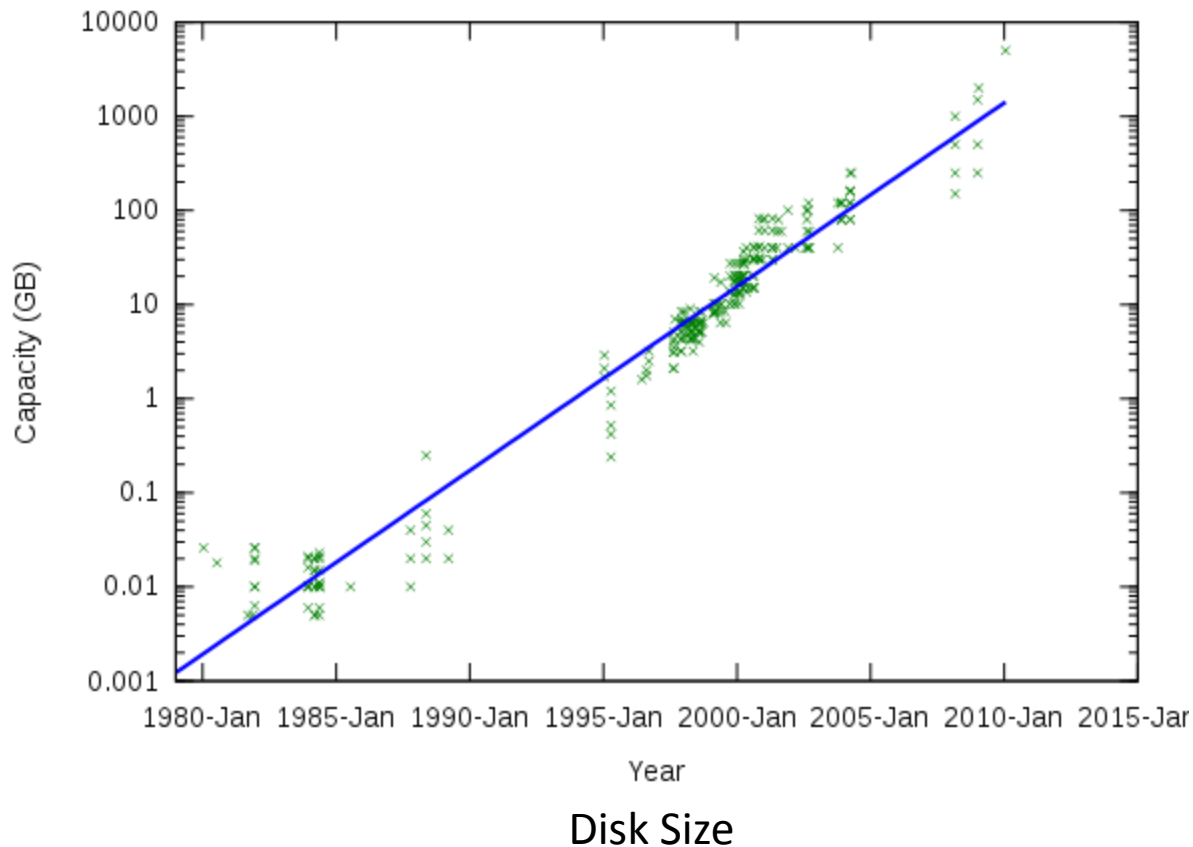
CPU Speed scaled by 1500x over 30 years

Wirth's Law (1995):

“Software is getting slower more rapidly than hardware becomes faster”

# Hardware Trends: Memory/Disk

- Memory and Disk Size have also followed exponential trends:



# Hardware Trends: Power



# 2020 IT Carbon Footprint

## IT footprints

Emissions by sub-sector, 2020

820m tons CO<sub>2</sub>

2007 Worldwide IT carbon footprint:  
2% = 830 m tons CO<sub>2</sub>  
Comparable to the global aviation industry

Expected to grow to 4% by 2020

PCs, peripherals and printers  
57%

Telecoms infrastructure and devices  
25%

360m tons CO<sub>2</sub>



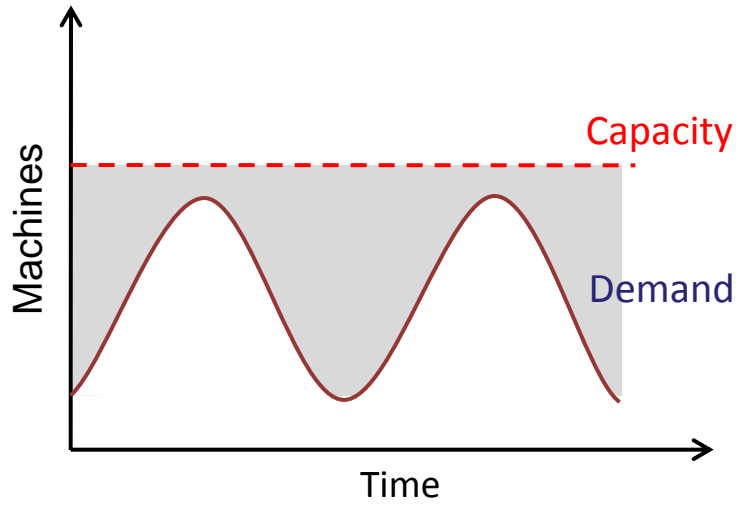
Source: The Climate Group

260m tons CO<sub>2</sub>

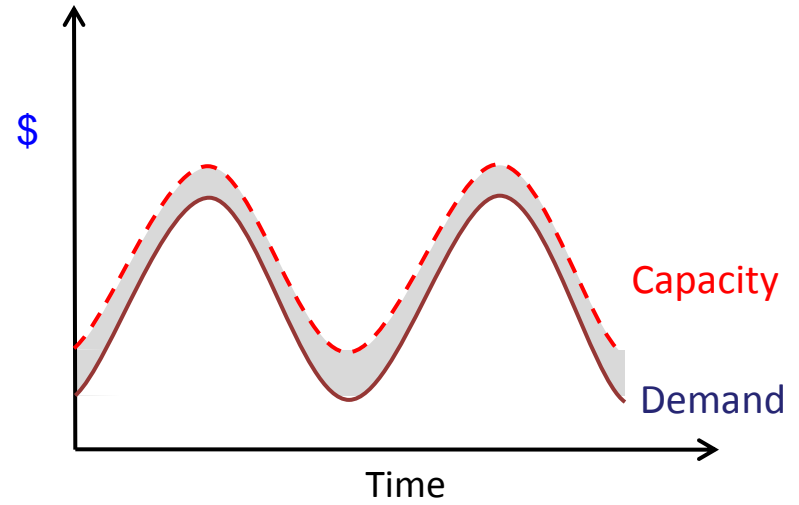
**Total emissions: 1.43bn tonnes CO<sub>2</sub> equivalent**

