CS 6240: Parallel Data Processing in MapReduce: Module 1

Mirek Riedewald

Why Parallel Processing?

- Answer 1: Big Data

How Much Information?

- Source: http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/execsum.htm
- 5 exabytes (10^{18}) of new information from print, film, optical storage in 2002
  - 37,000 times Library of Congress book collections (17M books)
- New information on paper, film, magnetic and optical media doubled between 2000 and 2003
- Information that flows through electronic channels—telephone, radio, TV, Internet—contained 18 exabytes of new information in 2002

Web 2.0

- Billions of Web pages, social networks with millions of users, millions of blogs
  - How do friends affect my reviews, purchases, choice of friends?
  - How does information spread?
  - What are “friendship patterns”?
    - Small-world phenomenon: any two individuals are likely to be connected through a short sequence of acquaintances

Facebook Statistics

- Active number of users as reported by the Associated Press in January 2013):
  - 1 million in late 2004
  - 5.5 million in late 2005
  - 12 million in late 2006
  - 20 million in 04/2007
  - 50 million in 10/2007
  - 100 million in 08/2008
  - 200 million in 04/2009
  - 400 million in 02/2010
  - 800 million in 09/2011, 1.01 billion in 09/2012, 1.06 billion in 12/2012
- As reported on Facebook's homepage in August 2011, 30 billion pieces of content were shared each month. Content includes Web links, news stories, blog posts, notes, and photo albums. The average user created 90 pieces of content per month.

Business World

- Fraudulent/criminal transactions in bank accounts, credit cards, phone calls
  - Goal: real-time detection
- Retail
  - What products are people buying together?
  - What promotions will be most effective?
- Marketing
  - Which ads should be placed for which keyword query?
  - What are the key groups of customers and what defines each group?
- Spam filtering

eScience Examples

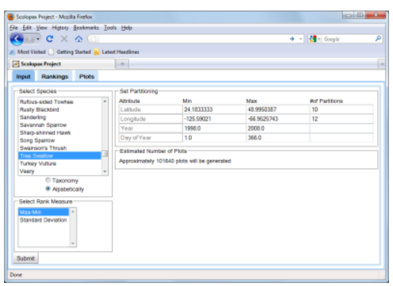
- Genome data
  - Single human genome has approx. 3 billion base pairs
- Large Hadron Collider: answer questions related to our understanding of the Universe
  - Approx. 15 petabytes of raw data per year
- Sloan Digital Sky Survey: “map the universe”
  - From 2000 to 2008, it obtained images covering more than a quarter of the sky and created 3-dimensional maps containing more than 930,000 galaxies and more than 120,000 quasars.
- DataONE: “universal access to data about life on earth and the environment”
- Cornell Lab of Ornithology
  - More than 100 million crowd-sourced bird sightings from all over the world.
  - Joining the reports based on time and location with data about climate, weather, human census information, habitat, etc. creates a rich resource for exploring the relationship between bird population trends and these factors.

Our Scolopax Project

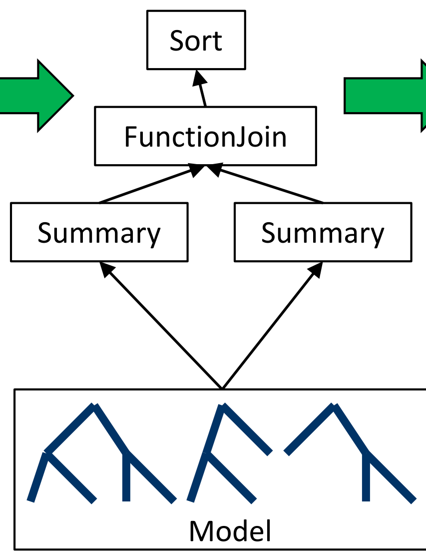
- Search for patterns in prediction models based on user preferences

  Make this as easy and fast as Web search

User-friendly query language (broad class of patterns)



Formal language (for query optimization)



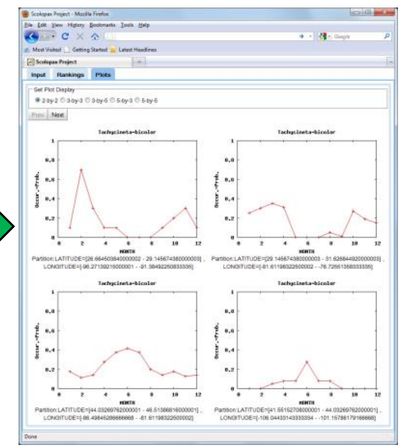
Optimizer (execution in distributed system)

Pattern ranking alg.

Function join alg.

Pattern creation alg. (low cost, parallel, approximate)

Pattern evaluation



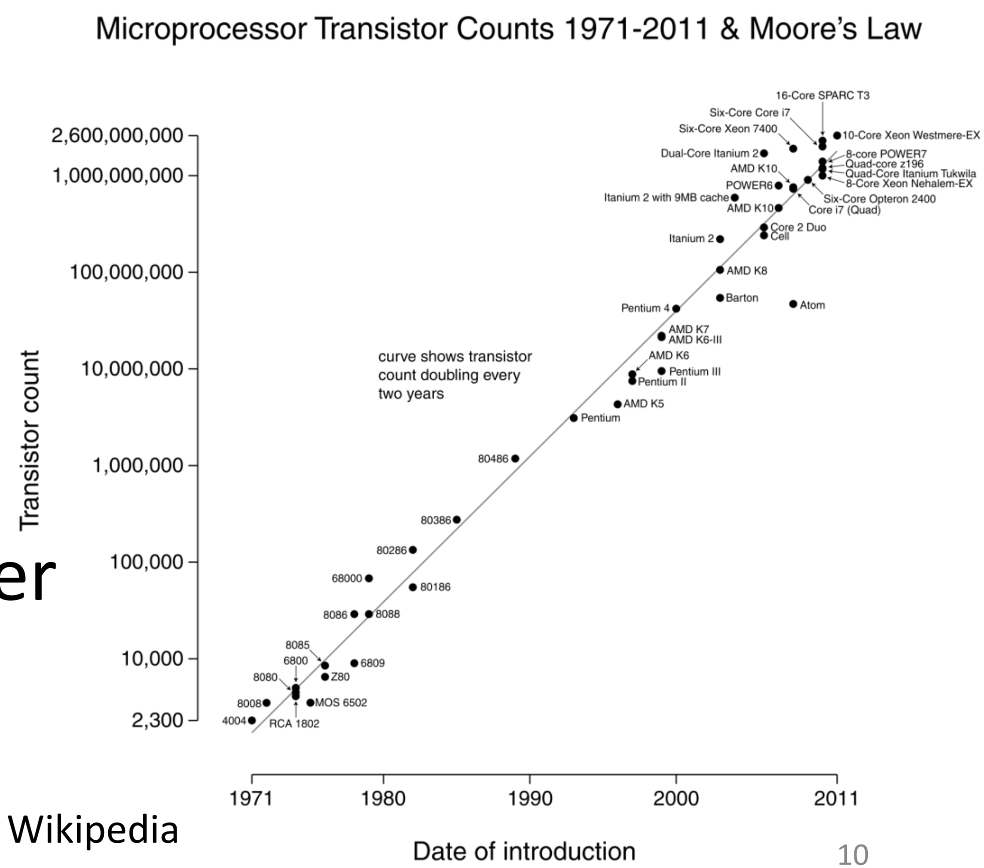
Data mining models (distributed training, evaluation, confidence computation)

Why Parallel Processing?