Source: David Patterson

# Virtualization Economics 101



“Statically provisioned”  
data center

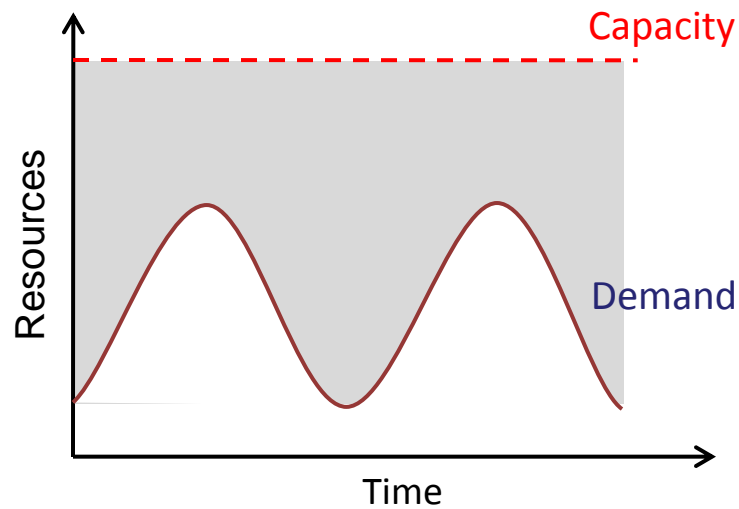


“Virtual” data center  
in the cloud

 Unused resources

# Risk of Under-utilization

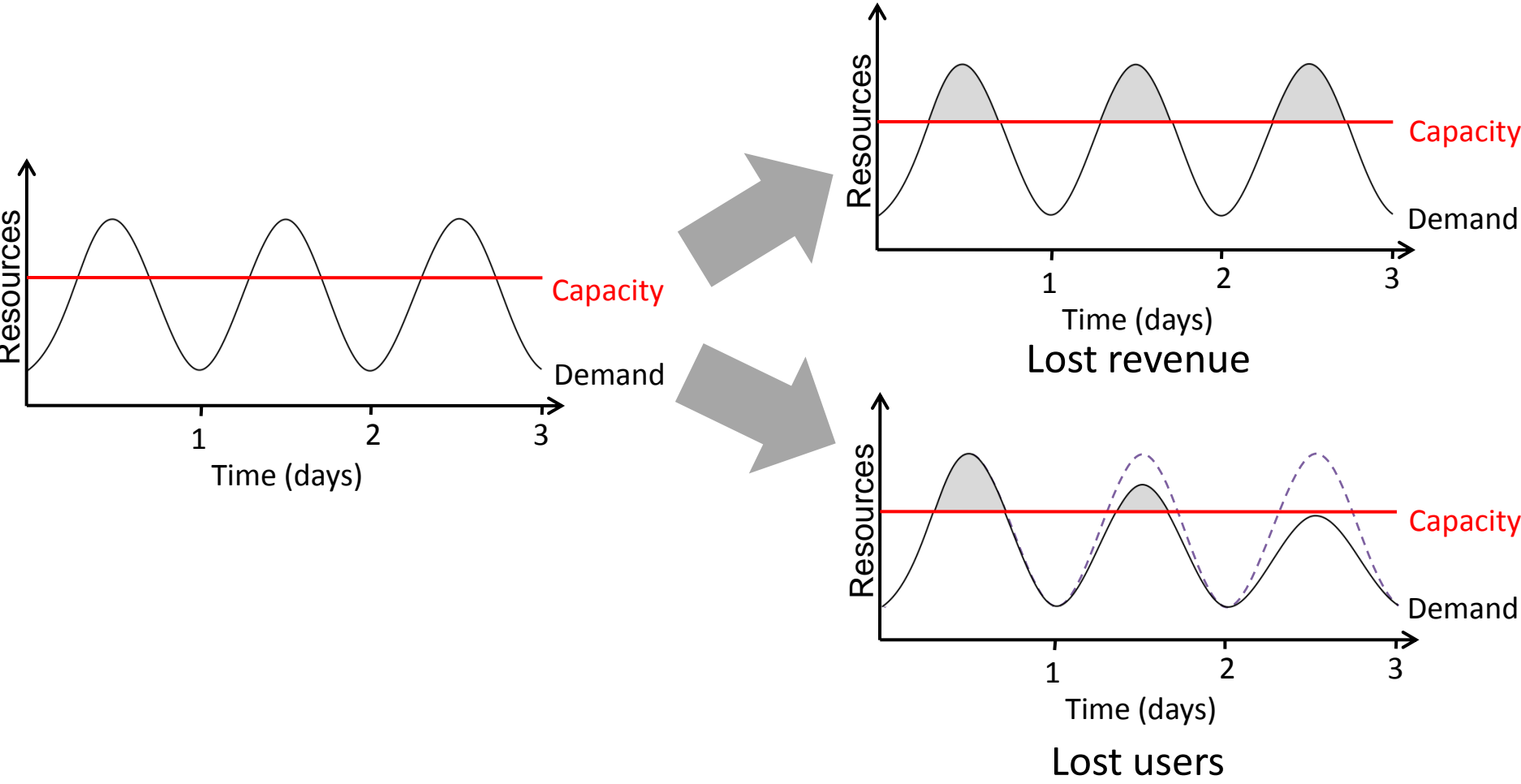
- Underutilizations occur if “peak” predictions too optimistic



Unused resources

Static data center

# Risk of Under-Provisioning



# “Risk Transfer” to Cloud

- Not (just) Capital Expense vs. Operation Expense!
- “Cost associativity”: 1,000 CPUs for 1 hour same price as 1 CPUs for 1,000 hours (@\$0.08/hour)
- *Major enabler* for SaaS startups
  - *Animoto* traffic doubled every 12 hours for 3 days when released as Facebook plug-in
  - Scaled from 50 to >3500 servers
  - *...then scaled back down*
- Gets IT gatekeepers out of the way
  - not unlike the PC revolution



# Course Outline

- Virtualization
  - Dynamic recompilation, JVM, LLVM, ...
  - Vmware, Xen, HVM, ...
  - CPU virtualization, memory virtualization, I/O device virtualization
  - Applications
- Cloud Computing
  - Parallel Programming
  - Data in the Cloud (MapReduce, BigTable, ...)
  - Present Day Clouds
  - Private/Public Clouds

# Next Lecture

- Compiling Java Just In Time
- Xen and the Art of Virtualization