- Answer 1: Big Data
- Answer 2: hardware trends
  - Multi-core CPUs

The Good Old Days

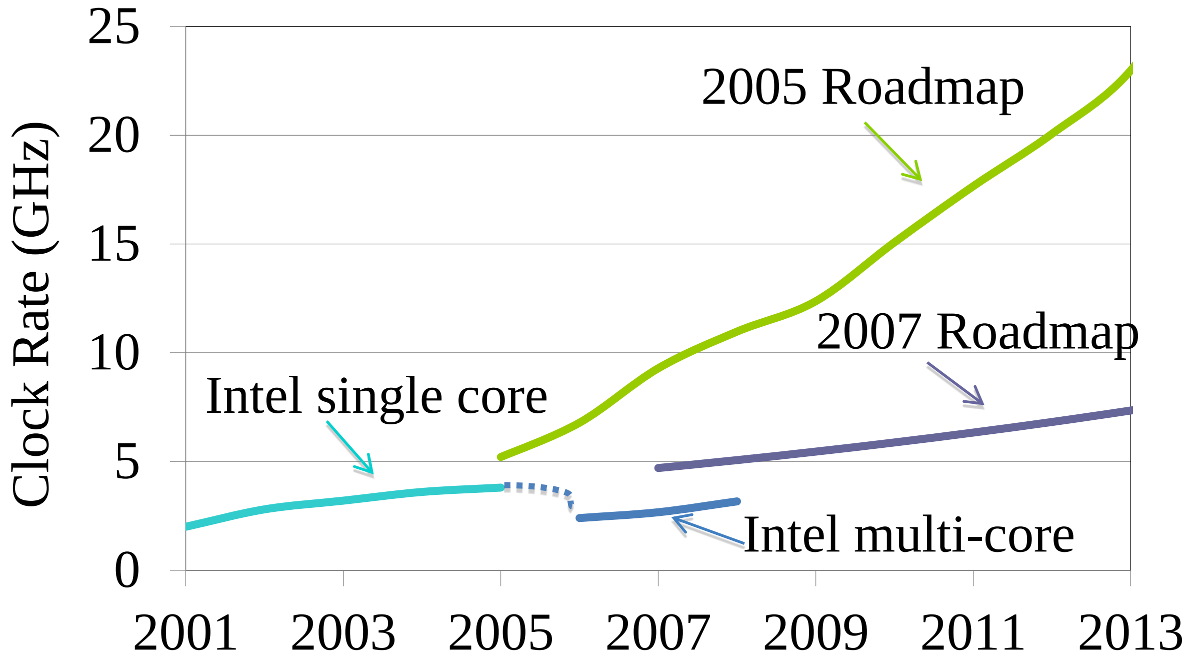
- Moore's Law: number of transistors that can be placed inexpensively on an integrated circuit doubles every ~2 years
- Computational capability improved at similar rate
  - Sequential programs became automatically faster
- Parallel computing never became mainstream



Source: Wikipedia

"New" Realities

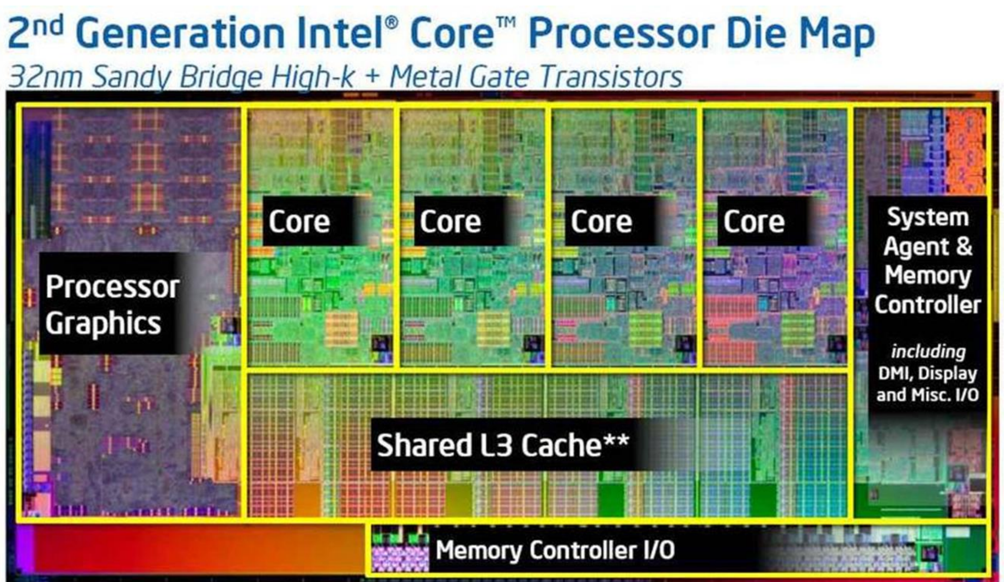
- "Party" ended around 2004
- Heat issues prevent higher clock speeds
- Clock speed remains below 4 GHz



Source:
Dave Patterson, UCB

Multi-Core CPUs

- Clock speed stagnates, but number of cores increases
  - Core = processor that shares chip with other cores
- Cores typically share some cache, memory bus, access to same main memory
- Need to keep multiple cores busy to exploit additional transistors on chip
  - Multi-threaded applications



Die	Number of Transistors (mio)	Die size with Scribe (mm2)
4+2	995	216
2+2	624	149
2+1	504	131

** Cache is shared across all 4 cores and processor graphics

Important Numbers (Source: Google's Jeff Dean @LADIS'09)

L1 cache reference	**0.5**
Branch mispredict	5
L2 cache reference	**7**
Mutex lock/unlock	25
Main memory reference	**100**
Compress 1 KB with Zippy	3,000
Send 2 KB over 1 Gbps network	20,000
Read 1 MB sequentially from memory	**250,000**
Round trip within same data center	500,000
Disk seek	**10,000,000**
Read 1 MB sequentially from disk	**20,000,000**
Send packet CA -> Holland -> CA	150,000,000

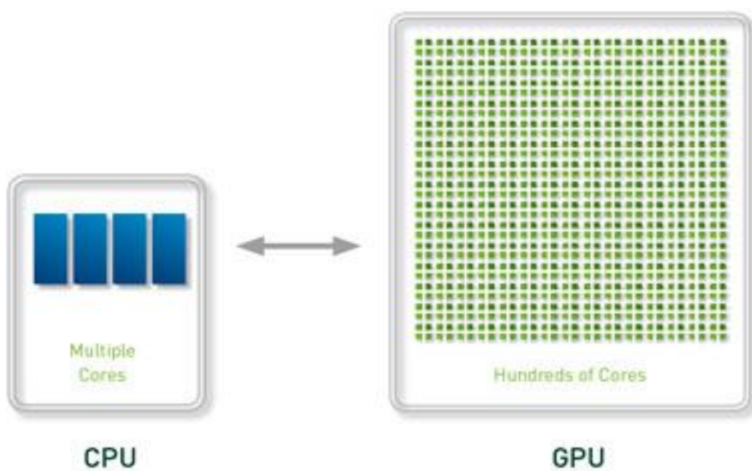
All times in ns.

Why Parallel Processing?

- Answer 1: Big Data
- Answer 2: hardware trends
  - Multi-core CPUs
  - GPU computing

GPU vs. CPU

- Optimized for massively parallel graphics processing
- Now also used for general computations
- Challenge: how to write code that keeps 100s of "simple" cores busy?
  - Example: NVIDIA developed CUDA



Source: NVIDIA

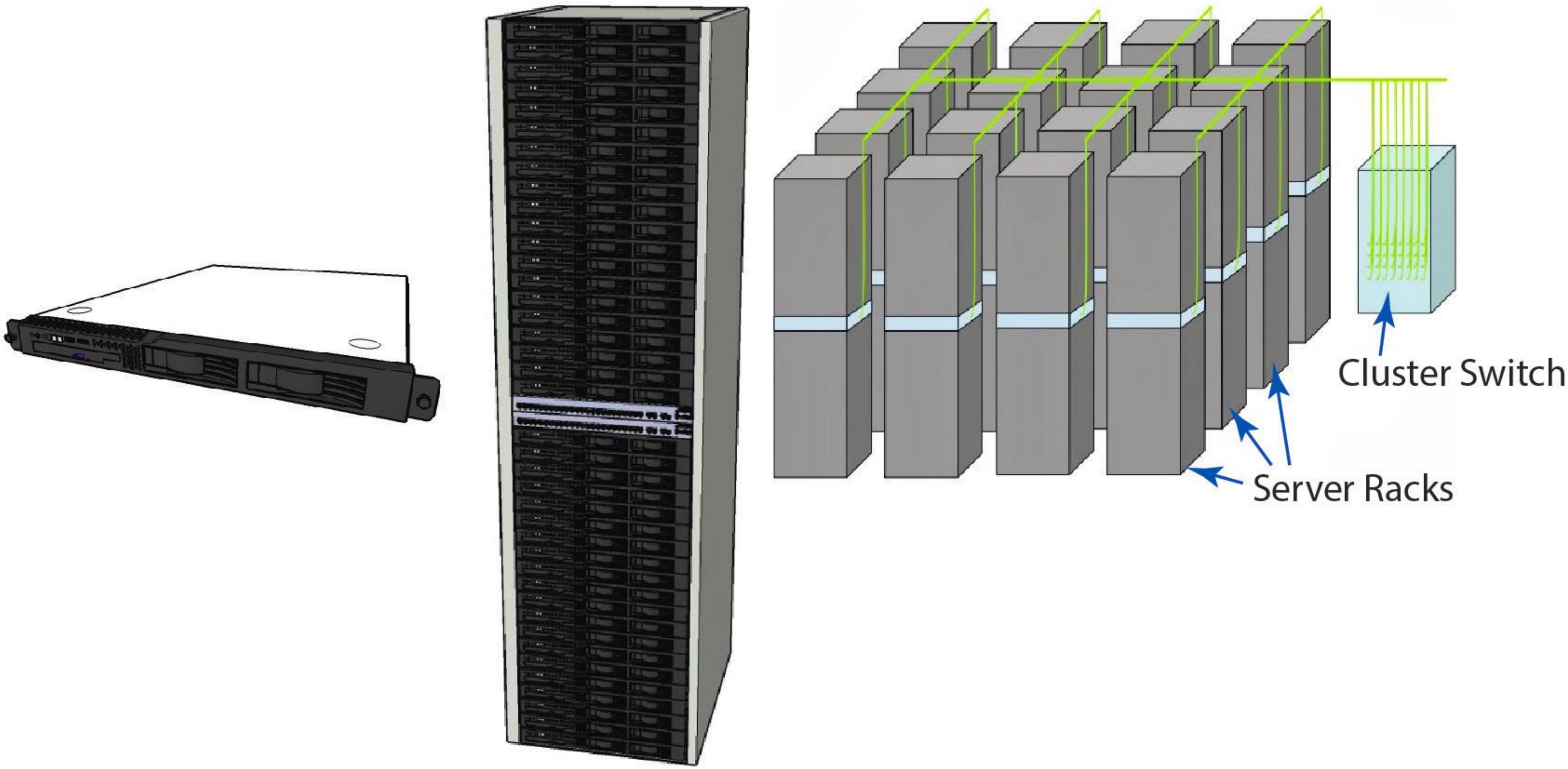
Why Parallel Processing?

- Answer 1: Big Data
- Answer 2: hardware trends
  - Multi-core CPUs
  - GPU computing
  - Data center as a computer

Warehouse-Scale Computer (WSC)

- Hundreds or thousands of commodity PCs
  - Better cost per unit of computational capability than specialized hardware due to economies of scale of commodity hardware
  - Easy to “scale out” by adding more machines
- Organized in racks in data centers
- Relatively homogenous hardware and system software platform with common system management layer

Basic Architecture

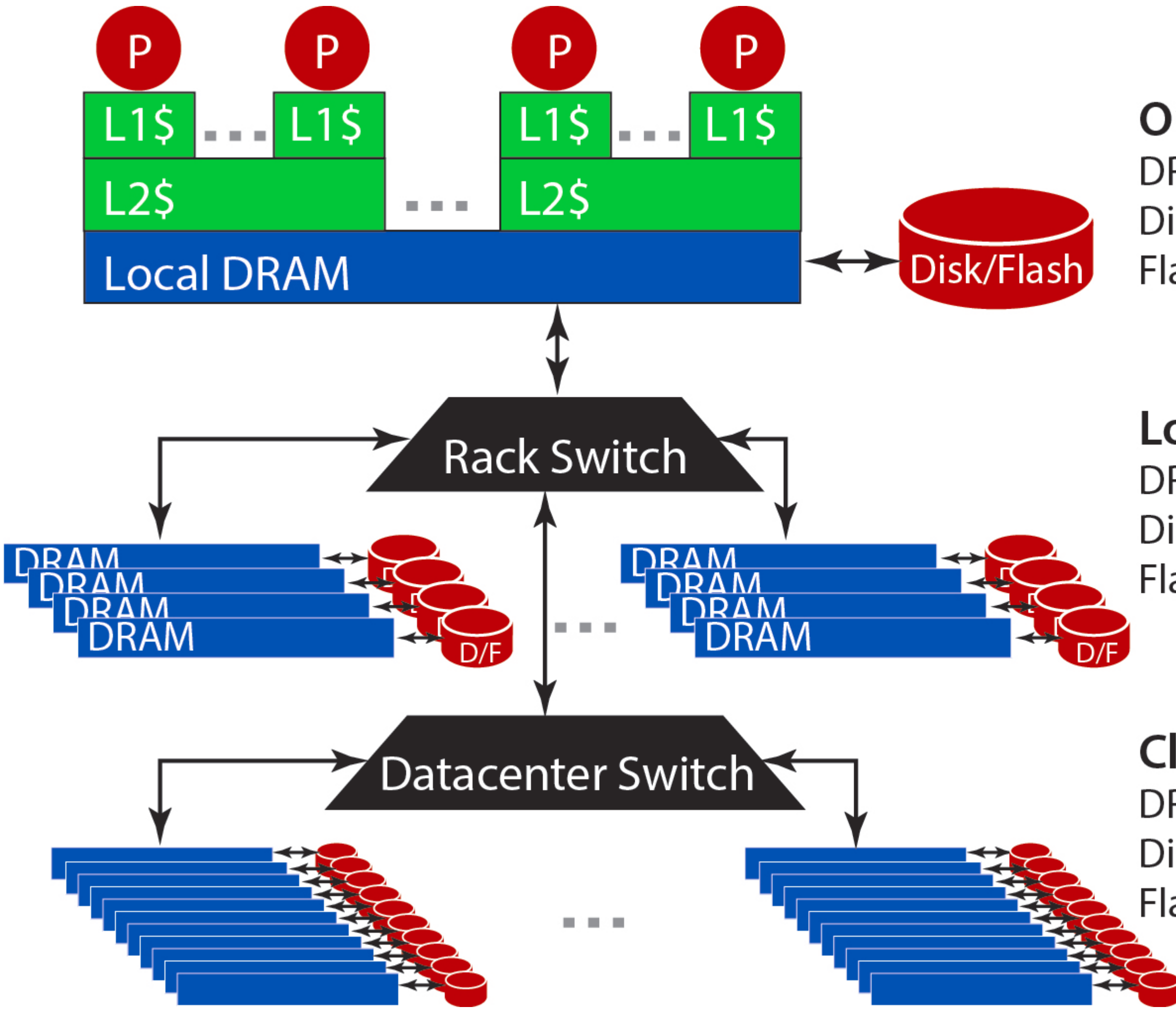


Source: Barroso, Clidaras, Holzle (2013)

Typical Specs

- Low-end servers in 1U enclosure in 7’ rack
- Rack-level switch with 1- or 10-Gbps links
- Connected by one or more cluster switches
  - Can include >10,000 servers
- Local (cheap) disks on each server
  - Managed by global distributed file system
- Might have Network Attached Storage (NAS) devices for more centralized storage solution

Storage Hierarchy



One Server

DRAM: 16 GB, 100 ns, 20 GB/s
Disk: 2T B, 10 ms, 200 MB/s
Flash: 128 GB, 100 us, 1 GB/s

Local Rack (80 servers)

DRAM: 1 TB, 300 us, 100 MB/s
Disk: 160 TB, 11 ms, 100 MB/s
Flash: 20 TB, 400 us, 100 MB/s

Cluster (30 racks)

DRAM: 30 TB, 500 us, 10 MB/s
Disk: 4.80 PB, 12 ms, 10 MB/s
Flash: 600 TB, 600 us, 10 MB/s

Source: Barroso, Clidaras, Holzle (2013)

Programming WSCs

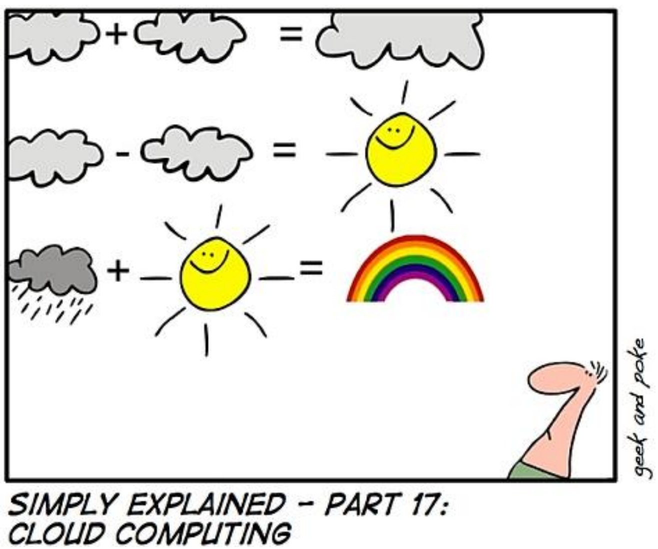
- Build cluster infrastructure and services that hide architecture complexity from developers
  - Program it like a single big computer, but avoid inefficient code
- Need easy way to keep hundreds or thousands of CPUs busy
- Handle failures transparently
  - With >1000 commodity machines, failures are the norm, not the exception
  - Developers want to focus on their application, not how to deal with failures of hardware and low-level services
- This is where MapReduce shines!

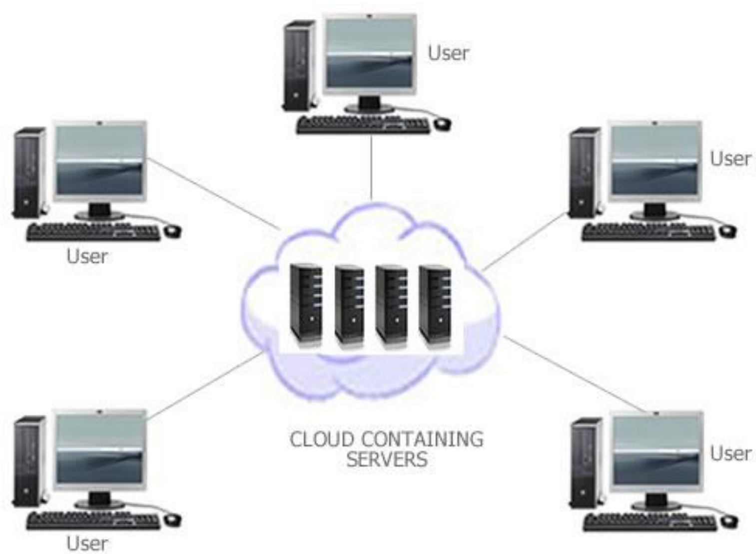
More and more, parallel data processing is performed in the “Cloud”. What is the “Cloud”?

Some content based on [Brian Hayes. Cloud Computing. *Communications of the ACM* 51(7), 2008]

The Cloud

- Basic idea: customers use virtual resources without knowing details about underlying hardware
- Goal: Move data and programs from desktop PCs and corporate server rooms into Cloud
  - Related concepts: on-demand computing, software as a service (SaaS), Internet as platform
- Starts to replace shrink-wrap software
  - MSFT Word on desktop PC vs. Google Docs or Office 365





Cloud Computing Variants

- Reserve virtual machines to create a virtual cluster; run your own application(s) there.
  - Examples: Amazon Web Services (e.g., Elastic MapReduce), IBM SmartCloud, Google App Engine, Microsoft Windows Azure, Force.com
- Connect through a Web browser to an existing application running in the Cloud.
  - Examples: Google Docs, Acrobat.com, Microsoft Office 365
- Build your own application on top of services offered by a Cloud provider.
  - Typical services: database, document management, Web design, workflow management, data analytics.
  - Examples: Salesforce.com, Microsoft Dynamics, IBM Tivoli Live

Back to the Future…

- 1960s: Hub-and-spoke configuration
  - “Dumb” terminals access a mainframe machine through phone lines
- 1980s: Client-server model
  - PCs take over functionality and data from mainframe

…or not?

- Cloud ≠ 1960's hub
  - Client can communicate with many servers at the same time
  - Servers can communicate with each other
- Still, functions migrate away to data centers
  - Storage, computing, high bandwidth, and careful resource management in core
  - End users initiate requests from fringe

Why Clouds?

- Leverage economies of scale
  - Software and hardware installation, configuration, and maintenance are easier for a homogeneous computing infrastructure
  - Single instance of MS Word cannot utilize 100 cores, but 100 instances can
- Elasticity: grow and shrink capacity on demand

Challenges

- Scalability
  - Many users, many applications
- Complex interactions
  - Client invokes programs on multiple servers, server talks to multiple clients
- Browser is limited compared to traditional OS
  - Limited functionality
  - Fewer development tools

More Challenges

- Heterogeneous environment
  - Database backend with SQL
  - JavaScript, HTML at client
  - Server app written in PHP, Java, Python
  - Information exchanged as XML
- New role for open source movement?
  - Open source word processor vs. running a service

Probably the Biggest Hurdles in Practice Are…

- Privacy, security, reliability
  - What if the service is not accessible?
  - Who owns the data?
  - Lose access to data if bill not paid?
  - Guarantee that deleted documents are really gone?
  - Data protection, e.g., against leak to third party or against government access?